A Relative Adequacy Framework for Multi-model Management in Design Optimization

Ahmed H. Bayoumy

PhD Candidate
Department of Mechanical Engineering
McGill University
Montreal, Quebec H3A 0C3, Canada
Email: ahmed.bayoumy@mail.mcgill.ca

Michael Kokkolaras

Associate Professor
Department of Mechanical Engineering
McGill University
Montreal, Quebec H3A 0C3, Canada
Email: michael.kokkolaras@mcgill.ca

We consider the problem of selecting among different physics-based computational models of varying, and oftentimes not assessed, fidelity for evaluating the objective and constraint functions in numerical design optimization. Typically, higher-fidelity models are associated with higher computational cost. Therefore, it is desirable to employ them only when necessary. We introduce a relative adequacy framework that aims at determining whether lower-fidelity models (that are typically associated with lower computational cost) can be used in certain areas of the design space as the latter is being explored during the optimization process. We implement our approach by means of a trust-region management framework that utilizes the mesh adaptive direct search derivative-free optimization algorithm. We demonstrate the link between feasibility and fidelity and the key features of the proposed approach using two design optimization examples: a cantilever flexible beam subject to high accelerations and an airfoil in transonic flow conditions.

1 Introduction

Numerical engineering design optimization requires computational models to predict system behavior in large and multi-dimensional design spaces. But what constitutes an adequate model? How can we characterize it? How do we choose among different models during the optimization process depending on their fidelity level in different areas of the design space? According to Gross, fidelity is the degree to which a model or simulation reproduces the state and behavior of a real-world object, feature or condition; therefore, fidelity is a measure of the realism of a model or simulation [1]. In the literature, the term "multi-fidelity" is

used as an adjective for several terms: physics [2, 3], surrogates [4], approximations [5], analysis [6], optimization [7], mapping [8] and multidisciplinary design optimization (MDO) [9]. Therefore, the use of the term can differ substantially, which can lead to misconceptions and inappropriate methods for multi-model management. The objective of this work is to develop a framework for managing the use of physics-based models of varying fidelity regardless of their disciplinary context. The proposed methodology will support decisions related to which model(s), and at which computational cost, should be used during the design exploration of the optimization process.

Typically, when model fidelity increases, so does computational cost as argued in [10]. However, this is not a law, and there can be exceptions. In this work, we may sometimes expect that a high-fidelity model has higher computational cost than a low-fidelity model, but our methodology does not depend on such assumptions. Existing multi-fidelity optimization (MFO) approaches attempt to calibrate lowfidelity models or replace low-fidelity analysis results using data from higher fidelity analyses [11]. Moreover, the performance of multidisciplinary systems is obtained not only by conducting the analysis of each discipline but also by capturing and accounting for their interactions. Therefore, quantifying system interactions using high-fidelity (HF) models is a computationally intensive process, because all disciplinary models have to be repeatedly exercised at each iteration until they achieve multidisciplinary convergence. A possible remedy is to use low-fidelity (LF) models in lieu of HF ones as long as adequate accuracy is maintained. Furthermore, intermediate-fidelity models may be possible to obtain. But how should multiple models be managed to achieve convergent solutions for single and multidisciplinary analysis? Before we discuss how to manage multiple models, we have to define how to assign fidelity levels to the available models.

The term multi-fidelity models (MFM) is used in the literature to denote using more than one model to predict the same phenomenon. To the best of our knowledge, the most recent literature reviews on the MFO topic are the ones of Fernández-Godino et al. [12] and Peherstorfer et al. [10]. Fernández-Godino et al. classified MFM into two categories: multi-fidelity surrogate models (MFSM) and multifidelity hierarchical models (MFHM). In both of these, the HF model is used as the reference to validate, calibrate, or replace lower-fidelity models or surrogate models. However, the adequacy level of the HF model may change at different points of the design space. Consequently, considering it as an absolute criterion for model selection may mislead the optimization process as the latter explores the design space. Peherstorfer et al. concluded that relative relationships among models may be more effective than the hierarchy assumed by existing MF methods, and recommended such an approach for future developments of MFO frameworks [10].

Peherstorfer et al. characterize MFM approaches as multi-fidelity management strategies that define how different models are employed during the optimization process and how outputs from different models may be combined [10]. They also classify MFM methods according to adaptation, fusion, and filtering. Adaptation enhances the output of the LF model with information from the HF model while the computation proceeds and the surrogate model is updated at each iteration. Methods based on fusion evaluate LF models and HF models and then combine information. An example of fusion is the co-Kriging method [6]. Filtering methods invoke the HF model following the evaluation of an LFM filter. That is, the HF model is used only if the LF model is inaccurate [6]. The authors include MFHM in the adaptation and filtering management strategies.

Tables 1 and 2 list works that use MFHM, distinguishing which of them was used as LF model and which as HF model, in two different disciplines (structural mechanics and fluid mechanics). As shown in these two tables, differ-

Table 1: Papers in structural mechanics organized by level of fidelity of models used (Lin: Linear, Non: Nonlinear, M: Single-fidelity model)

Reference	Discretiz	zation	Linearity		
Reference	Coarse	Fine	Lin	Non	
[13], [14]	LF	HF	LF	HF	
[15],	LF	HF	-	M	
[16], [17], [9]	LF	HF	M	-	
[18], [19], [20]	LF	HF	LF	HF	

ent combinations may result in changing the fidelity level of the same model. Therefore, the fidelity level of a model is

Table 2: Papers in fluid mechanics organized by level of fidelity of models used (AN: Analytical, EM: Empirical, LI: Linear, PF: Potential flow, EU: Euler, RANS: Reynolds-averaged Navier Stokes)

Reference	AN	EM	LI	PF	EU	RANS
[21], [22]	LF		HF			
[7], [23]		LF	HF			
[15]		LF				HF
[19]			LF		HF	
[9]			LF			HF
[24]				LF		HF
[3]					LF	HF

relative to the other available models.

Fernández-Godino et al. stated in their review that there is no clear relationship between the computational cost ratio of the used models to HF models and the computational cost ratio of the MFO process to the HF optimization process [12]. Figure 1 depicts the cost ratio of the MFO process to the optimization process using high-fidelity reference models (RFO) vs. the cost ratio of a single analysis of the low-fidelity model (LFA) to a single analysis of the high-fidelity reference model (RFA) for 18 papers that perform optimization using MFM. It shows that employing multiple models

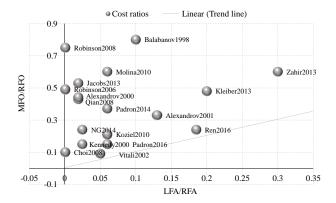


Fig. 1: Cost ratio of the optimization process using MFM (MFO) to the optimization process using high (reference)-fidelity models (RFO) vs. cost ratio of a single analysis of the low-fidelity model (LFA) to a single analysis of the high(reference)-fidelity model (RFA); reproduced from [12], which uses data reported in [2, 5, 25–35, 17, 36–39]

with low (LFA/RFA) cost ratio in the optimization process does not necessarily yield a big reduction in (MFO/RFO) cost ratio. Moreover, these ratios may change for different applications or optimization problem formulations even if the models stay the same [12].

The various types of models (theoretical, mathematical, and computational) and errors (observational, modeling, and

discretization) involved in the verification and validation process are presented in [40]. In our work, we refer to discrepancies among models as inadequacies and use the hierarchy shown in Figure 2. Different approximation models A_i may be available for a physical phenomenon. These models can be solved using different numerical methods N_{i_j} . Finally, parameter values $\mathbf{p}_{i_{j_k}}$ pertaining to the numerical method used define a particular model instantiation M_{i_k} .

Model inadequacy represents a model's inability to predict the true behavior of a physical phenomenon [41]. An adequacy level is assigned to each model instantiation based on absolute and relative measures. These measures require a reference model to be selected based on the highest computational cost among available models. We use the term absolute model inadequacy (AMI) to describe the difference between a model's prediction and the reference model's prediction. If more than two models are available, we use the term relative model inadequacy (RMI) to describe the difference in the output between two different models other than the reference model.

Model selection is then based on the capability to attain feasible solutions, maintaining adequate accuracy, and reducing computational cost. Oberkampf et al. discussed the relationship between prediction and validation [42]. Our methodology adopts their point of view and applies it to our proposed relative adequacy framework (RAF).

2 Proposed Methodology

We will illustrate our approach by means of a running example. In this example, computational models are used to compute the dynamic free response of a cantilevered three-dimensional beam under the effect of gravitational acceleration. The beam is fixed at one end as shown in Figure 3. Let us first consider modeling this problem using two approximation theories: a Timoshenko beam element with 6 degrees of freedom (DOF) and an absolute nodal coordinate formulation (ANCF) for a beam element with 12 DOF. The derived computational models are then exercised using the parameter values listed in Table 3, i.e., for varying gravitational acceleration. Initially, the beam is horizontal and at rest.

Table 3: Flexible beam dimensions and material properties

E (Pa)	$g_v (\text{m/s}^2)$	L (m)	$C(m^2)$	$\rho \left(kg/m^{3}\right)$	$I(\mathrm{m}^4)$	$l\left(\mathbf{m}\right)$
1e10	[-10,-100]	1	0.0001	7500	8.3e-10	0.01

As can be seen in Figure 4, the free response of the beam's tip changes smoothly until $g_{\nu} = 40 \text{m/s}^2$; then the tip starts to suffer from high-frequency fluctuations. ANCF captures the high-frequency modes and large displacements successfully. However, the Timoshenko beam model fails to predict the response accurately throughout the parameter space. The inadequacy between the two models increases with increasing acceleration.

Furthermore, Figure 5 shows how the Timoshenko beam model cannot converge to attainable solutions for gravity accelerations higher than 50m/s^2 . Consequently, if only the Timoshenko beam model is used, inaccurate predictions (if at all attainable) will misguide the optimization process. On the other hand, ANCF has high computational cost; however, it has similar fidelity with the Timoshenko beam model for predicting the dominant frequency within the acceleration range $g_v \in [-10, -20]\text{m/s}^2$ as depicted in Figure 6.

2.1 Relative Adequacy Framework

Let us consider the general constrained optimization problem

$$\min_{\mathbf{d} \in \mathbb{R}^n} f(\mathbf{d})$$
subject to $g_u(\mathbf{d}) \le 0, \ u = 1, 2, \dots, U.$

For example, suppose that we have one objective function and one constraint. When a numerical method N_{ij} is used with parameter values \mathbf{p}_{ij_k} to evaluate a function f, the corresponding function value is denoted by f_{ij_k} . Given available models n_m , a reference model is chosen and assigned the new index naught (M_0) . New indices are also assigned to the remaining n_m-1 models. In the remainder of the text, we will use $l \in \{1,2,...(n_m-1)\}$ and $h \in \{1,2,...(n_m-1)\}$ with $l \neq h$ to consider two models at a time. The number of possible absolute inadequacies is $n_a = n_m - 1$, while the number of possible relative inadequacies between the $n_m - 1$ models is $n_r = \frac{(n_m-1)(n_m-2)}{2}$. The AMI of model h is given by

$$\mathbf{\varepsilon}_h = [\varepsilon_{f_h} \ \varepsilon_{G_h}], \tag{2}$$

where ε_{f_h} and ε_{G_h} are model inadequacies of the objective and constraint functions, respectively. They are defined as

$$\varepsilon_{f_h} = |f_0 - f_h| \quad \text{and} \quad \varepsilon_{G_h} = |g_0 - g_h|$$
 (3)

with respect to the reference model M_0 .

At this point, model inadequacies can be represented graphically in an inadequacy space. The number of dimensions in the inadequacy space depends on the number of functions, i.e., $\mathbf{\epsilon}_h \in \mathbf{R}^{1+U}$. To keep dimensionality at two, we define one aggregate constraint function that sums constraint violations (if any) [43]. Consequently, the inadequacy vector is reduced to consisting of two components $\mathbf{\epsilon}_{f_h}$ and $\mathbf{\epsilon}_{G_h}$.

In the example of Figure 2, if we choose $M_{1_{1_1}}$ as a reference model, we have 13 remaining models; thus, $n_a=13$ and $n_r=78$. After specifying a reference model, a relative adequacy matrix (RAM) $\tilde{\mathbf{e}}$ of size $((n_m-1)\times(n_m-1))$ is built to summarize the relative inadequacies among remaining models. For illustration, we will show the relative inadequacy space and the RAM for model instantiations $M_{1_{1_2}}$, $M_{1_{1_3}}$, $M_{1_{2_1}}$, $M_{1_{2_2}}$, $M_{1_{3_1}}$ and $M_{1_{3_2}}$ (recall that M_{1_1} is our reference model). The new indices are listed in Table 4.

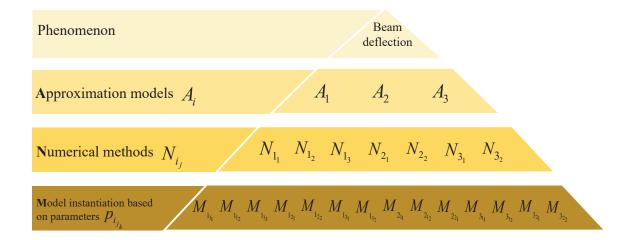


Fig. 2: Example of derivation of model instantiations

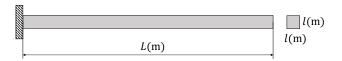


Fig. 3: Flexible cantilevered beam

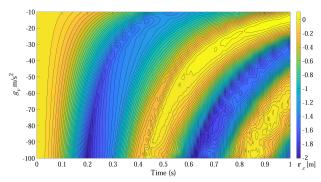
Table 4: New indices of considered models

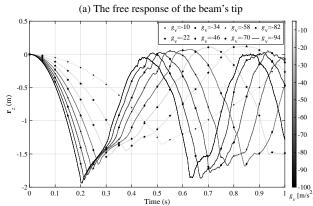
$M_{1_{1_1}}$	\rightarrow	M_0
$M_{1_{1_2}}$	\rightarrow	M_1
$M_{1_{1_3}}$	\rightarrow	M_2
$M_{1_{2_1}}$	\rightarrow	M_3
$M_{1_{2_2}}$	\rightarrow	M_4
$M_{1_{3_1}}$	\rightarrow	M_5
$M_{1_{3_2}}$	\rightarrow	M_6

Figure 7 depicts absolute model inadequacies, i.e., the values of $||\mathbf{\varepsilon}_h||_1$ for h = 1, 2, ..., 6 with M_0 being the reference model. Table 5 shows the RAM, whose elements are given by

$$\tilde{\boldsymbol{\varepsilon}}(h,l) = ||\boldsymbol{\varepsilon}_{l}||_{1} - ||\boldsymbol{\varepsilon}_{h}||_{1}
= |\boldsymbol{\varepsilon}_{f_{l}}| + |\boldsymbol{\varepsilon}_{G_{l}}| - |\boldsymbol{\varepsilon}_{f_{h}}| - |\boldsymbol{\varepsilon}_{G_{h}}|
= (|\boldsymbol{\varepsilon}_{f_{l}}| - |\boldsymbol{\varepsilon}_{f_{h}}|) + (|\boldsymbol{\varepsilon}_{G_{l}}| - |\boldsymbol{\varepsilon}_{G_{h}}|)
= \tilde{\boldsymbol{\varepsilon}}_{f_{hl}} + \tilde{\boldsymbol{\varepsilon}}_{G_{hl}}.$$
(4)

Each element $\tilde{\mathbf{\epsilon}}(h,l)$ in the RAM reflects a relative rank between pairs of models. For example, the positive value of $\tilde{\mathbf{\epsilon}}(1,3)$ indicates higher adequacy of model M_1 relative to M_3 since $||\mathbf{\epsilon}_3||_1 > ||\mathbf{\epsilon}_1||_1$. On the contrary, the negative value of $\tilde{\mathbf{\epsilon}}(1,6)$ indicates lower adequacy of model M_1 relative to M_6 since $||\mathbf{\epsilon}_1||_1 > ||\mathbf{\epsilon}_6||_1$. Figure 8 is a graphical represen-





(b) The free response of the beam's tip at selected gravitational acceleration values

Fig. 4: Tip displacements in z-direction (ANCF)

tation that depicts relative inadequacies for each element in the RAM.

2.2 Penalizing Relative Inadequacy

Let us assume that M_l has lower computational cost than M_h . Our objective is to minimize $\tilde{\epsilon}_{hl}$ so we can improve the fidelity of M_l to approach the one of M_h as much as possible while keeping the low computational cost of M_l .

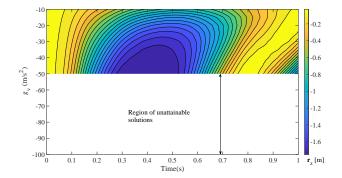


Fig. 5: Tip displacements in z-direction (Timoshenko)

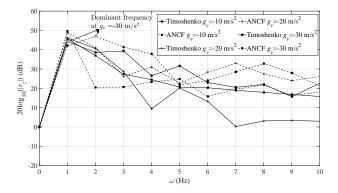


Fig. 6: Frequency response of the beam's tip at selected gravitational acceleration values

Table 5: An example of the relative adequate matrix

	M_1	M_2	M_3	M_4	M_5	<i>M</i> ₆
M_1	0	0.2	0.1	0.4	-0.5	-0.7
M_2	-0.2	0	-0.1	0.2	-0.7	-0.9
M_3	-0.1	0.1	0	0.3	-0.6	-0.8
M_4	-0.4	-0.2	-0.3	0	-0.9	-1.1
M_5	0.5	0.7	0.6	0.9	0	-0.2
<i>M</i> ₆	0.7	0.9	0.8	1.1	0.2	0

A traditional penalty function ϕ_h penalizes constraint violations in the objective function:

$$\phi_h(\mathbf{d}) = f_h(\mathbf{d}) + r_c \mathcal{F}_h(\mathbf{d}), \tag{5}$$

where r_c is the penalty term parameter and $\mathcal{F}_h = \sum_{u=1}^{U} max(g_{h_u}, 0)^2$ is the aggregate constraint function. We adopt this approach and modify the penalty function to take into account model's fidelity:

$$\phi_h(\mathbf{d}) = f_l(\mathbf{d}) + \tilde{\varepsilon}_{f_{hl}}(\mathbf{d}) + r_c \mathcal{F}_l(\mathbf{d}) + \tilde{\varepsilon}_{G_{hl}}(\mathbf{d}), \tag{6}$$

where

$$r_c = \begin{cases} 1 & \text{if } \mathcal{F}_h(\mathbf{d}) \le 0\\ \text{inf otherwise.} \end{cases}$$
 (7)

By utilizing Eq. (4), this function can be simplified to

$$\phi_h(\mathbf{d}) = \phi_l(\mathbf{d}) + \tilde{\varepsilon}_{hl}(\mathbf{d}). \tag{8}$$

We then minimize this penalty function using the derivativefree optimization algorithm described in the next section.

2.3 Mesh Adaptive Direct Search

Gradients of functions evaluated by means of computational models may not exist or may not be prone to reliable approximations at a reasonable cost [44]. Moreover, these models may be prone to severe numerical noise or fail to return a value. In these cases, we need to resort to derivative-free optimization algorithms. In this work we use the mesh adaptive direct search (MADS) algorithm of [45], which is supported by convergence properties.

MADS ensures global convergence to a solution satisfying local optimality conditions based on the Clarke calculus for nonsmooth functions [46]. Each MADS iteration consists of two steps: the search and the poll. The search step allows functions evaluation at any finite number of trial points generated by evaluating the objective function on a scaled discretized space of variables, called the *mesh* \mathcal{M} . The search favours exploration of the design space and allows users to implement any appropriate method, given their knowledge of the problem [47]. The mesh is centred around the incumbent solution $\mathbf{d}_k \in \mathbb{R}^n$ where the mesh size is parametrized by Δ_m . The poll step involves local exploration around \mathbf{d}_k . During this step, the distance between the trial point and the current incumbent solution is dictated by the poll size parameter Δ_p [48]. If the search fails to find an improvement, the poll proposes trial points around the incumbent solution following polling conditions. The optimization process either converges when the mesh size parameter is smaller than a user-specified threshold or terminates when a pre-described function evaluation budget is exhausted [47].

2.4 Our Search Step Strategy

We use the search step to solve a surrogate optimization problem. In this problem, the SM are not built for the output of the problem functions, but for the models' inadequacies. The purpose is to mitigate the discrepancy among available models to approach the adequacy level of the reference model while favoring the utilization of low-cost models. A Latin Hypercube (LH) is used to populate a set of design points \mathbf{S}_k to obtain a set of observations and build a surrogate model $\hat{\mathbf{E}}_{hl}$ which can be used to predict the value of $\tilde{\mathbf{E}}_{hl}(\mathbf{d}_k)$ for a given $\mathbf{d}_k \not\in \mathbf{S}_k$. The difference between the predicted inadequacy in the current iteration and the previous one is

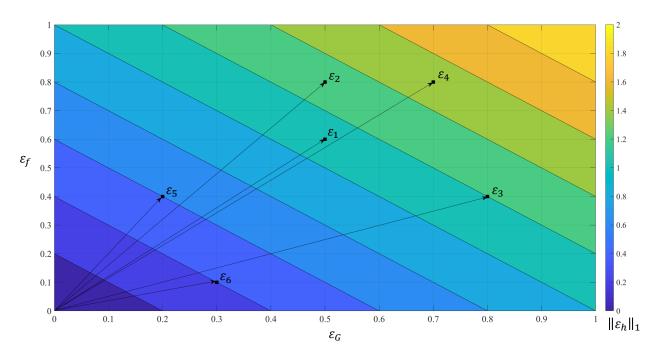


Fig. 7: Absolute inadequacies among models $M_1, M_2, ..., M_6$ with M_0 being the reference model

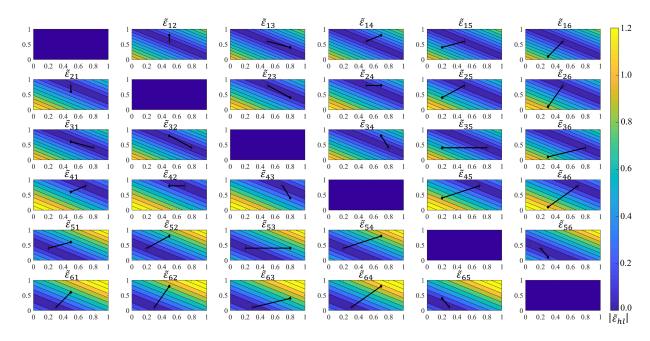


Fig. 8: Relative inadequacies among models M_h and M_l

calculated as follows

$$\Lambda_{hl}(\mathbf{d}_{k+1}) = \begin{bmatrix} \hat{\hat{\mathbf{\epsilon}}}_{f_{hl}}(\mathbf{d}_{k+1}) - \hat{\hat{\mathbf{\epsilon}}}_{f_{hl}}(\mathbf{d}_{k}) \\ \hat{\hat{\mathbf{\epsilon}}}_{G_{hl}}(\mathbf{d}_{k+1}) - \hat{\hat{\mathbf{\epsilon}}}_{G_{hl}}(\mathbf{d}_{k}) \end{bmatrix}.$$
(9)

In this work, we use polynomial response surfaces (PRSs) to build the SM [49–51]. The surrogate problem is defined as

$$\min_{\mathbf{d}_k \in \mathbb{R}^n} \quad \hat{\phi}_h = \phi_l(\mathbf{d}_k) + (\hat{\tilde{\epsilon}}_{hl}(\mathbf{d}_k) + ||\Lambda_{hl}(\mathbf{d}_k)||_1)$$
subject to
$$||\Lambda_{hl}(\mathbf{d}_k)||_1 \le \Delta_{hl}^r(\mathbf{d}_k),$$

$$(10)$$

where $\Delta_{hl}^{r}(\mathbf{d}_{k})$ is the radius of a trust region in the inadequacy space where surrogate model predictions are assumed to be reliable. For example, consider that we have four model instantiations to solve the problem in Eq. (10): M_{0} (used as the reference model), M_{1} , M_{2} and M_{3} . The predicted relative inadequacies $\hat{\epsilon}_{12}(\mathbf{d}_{k})$, $\hat{\epsilon}_{13}(\mathbf{d}_{k})$ and $\hat{\epsilon}_{23}(\mathbf{d}_{k})$ are bounded by trust regions with radii $\Delta_{12}^{r}(\mathbf{d}_{k})$, $\Delta_{13}^{r}(\mathbf{d}_{k})$ and $\Delta_{23}^{r}(\mathbf{d}_{k})$, respectively. In the next iteration, SM are used to predict $\hat{\epsilon}_{13}(\mathbf{d}_{k+1})$, $\hat{\epsilon}_{12}(\mathbf{d}_{k+1})$ and $\hat{\epsilon}_{23}(\mathbf{d}_{k+1})$.

In the example illustrated in Figure 9, $||\Lambda_{12}(\mathbf{d}_{k+1})||_1$ and $||\Lambda_{23}(\mathbf{d}_{k+1})||_1$ lie outside their trust regions, while

 $||\Lambda_{13}(\mathbf{d}_{k+1})||_1$ lies inside its trust region. Consequently, the

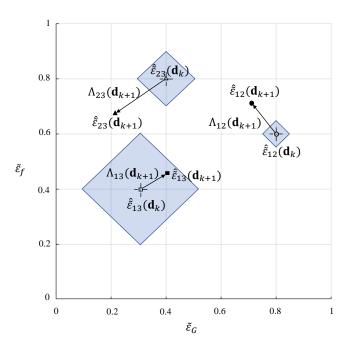


Fig. 9: Trust regions of the relative inadequacy space

predicted inadequacy $\hat{\mathbf{\epsilon}}_{13}(\mathbf{d}_{k+1})$ is accepted to enhance the output of the model $\phi_3(\mathbf{d}_{k+1})$. Utilizing the trust region in the relative inadequacy space bounds the predicted inadequacy to increase the trust in the estimated inadequacies and in the approximated objective function. If more than one predicted inadequacies lie inside their adjacent trust regions, then the inadequacy of the lowest cost model is utilized.

The trust region radius is updated as

$$\Delta_{hl}^{r}(\mathbf{d}_{k+1}) = \begin{cases} c_1 \Delta_{hl}^{r}(\mathbf{d}_k) & \text{if } r_{hl}(\mathbf{d}_k) < R_1 \\ c_2 \Delta_{hl}^{r}(\mathbf{d}_k) & \text{if } r_{hl}(\mathbf{d}_k) > R_2 \\ \Delta_{hl}^{r}(\mathbf{d}_k) & \text{otherwise,} \end{cases}$$
(11)

where

$$r_{hl}(\mathbf{d}_{k+1}) = \frac{\phi_h(\mathbf{d}_k) - \phi_h(\mathbf{d}_{k+1})}{\hat{\phi}_h(\mathbf{d}_k) - \hat{\phi}_h(\mathbf{d}_{k+1})}.$$
 (12)

The trust region method used here is the one reported in [52] as implemented by March and Willcox in [53]. The positive parameters $R_1 < R_2 < 1$ and $c_1 < 1, c_2 > 1$ are chosen to define the ranges of the ratio value for which reduction or enlargement are necessary [54]. In our implementation we have used the values $R_1 = 0.01$, $R_2 = 0.9$, $c_1 = 0.5$ and $c_2 = 1.5$. In case multiple estimated inadequacies exist within the trust region, the least expensive model is selected. If all estimated inadequacies lie outside the trust region, then all models are

evaluated to update the SM. The process of evaluating candidates using estimated model inadequacy within a trust region allows for an improvement of the solution with fewer evaluations. The modified MADS pseudo-algorithm using the RAF and the trust-region approach is provided in Algorithm 1. Figure 10 depicts a flow chart of the proposed algorithm.

2.5 Scaling Properties of our Algorithm

We investigate the scalability of our proposed algorithm to larger number of variables by solving the following test problem from [45]

$$\min_{\mathbf{d} \in \mathbb{R}^n} \qquad \sum_{i=1}^n d_i + \varepsilon_{f_h}$$
subject to
$$\sum_{i=1}^n d_i^2 + \varepsilon_{g_h} \le 3n.$$
(13)

To consider multiple models of varying fidelity and computational cost, we introduce what we call inadequacy terms to the objective and constraint functions as listed in Table 6. Using the reference model (i.e., for h = 0), there is a single optimal solution to the problem: every component of the vector **d** is equal to $-\sqrt{3}$ and the optimal objective function value is $-\sqrt{3}n$.

We solved the problem using our RAF-MADS algorithm for 6 values of n, namely 5, 10, 20, 30, 40, and 50. For each value of n, we solved the problem 5 times using different initial guesses. We used a convergence criterion of $\Delta_p \leq 10^{-12}$ and a termination criterion of a maximum of 600n function evaluations. In the search step, we used a 5n point Latin hypercube sample at the initial iteration and a n/5 random search at other iterations (if necessary).

Figure 11 depicts how the average (over the 5 runs for each n) number of function evaluations increases with n. The plot shows an increase in the reduced cost ratio: the fact that RAF-MADS utilizes relative relationships among models leverages the reduction in the number of evaluations of the reference model M_0 . We fitted an exponential to the obtained data, obtaining the expression

$$n_{feval} = 34.11 \times n^{1.55},$$
 (14)

where n_{feval} is the number of function evaluations. We tested this equation at n = 70 and obtained a result associated with a 4% relative residual error (see Figure 11).

3 Numerical Examples

We demonstrate the application of our proposed methodology on constrained design optimization problems considering both computational structural dynamics (CSD) and computational fluid dynamics (CFD) disciplines.

The CSD design problem of a flexible beam is used to investigate the impact of framework parameters on the

Algorithm 1 The MADS optimization pseudo-algorithm using relative model adequacies in trust regions

```
[1] Initialization (k = 0)
      Set maximum number of evaluations k_{\text{max}}.
      Set convergence tolerance (typically machine precision)
      Set initial poll and mesh sizes
      Set initial trust radiuses \Delta_{hl}^{r}(\mathbf{d}_{k}) = 0 and inadequacy distance ||\Lambda_{hl}(\mathbf{d}_{k})||_{1} = +\inf
      Populate initial guesses using Latin Hypercube (LH) and store them in the cache matrix \mathbf{D}_k
      Evaluate \phi_k using Eq. (5) \forall \mathbf{d}_k \in \mathbf{D}_k
      Set \phi_{\min} = \phi_k for \mathbf{d}_k that returns minimum values for f_k and \mathcal{F}_k
[2] Iteration
      [2.1] Search
            If k > 0, then
                   Populate possible points using LH and store them in S_k
                   Update the RAM using built SM
                   Compute the difference among model inadequacies of the current iteration and the previous one
            End if
            If ||\Lambda_{hl}(\mathbf{d}_k)||_1 > \Delta_{hl}^r(\mathbf{d}_k), then
                   Evaluate \phi_k \ \forall \mathbf{d}_k \in \mathbf{S}_k  where \mathbf{S}_k \subset \mathcal{M}, but \mathbf{d}_k \notin \mathbf{D}_k
                   Disregard models that fail to return values
                   Calculate model cost ratios
                   Select the reference model that has the highest computational cost
                   Generate the RAM using Eq. (4)
                   Build a surrogate model for each inadequacy \hat{\tilde{\epsilon}}_{hl}(\mathbf{d}_k) in the RAM
                   Evaluate trust ratios and trust radiuses using Eq. (12) and Eq. (11), respectively
                   If \phi_k < \phi_{\min}, then
                         Update \phi_{\min} = \phi_k and \mathbf{d}_{\min} = \mathbf{d}_k
                         Go to 3
                   End if
            Otherwise,
                   Select the model inadequacy that lies inside the trust region adjacent to the model with the lowest cost ratio
                   Enhance the model output using the selected model inadequacy
                   If \hat{\phi}_k < \phi_{\min}, then
                         Update \phi_{\min} = \hat{\phi}_k and \mathbf{d}_{\min} = \mathbf{d}_k
                         Go to 3
                   End if
            End if
      [2.2] Poll
            Determine the polling directions and frame
            Update the RAM using the SM built in 2.1
            Compute the difference between model inadequacies of the current iteration and the previous one
            Select the model inadequacy that lies inside the trust region adjacent to the model with the lowest cost ratio
            If all inadequacies lie outside their adjacent trust region, then go to 2.1
            Evaluate \hat{\phi}_k at each design point in the polling set using the model adjacent to the selected inadequacy
            Enhance the model output using the selected model inadequacy
            If \hat{\phi}_k < \phi_{\min}, then
                   Update \phi_{\min} = \hat{\phi}_k and \mathbf{d}_{\min} = \mathbf{d}_k
            End if
[3] Updates
      k \leftarrow k + 1
      Update \Delta_m, \Delta_P, \Delta_k^r, \mathbf{d}_k and \mathbf{D}_k
      If no stopping criteria is met, then go to 2.2
      Otherwise terminate the algorithm; return \mathbf{d}_{min} and \phi_{min}
```

Table 6: Models used in scalability study

Model	Inadequacies ε_{f_h} and ε_{g_h}	θ	а	b	Cost ratio (M_h/M_0)	
M_0	$\varepsilon_{f_0} = 0$		_	_	1	
	$\epsilon_{g_0} = 0$				•	
M_1	$\frac{\varepsilon_{f_1} = a \times \mathbf{d} _2}{\varepsilon_{g_1} = \sin(\theta \times \mathbf{d} _2)}$	- π/4	0.013	-	0.57	
	$\mathbf{\epsilon}_{g_1} = \sin(\mathbf{\theta} \times \mathbf{d} _2)$				0.57	
M_2	$\varepsilon_{f_2} = a \times \mathbf{d} _1^2$		0.005	_	0.42	
	$\epsilon_{g_2}=0$		0.005		0.42	
M_3	$\mathbf{\varepsilon}_{f_3} = \sin(b \times \mathbf{\theta} \times \mathbf{d} _2)$	- π		0.5	0.12	
<i>IVI</i> 3	$\mathbf{\epsilon}_{g_3} = \sin(\mathbf{\theta} \times \mathbf{d} _2)$, , , , , , , , , , , , , , , , , , ,		0.5	0.12	
M_{4}	$\varepsilon_{f_4} = \exp(-\sin(\mathbf{d} _1)) + b$		0.01	0.3	0.0071	
1414	$\varepsilon_{g_4} = \exp(-\sin(\mathbf{d} _1)) - a$		0.01	0.5	0.0071	

cost ratio and accuracy of the overall optimization process. The beam analysis and initial configuration are adopted from [55]. The problem has a low-dimensional design space but involves non-smooth functions and solutions that exhibit numerical noise. A total of 9 optimization runs are conducted by choosing six different sizes of initial trust regions, and three different sampling sizes at the initial step (k = 0). For each run, we allow a maximum of 100n black box evaluations; the optimization process can obviously stop earlier if the MADS termination criterion is satisfied.

The CFD design problem is adopted from the works of Koziel and Leifsson [56–60]. In these works, the authors proposed a hybrid approach of MFM that utilizes both MFHM and MFSM to solve a physics-based design optimization problem [56]. Moreover, they tested their proposed approach for different design space ranges of the airfoil shape optimization problem. To investigate the scalability of our proposed framework, we reproduce their approach and illustrate the different aspects between our proposed framework and their MFM method.

3.1 Structural Dynamics of Flexible Cantilever

The dynamic response of the cantilever beam described in Table 7 can be computed using three different models: absolute nodal coordinate formulation (ANCF) (A_1) , Euler-Bernoulli (A_2) , and Timoshenko (A_3) . The main difference

Table 7: Flexible beam dimensions and material properties

E (Pa)	$g_v (\text{m/s}^2)$	L (m)	$C(m^2)$	$\rho (kg/m^3)$	$I(\mathrm{m}^4)$
0.7e6	-30	d (1)	d (2)	5400	$I(\mathbf{d}(2))$

among the three models is the way the cross sectional rotations are calculated, which has a considerable effect on the simulation of deformation shapes under large loads. Euler-Bernoulli beam's cross sectional rotations are assumed to occur due to bending only, since the plane section remains normal to the longitudinal axis. However, in Timoshenko's theory they are assumed to be a sum of two contributions of bending and shear deformations. ANCF makes no assumptions on the sectional rotations which are represented by interpolation functions. This allows for the cross section to deform and change shape [61]. It is expected that this type of description, together with a three dimensional continuum mechanics approach, leads to more adequate results [62].

For isotropic material the vertical displacement r_z satisfies the following partial differential equation under the Euler-Bernoulli beam assumption

$$EI\frac{\partial^4 r_z}{\partial x^4} = q(x),\tag{15}$$

where q(x) is the distributed load which satisfies the following condition under the Timoshenko beam assumption [63, 64]:

$$EI\frac{\partial^4 r_z}{\partial x^4} = q(x) - \frac{EI}{\gamma C \lambda_g} \times \frac{\partial^2 q}{\partial x^2},$$
 (16)

where γ is the shear correction factor, and λ_g is the modulus of rigidity.

In ANCF the shape function is obtained using polynomials that are cubic in x and linear in y and z, where x is a material coordinate along the beam axis and y and z are the two other perpendicular coordinates. The configuration of the beam element is determined by the position and slope vectors of two end nodes N_p and N_q as shown in Figure 12. Each node is defined by one vector for the position \mathbf{r} and three vectors the slopes \mathbf{r}_x , \mathbf{r}_y and \mathbf{r}_z . Thus the element has

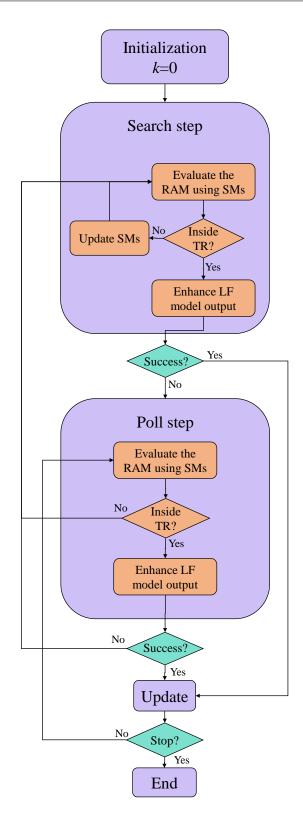


Fig. 10: Flow chart of the algorithm integrating mesh adaptive direct search with the relative adequacy framework using trust region principles

24 nodal coordinates given by the vector

$$\mathbf{e} = \left[\mathbf{r}_{p}^{\mathrm{T}} \mathbf{r}_{px}^{\mathrm{T}} \mathbf{r}_{py}^{\mathrm{T}} \mathbf{r}_{pz}^{\mathrm{T}} \mathbf{r}_{q}^{\mathrm{T}} \mathbf{r}_{qx}^{\mathrm{T}} \mathbf{r}_{qy}^{\mathrm{T}} \mathbf{r}_{qz}^{\mathrm{T}}\right]^{\mathrm{T}}.$$
 (17)

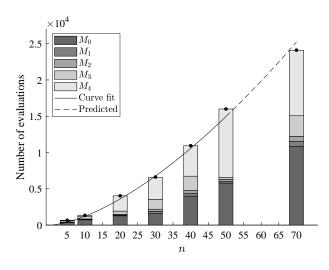


Fig. 11: Average number of models evaluation vs *n*

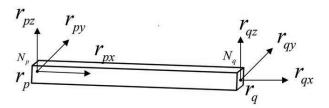


Fig. 12: ANCF beam

The location of an arbitrary point r in the beam is determined by the interpolation

$$\mathbf{r} = \mathbf{S}(x, y, z)\mathbf{e},\tag{18}$$

where ${\bf S}$ is the element shape function and ${\bf e}$ is the vector of nodal coordinates

$$\mathbf{S} = \begin{bmatrix} S_1 \mathbf{I} \ S_2 \mathbf{I} \ S_3 \mathbf{I} \ S_4 \mathbf{I} \ S_5 \mathbf{I} \ S_6 \mathbf{I} \ S_7 \mathbf{I} \ S_8 \mathbf{I} \end{bmatrix}^{\mathrm{T}}, \tag{19}$$

I is the 3×3 identity matrix, and

$$S_{1} = 1 - 3\xi^{2} + 2\xi^{3}$$

$$S_{2} = l(\xi - 2\xi^{2} + \xi^{3})$$

$$S_{3} = l(1 - \xi)\eta$$

$$S_{4} = l(1 - \xi)\zeta$$

$$S_{5} = 3\xi^{2} - 2\xi^{3}$$

$$S_{6} = l(-\xi^{2} + \xi^{3})$$

$$S_{7} = l\xi\eta$$

$$S_{8} = l\xi\zeta$$
(20)

with the nondimensional coordinates

$$\xi = x/l, \quad \eta = y/l, \quad \zeta = z/l \tag{21}$$

and l the initial length of the beam element. For Euler-Bernoulli and Timoshenko beam elements the shape functions S_1 , S_2 , S_5 and S_6 are calculated using Eq. (20), and $S_3 = S_4 = S_7 = S_8 = 0$. Using the finite-element method leads to solving the equations of motion in discretized form

$$[\mathbf{M}]\ddot{\mathbf{x}} + [\mathbf{K}]\mathbf{x} = \mathbf{Q}(\mathbf{x}), \tag{22}$$

where [M] and [K] are the global mass and stiffness matrices respectively, and Q is the global force vector. The mass and stiffness matrices can be found in [61, 65, 66]. The kinetic energy is calculated by

$$T = \frac{1}{2}\dot{\mathbf{e}}^{\mathrm{T}}[\mathbf{M}]\dot{\mathbf{e}}.\tag{23}$$

3.1.1 Problem Formulation

The beam kinetic energy (KE) $T(\mathbf{d})$ is minimized considering the beam cross section area C and the beam length L as design variables where $\mathbf{d} = [L \ C]^{\mathrm{T}}$. The maximum absolute displacement in the vertical direction of the beam tip for $t \in [0, 1.2\mathrm{s}]$ must be greater than or equal to 1.5m. The design domain is defined by the lower and upper bounds \mathbf{d} and \mathbf{d} of \mathbf{d} , where $\mathbf{d} = \begin{bmatrix} 1 & 0.1 \end{bmatrix}^{\mathrm{T}}$ and $\mathbf{d} = \begin{bmatrix} 4 & 1 \end{bmatrix}^{\mathrm{T}}$.

3.1.2 Modeling Parameters

The beam response can be obtained by means of three models named here "Euler-Bernoulli beam," "Timoshenko beam," and "ANCF." The optimization process is conducted considering a vector of parameters \mathbf{p} for each model, $\mathbf{p} = [ne \ \Delta t]$, where ne is the number of meshing elements and Δt is the time step. The beam is under the effect of gravitational acceleration $g_v = -30 \text{ m/s}^2$.

"ANCF" is assumed to be the reference model M_0 ; "Euler-Bernoulli beam" and "Timoshenko beam" are used as models M_1 and M_2 , respectively, where the numerical integration method is Newmark. We use the validated models reported in [55, 67, 68]. Parameter values for each model are shown in Figure 13.

3.1.3 Results

Figure 14 depicts three different feasible design regions obtained using M_0 , M_1 and M_2 . The high-cost model M_0 is used as a reference to improve both lower-fidelity models M_1 and M_2 while keeping the computational cost as low as possible. Computational cost ratios are summarized in Table 8. Table 9 summarizes the MFO/RFO and reduced cost ratios for nine optimization runs; the first six runs use different initial trust region radii and the last three runs use different initial sample sizes. We can conclude that when the initial

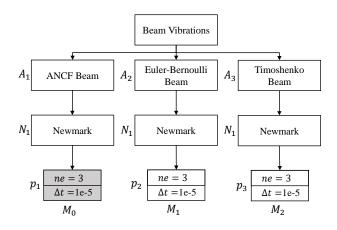


Fig. 13: Available models for predicting beam response

Table 8: Cost ratios among different models

Models	Cost ratio
M_2/M_0	0.15
M_1/M_0	0.1
M_1/M_2	0.66

trust radius increases, the computational cost decreases, but with an increase in the error of the obtained solution (defined as the relative deviation from the best-known solution of the problem). Figure 15 depicts the convergence history of the first six runs.

Figure 16 depicts the choice of models based on the multi-model management method at each iteration of the optimization algorithm for the first six optimization runs.

We observe that the higher-fidelity reference model is preferred in the early stages of the optimization process (where data are collected to improve the prediction capability of the lower-fidelity models), but with less frequency as the trust region is decreasing. We can also observe that the higher-fidelity reference model is preferred when the predictive capability of the lower-fidelity models deteriorates as new areas of the design space are explored as depicted in Figure 17. Figure 18 illustrates that there is no clear relationship between the size of the initial trust regions and the number of evaluations of M_0 in the first six runs. For the last three runs, we observe a decrease in the number of evaluations of M_0 and an increase in the number of evaluations of M_1 . This observation explains the linear increase in the reduced cost ratio, the reason being that the increase of the number of samples at the initial design step contributes to larger trust regions in the approximated model inadequacy as depicted in Figure 19.

3.2 Airfoil Shape Parameterization

We consider this problem to investigate the scalability of our proposed framework and illustrate its differences from existing MFO approaches. Koziel and Leifsson solved this

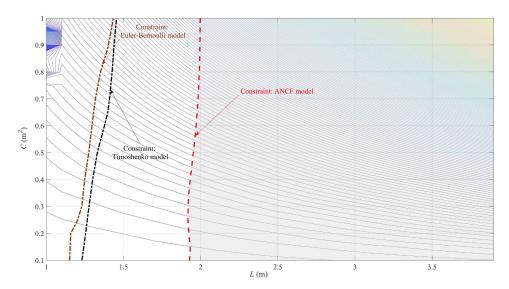


Fig. 14: The feasible design space obtained using three different models

Table 9: (MFO /RFO) cost ratios of the con-	ducted optimization runs
--------------------	---------------------------	--------------------------

Runs	Cost ratio	Reduced cost	$n_{\rm int}$	$\Delta^r_{hl}(\mathbf{d}_0)$	Error%
Run 01	60.55%	39.45%	27	0.8	0.5%
Run 02	55.61%	44.39%	27	0.9	1.1%
Run 03	48.72%	51.28%	27	2	2%
Run 04	48.89%	55.11%	27	2.5	2%
Run 05	27.46%	72.54%	27	3	4.4%
Run 06	25.20%	74.30%	27	3.2	7.2%
Run 07	24.9%	73.83%	30	3.2	6.8%
Run 08	23.4%	73.32%	36	3.2	6.1%
Run 09	22.1%	72.71%	42	3.2	5.8%

problem to study the relationship between the design problem dimensionality and the computational cost of the optimization process solved using their proposed algorithm [56].

NACA-four digits has a limited control to parameterize the airfoil shape with only three design variables [56]. However, shape parameterization techniques (e.g., Bèzier curves) can provide better control by invoking more design parameters (control points).

3.2.1 NACA and Bèzier Airfoils

In the NACA-four digits, the airfoil is defined by three design variables relative to the chord length, namely maximum ordinate of the mean camber m_c , chordwise position of the maximum ordinate p_c , and maximum thickness to chord ratio t_c [69].

Bèzier curves of order n_B are defined as

$$B(t_c) = \sum_{k_B=1}^{m_B} \sum_{i=0}^{n_B} (1 - b_B(k_B))^{n_B - i} b(k_B)^i P_B(i),$$
 (25)

where $P_B(i)$ are control points and b_B is a row vector with m_B components verying from zero to one. Control points location is selected to represent the airfoil shape. Koziel and Leifsson used two setups in their study: one with 6 and one with 8 control points (see Figure 20) [56].

The norm of the difference between the actual NACA0012 and the curve shapes generated by Eq. 25 is minimized to find the approximate Bèzier airfoil shapes as shown in Figure 20. Invoking more control points yields a better approximation to the actual airfoil shape. This improves the control of the airfoil shape during the optimization process. However, additional control points increase computational cost.

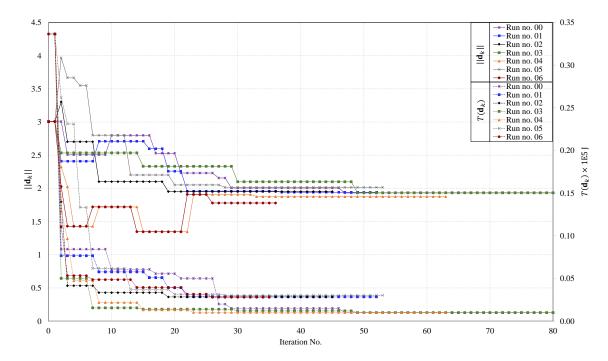


Fig. 15: Convergence history

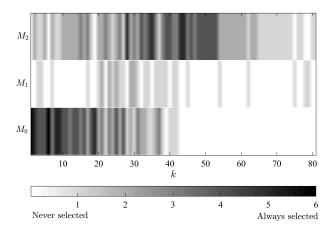


Fig. 16: Selected models in six optimization runs

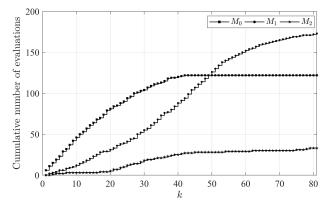


Fig. 17: Cumulative number of model evaluations in the first six runs

3.2.2 Problem Formulation

The transonic design problem considers a free-stream Mach number $\mathrm{Mach}_{\infty}=0.75$ and an angle of attack $\alpha=1^{\circ}$. The objective is to maximize the lift coefficient C_l subject to constraints on the wave drag coefficient $C_{\mathrm{dw,max}}=0.006$ and the non-dimensional cross-sectional area $A_{\mathrm{cs,min}}=0.075$. The initial designs in all cases are approximations of the NACA 2412 airfoil shape. These approximation shapes are obtained for each design case by minimizing the norm of the difference between NACA2412 and the airfoil shape obtained by parameterization. The initial designs are different in each design problem since the number of design variables differs. The design domain for each problem is defined by the lower and upper bounds \mathbf{d} and $\overline{\mathbf{d}}$ listed in Table 10. The

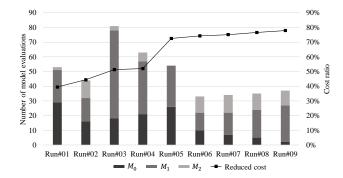
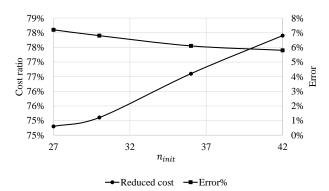


Fig. 18: Number of model evaluations and the reduced cost ratio in nine optimization runs

DP#1 DP#2 **DP#3** d $\overline{\mathbf{d}}$ $\overline{\mathbf{d}}$ d $\overline{\mathbf{d}}$ \mathbf{d}^* \mathbf{d}^* \mathbf{d}^* ₫ d \mathbf{d}_0 \mathbf{d}_0 \mathbf{d}_0 0.09 0.05 0.12 0.13 0.1134 0.06 0.2 0.03 0.06 0.05 0.2 0.05 d_1 0.4 0.3 0.6 0.5435 -0.06-0.2-0.05 -0.04-0.06 -0.2-0.05 -0.07 d_2 0.02 d_3 0 0.03 0.0178 0.3 0.2 0.8 0.36 0.3 0.2 0.6 0.25 d_4 _ 0.1 0.1 0.3 0.12 0.1 0.1 0.3 0.155 0.3 0.2 0.8 0.22 0.3 0.2 0.6 0.4 d_5 -0.1 -0.2-0.01 -.09 -0.1 -0.2-0.01 -0.045 d_6 d_7 _ _ _ _ 0.002 0.001 0.005 0.002 0.002 0.001 0.005 0.002 d_8 0.85 0.6 0.8 0.85 0.025 0.1 0.3 0.043 d_9 0.85 0.6 0.8 0.8 d_{10}

Table 10: Design variables of the three optimization design problems



-

-

-

Fig. 19: Reduced cost ratio and the solution accuracy in runs 7-9

optimization problem is summarized as

 d_{11}

$$\begin{aligned} & \min_{\mathbf{d} \in R^n} & -C_l(C_p(\mathbf{d})) \\ & \text{s.t.} & & 0.075 - A_{\text{cs}}(\mathbf{d}) \leq 0 \\ & & & C_{\text{dw}}(C_p(\mathbf{d})) - 0.006 \leq 0 \end{aligned}$$
 (26)

Nondimensional coefficients of lift C_l and wave drag C_{dw} are calculated from

$$C_l = -C_x \sin \alpha + C_z \cos \alpha,$$

$$C_{\text{dw}} = C_x \cos \alpha + C_z \sin \alpha.$$
(27)

The force coefficients C_x and C_z are calculated by integrating the pressure coefficient C_p counter clockwise around the

airfoil surface as

$$C_x = \oint C_p \sin \theta \, ds,$$

$$C_z = -\oint C_p \cos \theta \, ds,$$
(28)

-0.01

-0.005

where ds is the panel length on the surface of the airfoil and θ is the angle between the panel and the horizontal axis. The pressure coefficient is defined as

-0.025

-0.2

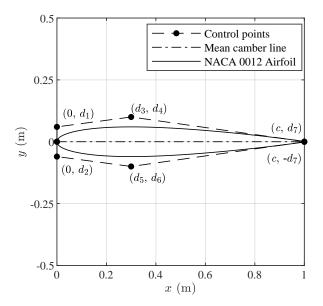
$$C_p = \frac{p - p_{\infty}}{0.5\rho_{\infty}V_{\infty}},\tag{29}$$

where p is the pressure in the flow-field, p_{∞} is the free-stream pressure, ρ_{∞} is the free-stream density, and V_{∞} is the free-stream velocity.

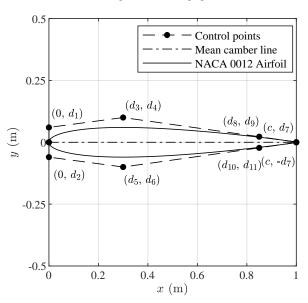
3.2.3 Derivation of Available Models

Transonic flow phenomenon occurs when there are mixed subsonic and supersonic local flows over a body moving with a speed near the speed of sound. Typically, the supersonic region of the flow is terminated by a shock wave where the flow slows down to subsonic speed. Across the shock there is an abrupt increase in pressure, temperature, density, and entropy. In this work, we consider a two-dimensional transonic flow past an airfoil section assuming a steady, inviscid, and adiabatic flow with no body forces.

Approximation Models Two approximation models are utilized to represent the transonic flow phenomenon with



(a) 6 control points and 7 shape parameters



(b) 8 control points and 11 shape parameters

Fig. 20: Bèzier curve approximations of the NACA0012 airfoil

the aforementioned assumptions: the Euler equations and the transonic small disturbance equation (TSDE). The Euler equations are a set of coupled non-linear partial differential equations that represent the conservation of mass, momentum, and energy [56]. The TSDE is an approximation of the Euler equations (thin airfoils at small angle of attack).

Solution Methods The Euler equations require a timemarching solution method. In the steady case, it is assumed that time-marching proceeds until a steady-state solution is reached. In this work, the temporal discretization of the governing equations is accomplished using an implicit timemarching algorithm. For the flow spatial discretization, a second-order upwind scheme is utilized. Fluent [70] is used to solve this approximation model. The TSDE can be solved with a finite difference scheme on a rectangular grid. We also used the legacy code (TSFOIL) to solve the TSDE; however, the finite difference scheme which is embedded in the code is inaccessible [71], which makes the tool a blackbox and thus amenable to our framework since our derivative-free optimization algorithm is ideal for blackbox problems.

Modeling parameters To set up multimodels we use the grid independence studies that exist in [56]. These studies were conducted for the initial airfoil configuration only; therefore, they may need to be repeated if the airfoil shape parameters are modified. Figure 21 depicts the available models that utilize modeling attributes used in the adopted independent studies; *RMSE* is the root-mean-square error used as the termination criterion of each model instantiation.

For the Euler model, the free-stream Mach number, static pressure, and angle of attack are prescribed at the far-field boundary. The solution domain boundaries are placed at 25 chord lengths in front of the airfoil, 50 chord lengths behind it, and 25 chord lengths above and below it. The computational grids are of structured curvilinear body-fitted C-topology with elements clustering around the airfoil and growing in size with distance from the airfoil surface as depicted in Figure 22. The number of nodes in the discretized flow domain is denoted n_{nodes} . For the TSDE model, the number of panels n_{panels} represents the discretized elements of the airfoil surface.

3.2.4 Results

The three design optimization problems with different number of variables (denoted DP#1, DP#2, and DP#3) are solved using our proposed RAF algorithm where the problem formulation is the same, except that the number of design variables is different as listed in Table 10. While all initial designs approximate the same NACA2412 airfoil shape parameters, there is a difference among them due to different number of design variables. Consequently, there is a considerable variation in the lift and drag coefficients of the initial designs. Figure 23 depicts some similarities among the optimized airfoil shapes; however, there are differences at the tailing-edge, where additional control points provide more degree of flexibility to modify the airfoil shape. The flexible shape parameterization impacts the lift coefficient as the optimized airfoil in DP#3 has the highest lift compared to DP#1 and DP#2, as summarized in Table 11.

Model instantiations, M_1 , M_2 and M_3 , have low execution CPU time (about 1 to 3 seconds depending on the selected number of panels) and for M_0 , M_4 , M_5 and M_6 evaluation takes higher CPU time (about 10 to 30 minutes depending on the selected number of nodes and the threshold value of the RMSE). The total CPU time of evaluating the low cost models in each optimization run is approximately from half to one evaluation of the reference model. The total number

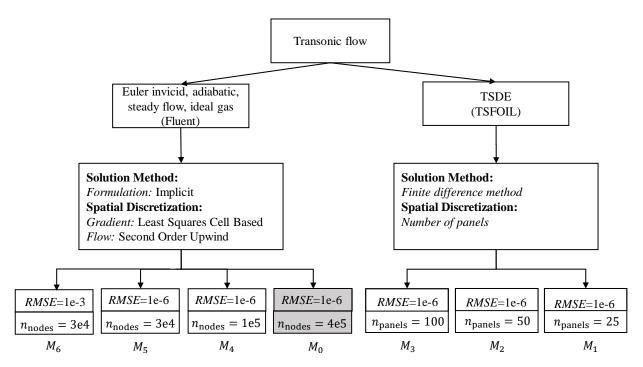


Fig. 21: Available model instantiations for predicting the aerodynamic characteristics of the airfoil

Table 11: Results of the three design problems for 3, 7, and 11 optimization variables using either only the high-fidelity reference model M_0 or multiple models managed by our relative adequacy framework

Aerodynamic — characteristics	DP#1			DP#2			DP#3				
	Initial	Optim	ized		Initial	Optim	ized		Initial	Optim	ized
	Illitiai	M_0 only	RAF	•	IIIItiai	M_0 only	RAF		muai	M_0 only	RAF
C_l	0.492	0.496	0.479		0.452	0.685	0.672		0.741	0.752	0.781
$C_{ m dw}$	0.01	0.006	0.006		0.003	0.006	0.006		0.005	0.006	0.006
Area	0.077	0.08	0.077		0.078	0.088	0.082		0.074	0.085	0.076

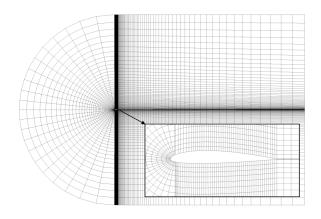


Fig. 22: A grid for the NACA 2412 airfoil

of models evaluation increases with additional design variables as depicted in Figure 24. From DP#1 to DP#2, the number of evaluations of M_0 increased by 1 (3 to 4), and from DP#2 to DP#3 by 2 (4 to 6). DP#1 needs 149 iterations

and DP#2 needs 247 iterations, giving an increase of 98 iterations with the additional 4 design variables. The number of iterations increases to 374 for DP#3, an increase of 127 for the additional 4 design variables. The overall computational cost rises accordingly. However, the reduced cost ratio of the optimization process increases by 2.4% from DP#1 to DP#2 and by 0.3% from DP#2 to DP#3. Figure 24 depicts a slope change which indicates that any further increase in the number of design variables may lead the reduced cost ratio to approach a plateau. A possible reason is that the number of the initial samples, which increases accordingly with the increase of the number of design variables, is part of the search step that contributes to reducing the total number of design iterations of the MFO compared to the RFO process. Further increase in the number of design variables may be accompanied by a significant increase in the cost ratio if the number of evaluations of the reference model approaches the number of design iterations of the RFO.

Figure 25 depicts the increase in the number of model evaluations for each model instantiation with increasing

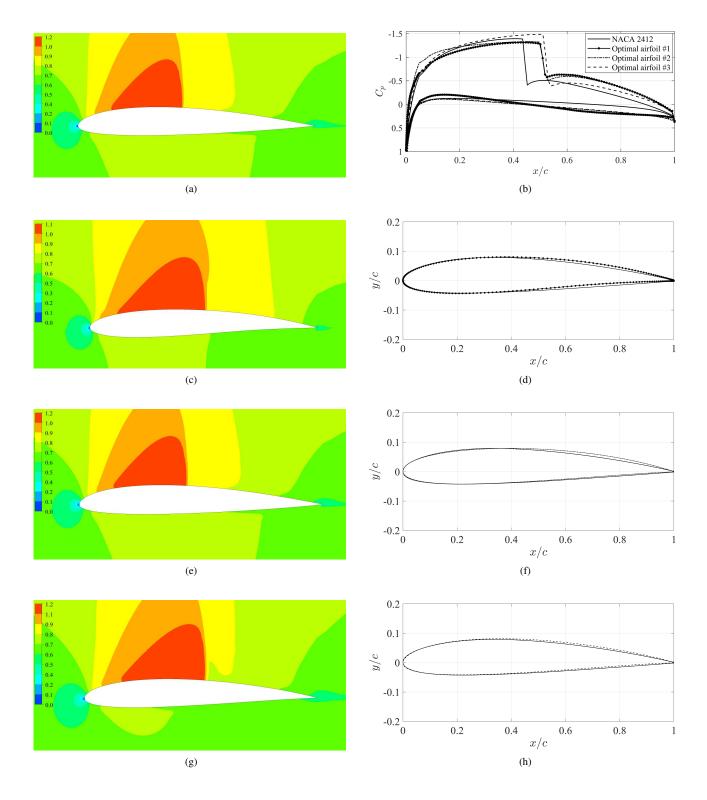


Fig. 23: (a) Mach number contours for the initial design of problem DP#3 ($M_{\infty} = 0.70$, $\alpha = 1^{\circ}$); (b) pressure coefficient on the surface of initial and optimal designs; (c), (e), and (g) Mach number contours of the optimal design for problems DP#1, DP#2, and DP#3, respectively; (d), (f), and (h) corresponding optimal airfoil shapes

number of design variables. The use of multiple models contributes to reducing the number of design iterations as the utilized relative relationships among model instantiations help to eliminate the deterioration in the inexpensive models by

enhancing their predictive capabilities without the need to evaluate the reference model. Koziel and Leifsson solved the same three cases of the design optimization problem using their MFSM approach [56]. Figure 25 plots the total number

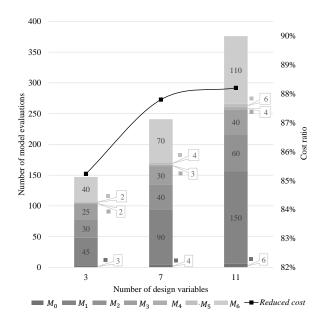


Fig. 24: Change in number of model evaluations and process cost ratio for increasing number of design variables

of evaluations they obtained along with ours. While curves have similar trends, our algorithm requires considerable less (approximately one fifth) total number of evaluations. It should be noted that we present this plot for information only, and not as a comparison. Koziel and Leifsson consider only one high-fidelity model and one surrogate model while our approach considers several numerical instantiations of physics-based models (even though surrogate models can be considered as well).

Figure 26 depicts the number of selections of each model instantiation during the optimization process of the three problems. We observe that the reference model is selected only at the beginning of the optimization process; thereafter more inexpensive models are selected.

Independent studies are conducted for the same boundary and initial conditions to tune modeling parameters that ensure numerical stability. Typically, these studies are conducted for the initial design guess; however, design variables are altered during the optimization process, which may change the boundary conditions of the analysis problem. This may require repeating the independent studies to ensure the adequacy of the selected modeling parameters. Our proposed framework assess these studies for the selected samples at the initial step and during the optimization process (if the approximated model inadequacies are mistrusted). The exhausted cost for that assessment is compensated as the possibility of misleading the search process is minimized which is reflected on reducing the total number of design iterations. On the other hand, that assessment contributes to reducing the number of evaluations of the reference model during the optimization process. The main difference between our methodology and other existing MFHM or MFSM approaches is that we utilize relative relationships among available models regardless of their expected fidelity to enhance

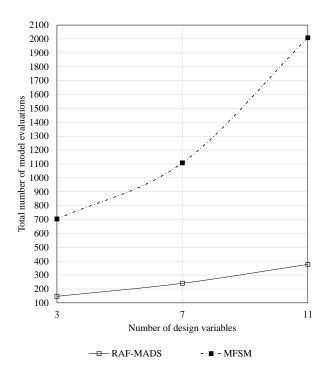


Fig. 25: Total number of model evaluations for each of the three cases required by the MFSM algorithm of Koziel and Leifsson and by our RAF-MADS algorithm. The MFSM curve is reproduced to the best of our ability based on the information available in [56]

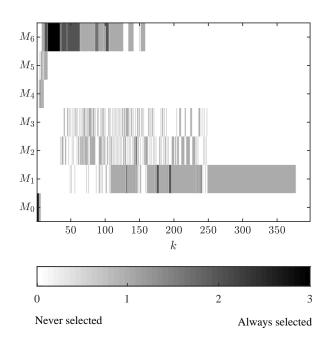


Fig. 26: Selected models during the optimization problem

inexpensive models and reduce the number of evaluations of expensive models.

4 Concluding Summary

This work introduces a framework for managing the use of models (whether physics- or data-based) of varying fidelity regardless of their disciplinary context. The framework considers a set of available models and assigns an adequacy level to each model based on its absolute and relative error relative to the other models and a reference model, respectively. This requires a reference model which is selected based on the highest computational cost among available models. Model inadequacies are represented graphically in a two-dimensional inadequacy space. Relative model inadequacies are concatenated in a relative adequacy matrix. The elements of this matrix are then used in an optimization problem formulation where a penalty function is minimized using the mesh adaptive direct search (MADS) algorithm. The MADS search step is utilized to solve a surrogate of the original optimization problem within trust regions where polynomial response surfaces are built to predict model inadequacies. This enables multiple models to make considerably faster predictions and attain solutions when non-linear phenomena emerge.

We demonstrate the proposed framework by solving two design optimization problems that have different disciplinary structures. The first problem is the optimization of a cantilevered flexible beam using three models. The presented results of this numerical example showed a significant reduction in the computational cost of the optimization process. The second problem is the airfoil shape optimization that utilizes Bèzier airfoils to enhance its aerodynamic characteristics. To investigate the scalability of our proposed framework, the size of the design space is scaled up by invoking more control points to the Bèzier airfoil geometry. The results show that both reduced cost ratio and optimization computational cost increases linearly with increased number of design variables.

The initial cost ratio (at the initial design step) increases with increasing number of available models and/or design variables. However, the reduced cost ratio does not decrease sharply. A possible reason is that the increase in the number of initial model evaluations increase the trust in approximated model inadequacy. Moreover, the evaluation of models is conducted at the search step. This is reflected in the reduction in number of reference model evaluations during the optimization process and the reduction in total number of design iterations. Independent studies are conducted for the initial design guesses and may be repeated (if necessary) during the optimization process to ensure numerical stability. We found that utilizing independent studies can assist the RAF to alleviate the initial cost ratio when it is used to solve design optimization problems with larger number of variables. In this work, we demonstrated the proposed framework using physics based models; however, it can be applied to any types of models, including (data-based) surrogate models.

Ongoing and future research focuses on extending the relative adequacy framework for solving multidisciplinary design optimization problems.

Acknowledgments

This work has been supported partially by FRQNT grant 2015-PR-182098 and NSERC grants CRDPJ 513922-17 and RGPIN 436193-18; the authors are grateful for this partial support, which does not constitute an endorsement of the opinions expressed in this paper. The first author is grateful to the Faculty of Engineering at McGill University for its partial support through a McGill Engineering Doctoral Award.

References

- [1] D. C. Gross. Fidelity Implementation Study Group Report. (No. SISO-REF-002–1999). Simulation Interoperability Standards Organization (SISO), 1999.
- [2] N. Alexandrov, E. Nielsen, R. Lewis, and W. Anderson. First-Order Model Management with Variable-fidelity Physics Applied to Multi-Element Airfoil Optimization. In 8th Symposium on Multidisciplinary Analysis and Optimization. AIAA, 2000.
- [3] N. Alexandrov, R. Lewis, C. Gumbert, L. Green, and P. Newman. Optimization with Variable-fidelity Models Applied to Wing Design. In 8th Symposium on Multidisciplinary Analysis and Optimization. AIAA, 2000.
- [4] Y. Kuya, K. Takeda, X. Zhang, and A. Forrester. Multifidelity Surrogate Modeling of Experimental and Computational Aerodynamic Data Sets. *AIAA journal*, 49 (2):289–298, 2011.
- [5] R. Vitali, R. Haftka, and B. Sankar. Multi-fidelity Design of Stiffened Composite Panel with a Crack. Structural and Multidisciplinary Optimization, 23(5):347–356, 2002.
- [6] A. Forrester, A. Sobester, and A. Keane. Engineering Design via Surrogate Modelling: A Practical Guide. John Wiley & Sons, 2008.
- [7] A. Forrester, A. Sóbester, and A. Keane. Multi-fidelity Optimization via Surrogate Modelling. In *Proceedings of the Royal Society of London A: Mathematical*, *Physical and Engineering Sciences*, volume 463, pages 3251–3269. The Royal Society, 2007.
- [8] S. Koziel and J. Bandler. Multi-fidelity Space Mapping Modeling of Microwave Devices with Double Coarse Model Processing and Functional Approximation. In *IEEE MTT-S International Microwave Symposium Di*gest, pages 969–972, June 2009.
- [9] A. DeBlois and M. Abdo. Multi-fidelity Multidisciplinary Design Optimization of Metallic and Composite Regional and Business Jets. AIAA Paper, 9191: 2011, 2010.
- [10] B. Peherstorfer, K. Willcox, and M. Gunzburger. Survey of Multifidelity Methods in Uncertainty Propagation, Inference, and Optimization. *SIAM Review*, 60 (3):550–591, 2018.
- [11] D. Allaire, K. Willcox, and O. Toupet. A Bayesian-Based Approach to Multifidelity Multidisciplinary Design Optimization. In 13th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, 2010.
- [12] M. Fernández-Godino, C. Park, N. Kim, and R. Haftka. Review Of Multi-fidelity Models. *arXiv preprint*

- arXiv:1609.07196, 2016.
- [13] J. Alonso, P. LeGresley, E. Van Der Weide, J. Martins, and J. Reuther. pyMDO: A Framework for High-fidelity Multi-Disciplinary Optimization. In 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, page 4480, 2004.
- [14] P. Piperni, A. DeBlois, and R. Henderson. Development of a Multilevel Multidisciplinary-Optimization Capability for an Industrial Environment. *AIAA journal*, 51(10):2335–2352, 2013.
- [15] A. Variyar, T. Economon, and J. Alonso. Multifidelity Conceptual Design and Optimization of Strut-Braced Wing Aircraft Using Physics Based Methods. In *54th AIAA Aerospace Sciences Meeting*, page 2000, 2016.
- [16] K. Alston and T. Winter. Balancing High Fidelity MDAO with Robust System Design. In 49th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition, page 884, 2011.
- [17] V. Balabanov, B. Grossman, L. Watson, W. Mason, and R. Haftka. Multifidelity Response Surface Model for HSCT Wing Bending Material Weight. In 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, page 4804, 1998.
- [18] M. Hutchison, E. Unger, W. Mason, B. Grossman, and R. Haftka. Variable-Complexity Aerodynamic Optimization of a High-Speed Civil Transport Wing. *Journal of Aircraft*, 31(1):110–116, 1994.
- [19] D. Rajnarayan, A. Haas, and I. Kroo. A Multifidelity Gradient-Free Optimization Method and Application to Aerodynamic Design. In *Proceedings of the 12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Victoria, Etats-Unis*, 2008.
- [20] R. Vitali, O. Park, R. Haftka, and B. Sankar. Structural Optimization of a Hat Stiffened Panel by Response Surface Techniques. In 38th Structures, Structural Dynamics, and Materials Conference, page 1151, 1997.
- [21] Y. Goldfeld, K. Vervenne, J. Arbocz, and F. Van Keulen. Multi-fidelity Optimization of Laminated Conical Shells for Buckling. Structural and Multidisciplinary Optimization, 30(2):128–141, 2005.
- [22] J. Umakant, K. Sudhakar, P. Mujumdar, and C. Rao. Ranking Based Uncertainty Quantification for A Multifidelity Design Approach. *Journal of aircraft*, 44(2): 410–419, 2007.
- [23] D. Allaire and K. Willcox. A Mathematical and Computational Framework for Multifidelity Design and Analysis with Computer Models. *International Journal for Uncertainty Quantification*, 4(1), 2014.
- [24] S. Bahrami, C. Tribes, C. Devals, T. Vu, and F. Guibault. Multi-fidelity Shape Optimization of Hydraulic Turbine Runner Blades Using a Multi-Objective Mesh Adaptive Direct Search Algorithm. *Applied mathematical modelling*, 40(2):1650–1668, 2016.
- [25] T. Robinson, M. Eldred, K. Willcox, and R. Haimes. Surrogate-Based Optimization Using Multifidelity Models with Variable Parameterization and Corrected Space Mapping. AIAA Journal, 46(11):2814–2822, 2008.

- [26] T. Robinson, M. Eldred, K. Willcox, and R. Haimes. Strategies for Multifidelity Optimization with Variable Dimensional Hierarchical Models. In *Proceedings* of the 47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference (2nd AIAA Multidisciplinary Design Optimization Specialist Conference), Newport, RI, pages 2006–1819, 2006.
- [27] S. Choi, J. Alonso, I. Kroo, and M. Wintzer. Multifidelity Design Optimization of Low-Boom Supersonic Jets. *Journal of Aircraft*, 45(1):106–118, 2008.
- [28] J. Jacobs, S. Koziel, and S. Ogurtsov. Computationally Efficient Multi-fidelity Bayesian Support Vector Regression Modeling of Planar Antenna Input Characteristics. *IEEE Transactions on Antennas and Propagation*, 61(2):980–984, 2013.
- [29] P. Qian and C. Wu. Bayesian Hierarchical Modeling for Integrating Low-Accuracy and High-Accuracy Experiments. *Technometrics*, 50(2):192–204, 2008.
- [30] L. Ng and K. Willcox. Multifidelity Approaches for Optimization Under Uncertainty. *International Journal* for Numerical Methods in Engineering, 100(10):746– 772, 2014.
- [31] M. Kennedy and A. ÓHagan. Predicting the Output from a Complex Computer Code when Fast Approximations are Available. *Biometrika*, 87:1–13, 2000. doi: 10.1093/biomet/87.1.1.
- [32] A. Padron, J. Alonso, and M. Eldred. Multi-fidelity Methods in Aerodynamic Robust Optimization. In 18th AIAA Non-Deterministic Approaches Conference, page 0680, 2016.
- [33] S. Koziel. Multi-fidelity Multi-Grid Design Optimization of Planar Microwave Structures with Sonnet. *International Review of Progress in Applied Computational Electromagnetics*, pages 719–724, 2010.
- [34] A. Padron, J. Alonso, F. Palacios, M. Barone, and M. Eldred. Multi-fidelity Uncertainty Quantification: Application to A Vertical Axis Wind Turbine Under an Extreme Gust. In 15th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, page 3013, 2014.
- [35] P. Molina-Cristóbal, A.and Palmer, B. Skinner, and G. Parks. Multi-fidelity Simulation Modelling in Optimization of A Submarine Propulsion System. In *Vehicle Power and Propulsion Conference (VPPC)*, pages 1–6. IEEE, 2010.
- [36] N. Alexandrov, R. Lewis, C. Gumbert, L. Green, and P. Newman. Approximation and Model Management in Aerodynamic Optimization with Variable-fidelity Models. *Journal of Aircraft*, 38(6):1093–1101, 2001.
- [37] J. Ren, L. Leifsson, S. Koziel, and Y. Tesfahunegn. Multi-fidelity Aerodynamic Shape Optimization Using Manifold Mapping. In 57th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, page 0419, 2016.
- [38] W. Kleiber, S. Sain, M. Heaton, M. Wiltberger, C. Reese, and D. Bingham. Parameter Tuning for a Multi-fidelity Dynamical Model of The Magnetosphere. *The Annals of Applied Statistics*, 7(3):1286–

- 1310, 2013.
- [39] M. Zahir and Z. Gao. Variable-fidelity Optimization with Design Space Reduction. *Chinese Journal of Aeronautics*, 26(4):841–849, 2013.
- [40] T. Oden, R. Moser, and O. Ghattas. Computer Predictions with Quantified Uncertainty, Part I. *SIAM News*, 43(9):1–3, 2010.
- [41] R. Lam. Surrogate Modeling Based on Statistical Techniques for Multi-fidelity Optimization. PhD thesis, Massachusetts Institute of Technology, 2014.
- [42] W. Oberkampf, T. Trucano, and C. Hirsch. Verification, Validation, and Predictive Capability in Computational Engineering and Physics. *Applied Mechanics Reviews*, 57(5):345–384, 2004.
- [43] Y. Ermoliev, A. Kryazhimskii, and A. Ruszczyński. Constraint Aggregation Principle In Convex Optimization. *Mathematical Programming*, 76(3):353–372, 1997.
- [44] J.J. Moré and S.M. Wild. Do you trust derivatives or differences? *Journal of Computational Physics*, 273: 268–277, 2014. doi: 10.1016/j.jcp.2014.04.056.
- [45] C. Audet and J.E. Dennis, Jr. Mesh adaptive direct search algorithms for constrained optimization. *SIAM Journal on Optimization*, 17(1):188–217, 2006. doi: 10.1137/040603371. URL http://dx.doi.org/doi:10.1137/040603371.
- [46] Frank H Clarke. *Optimization and Nonsmooth Analysis*. SIAM, 1990. doi: 10.1137/1.9781611971309.
- [47] B. Talgorn, S. Le Digabel, and M. Kokkolaras. Statistical Surrogate Formulations for Simulation-Based Design Optimization. *Journal of Mechanical Design*, 137 (2):021405–1–021405–18, 2015.
- [48] A. Conn, K. Scheinberg, and L. Vicente. *Introduction to Derivative-Free Optimization*. Society for Industrial and Applied Mathematics, 2009. doi: 10.1137/1.9780898718768.
- [49] E. Acar and M. Rais-Rohani. Ensemble of Metamodels with Optimized Weight Factors. *Structural and Multi-disciplinary Optimization*, 37(3):279–294, 2009.
- [50] N. Queipo, R. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, and P. Tucker. Surrogate-Based Analysis and Optimization. *Progress in aerospace sciences*, 41(1):1–28, 2005.
- [51] C. Audet, M. Kokkolaras, S. Le Digabel, and B. Talgorn. Order-Based Error for Managing Ensembles of Surrogates in Mesh Adaptive Direct Search. *Journal of Global Optimization*, 70(3):645–675, 2018.
- [52] A. Conn, K. Scheinberg, and L. Vicente. Global Convergence of General Derivative-free Trust-Region Algorithms to First-and Second-Order Critical Points. SIAM Journal on Optimization, 20(1):387–415, 2009.
- [53] A. March and K. Willcox. Provably Convergent Multifidelity Optimization Algorithm not Requiring High-Fidelity Derivatives. *AIAA journal*, 50(5):1079–1089, 2012.
- [54] J. Rodriguez, J. Renaud, B. Wujek, and R. Tappeta. Trust Region Model Management in Multidisciplinary Design Optimization. *Journal of Computational and*

- Applied Mathematics, 124(1):139–154, 2000.
- [55] K. Dufva, J. Sopanen, and A. Mikkola. A Two-Dimensional Shear Deformable Beam Element Based on the Absolute Nodal Coordinate Formulation. *Journal of Sound and Vibration*, 280(3-5):719–738, 2005.
- [56] S. Koziel and L. Leifsson. Scaling Properties of Multi-Fidelity Shape Optimization Algorithms. *Procedia Computer Science*, 9:832–841, 2012.
- [57] S. Koziel and L. Leifsson. Surrogate-Based Modeling and Optimization. *Applications in Engineering*, 2013.
- [58] S. Koziel and L. Leifsson. Multi-Level CFD-Based Airfoil Shape Optimization with Automated Low-Fidelity Model Selection. *Procedia Computer Science*, 18:889–898, 2013.
- [59] S. Koziel and L. Leifsson. Simulation-Driven Design using Surrogate-Based Optimization and Variable-Resolution Computational Fluid Dynamic Models. *Journal of Computational Methods in Sciences and Engineering*, 12(1, 2):75–98, 2012.
- [60] L. Leifsson and S. Koziel. Multi-Fidelity Design Optimization of Transonic Airfoils using Physics-Based Surrogate Modeling and Shape-Preserving Response Prediction. *Journal of Computational Science*, 1(2): 98–106, 2010.
- [61] R. Yakoub and A. Shabana. Three Dimensional Absolute Nodal Coordinate Formulation for Beam Elements: Implementation and Applications. *Journal of Mechanical Design*, 123(4):614–621, 2001.
- [62] A. Schwab and J. Meijaard. Comparison of Three-Dimensional Flexible Beam Elements for Dynamic Analysis: Finite Element Method and Absolute Nodal Coordinate Formulation. In ASME 2005 international design engineering technical conferences and computers and information in engineering conference, pages 1341–1349. American Society of Mechanical Engineers, 2005.
- [63] S. Rao and F. Yap. *Mechanical Vibrations*, volume 4. Prentice Hall Upper Saddle River, 2011.
- [64] Weaver J., S. Timoshenko, and D. Young. *Vibration Problems in Engineering*. John Wiley & Sons, 1990.
- [65] A. Shabana. ANCF Consistent Rotation-Based Finite Element Formulation. *Journal of Computational and Nonlinear Dynamics*, 11(1):014502, 2016.
- [66] A. Shabana. Flexible Multibody Dynamics: Review of Past and Recent Developments. *Multibody system dynamics*, 1(2):189–222, 1997.
- [67] H. Sugiyama, H. Koyama, and H. Yamashita. Gradient Deficient Curved Beam Element using the Absolute Nodal Coordinate Formulation. *Journal of Computational and Nonlinear Dynamics*, 5(2):021001, 2010.
- [68] M. Omar and A. Shabana. A Two-Dimensional Shear Deformable Beam for Large Rotation and Deformation Problems, 2001.
- [69] H. Abbott and E. Von Doenhoff. *Theory of Wing Sections, Including a Summary of Airfoil Data*. Courier Corporation, 1959.
- [70] Inc ANSYS. ANSYS Fluent Theory Guide. *Canonsburg*, *PA*, page 794, 2019.

[71] E. Murman, F. Bailey, and M. Johnson. TSFOILA Computer Code for 2D Transonic Calculation Including Wind Tunnel Wall Effects and Wave Drag Evaluations. In NASA Conference on Aerodynamic Analysis Requiring Advanced Computers, 1975.