Generalizable Multifidelity Aerodynamic Wing Shape Design Optimization

Aobo Yang *
The Hong Kong University of Science and Technology (HKUST), Clear Water Bay, New Territories, Hong Kong SAR

Jichao Li †
*Institute of High Performance Computing, A*STAR, Singapore 138632, Singapore*

Rhea P. Liem ‡
Imperial College London, London, SW7 2AZ, United Kingdom

Despite the demonstrated success of traditional surrogate-based approaches in establishing the relationship between design space and aerodynamic performance for specific aerodynamic shape optimization (ASO) problems, they still lack generalization for broader optimization tasks. To overcome this limitation, we propose a gradient-based, data-driven ASO framework for multifidelity, high-dimensional optimization. By integrating deep learning and multifidelity approach with gradient-based optimization, our method efficiently scales to complex design tasks while using a single model to solve multiple ASO problems, thereby enhancing adaptability and efficiency. We use compact modal parameterization to capture the global representation of wing shapes while reducing design dimensionality. To enable fast and efficient design solutions, the multifidelity surrogate models are fine-tuned with residual learning to preserve accuracy and control model complexity before being integrated into gradient-based optimization. In this paper, we showcase the versatility of the derived multifidelity surrogate model by applying it to a range of single-point and multipoint ASO problems, producing aerodynamic optimization results that are closely aligned with those from adjoint-based CFD solvers, while substantially

*PhD candidate in Department of Mechanical and Aerospace Engineering. AIAA Student Member.

†Scientist, Fluid Dynamics Department. AIAA Member.

‡Reader in Sustainable Aviation, Department of Aeronautics (Imperial College London); also Adjunct Associate Professor, Department of Mechanical and Aerospace Engineering (HKUST); r.liem@imperial.ac.uk. AIAA Associate Fellow (Corresponding Author).

§This article is an extended version of our earlier conference paper presented at the AIAA Aviation and Aeronautics Forum and Exposition (AIAA AVIATION 2023 Forum), San Diego, California, USA, 12–16 June 2023; AIAA Paper 2023-3470. The journal version includes improved methodology, broader experiments, and new results and analyses.

reducing wall-clock time during the optimization stage after the initial CFD-driven training data have been prepared. With this approach, we offer a new data-driven pathway for tackling general high-fidelity ASO tasks.

I. Introduction

Aerodynamic shape optimization (ASO) is essential in aircraft design, as it facilitates the exploration of high-dimensional geometric design spaces and identifies optimal solutions by leveraging physics-based evaluations. Aerodynamically-optimized wing designs have smoother pressure distribution, resulting in reduced drag and improved aerodynamic efficiency. This in turn contributes to improving fuel efficiency and cost-effectiveness, which are important for commercial aviation. Computational fluid dynamics (CFD)-based ASO has been the *gold standard* in aircraft design, but its widespread use—especially in industry and government agencies—is yet to be achieved [1], owing to the high computational cost. Typical design processes require iterative development, testing, and modifications [2]. Therefore, to truly support the complex aircraft design processes (especially at the *design exploration* stage), we need a comprehensive ASO framework that is fast, robust, and yet accurate. Improving the computational efficiency of ASO can make aerodynamic performance assessments possible even at the conceptual aircraft design phase, where different candidate designs are still being evaluated. This necessitates replacing the high-fidelity CFD analyses—which can be computationally prohibitive—with faster and more efficient models. Recent research efforts have extended ASO formulation from single-point optimization (focusing solely on the nominal condition) to an operation-aware multipoint one (incorporating more realistic flight conditions into the problem formulation) [3–5]. Such formulation complicates the design space and increases the required computational cost, further underscoring the needs to substantially improve the computational efficiency of ASO frameworks.

This shift in design paradigm—to support industry needs—leads to the dichotomy in ASO methods, namely the physics-based approach [6] that assesses candidate designs using high-fidelity numerical solvers, and a surrogate-based approach [7] that accelerates optimization by training predictive models on data. The physics-based approach solves the governing equations of fluid dynamics, such as the Navier-Stokes equations, using CFD solvers [8–10]. By leveraging the adjoint method [11], this approach efficiently computes accurate gradients through adjoint equations, enabling direct optimization [12–15]. When integrated with CFD solvers, the adjoint-based framework has demonstrated notable success in tackling high-dimensional optimization problems [16–18] with both efficiency and precision. A surrogate-based approach emulates expensive simulations using regression or interpolation models [4, 19]. It is usually referred to as a data-driven model when it leverages large-scale precomputed data—such as simulation or experimental results—and employs machine learning or statistical methods to predict aerodynamic performance without directly solving fluid dynamics equations [20–22]. Compared to physics-based approaches, the surrogate-based ones typically offer higher computational efficiency, often at the expense of accuracy in aerodynamic estimations used in the optimization. To balance the trade-off between computational resources and solution accuracy, multifidelity design optimization [19, 23] has emerged as a suitable solution.

The key point of multifidelity design optimization lies in leveraging data of varying levels of fidelity, accuracy, and computational cost to efficiently explore a design space and perform optimization. As an example, we can combine the aerodynamic performance analyses obtained using coarse (low-fidelity) and fine (high-fidelity) meshes. Whilst evaluating aerodynamic performance on the coarse mesh might sacrifice accuracy, it still provides valuable aerodynamic insights, which is illustrated as follows. Figure 1 presents drag polars obtained using the same CFD solver (`ADflow`) on three mesh resolutions, namely L1, L2, and L3 (from fine to coarse). The CFD solver and meshes (especially L2 and L3) are also used in this work, which will be further described in Section II.B. Using the same CFD solver in this comparison isolates the discretization effects in the multifidelity evaluation, thereby ensuring a consistent comparison between different fidelity levels. The figure demonstrates that evaluations obtained using different fidelity levels can capture key aerodynamic trends, with the medium-fidelity results (L2) showing a closer alignment with the high-fidelity (L1) results. This comparison underscores the effectiveness of low-fidelity (LF) data in guiding high-fidelity (HF) designs while substantially reducing computational costs. In a multifidelity setup, the infusion of high-fidelity data can improve the estimation accuracy.

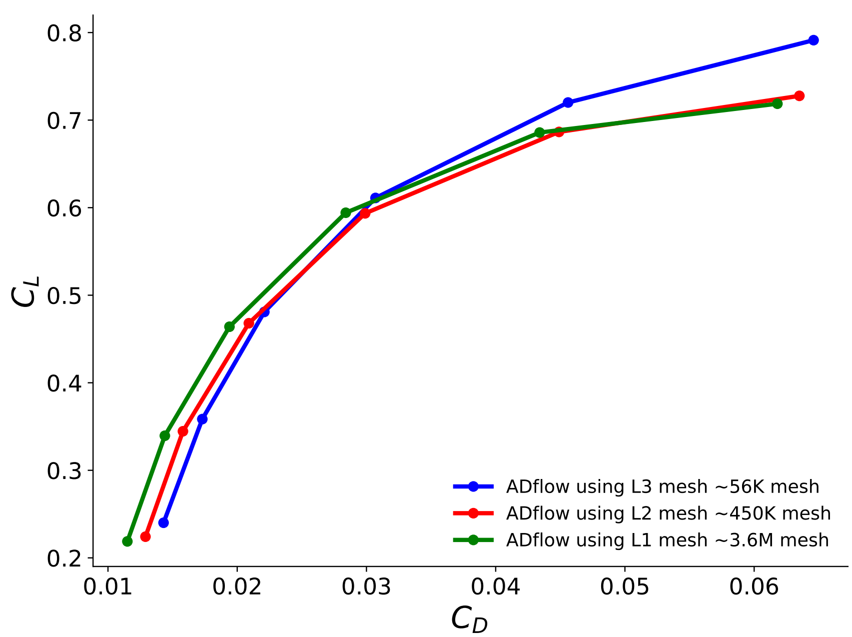
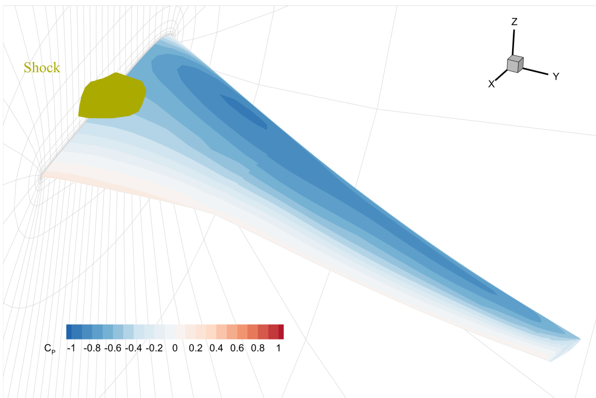
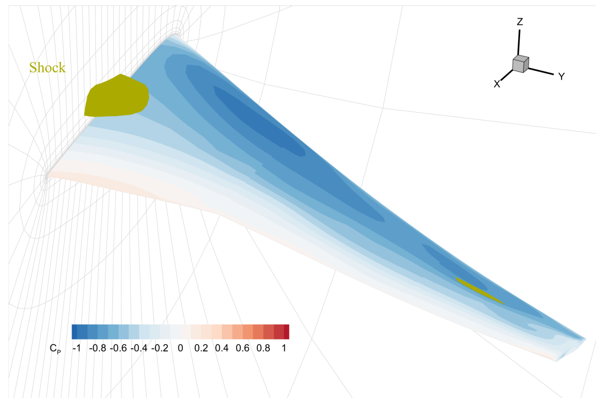


Fig. 1 Drag polar curves of aerodynamic data obtained using the same CFD solver on three fidelity levels.

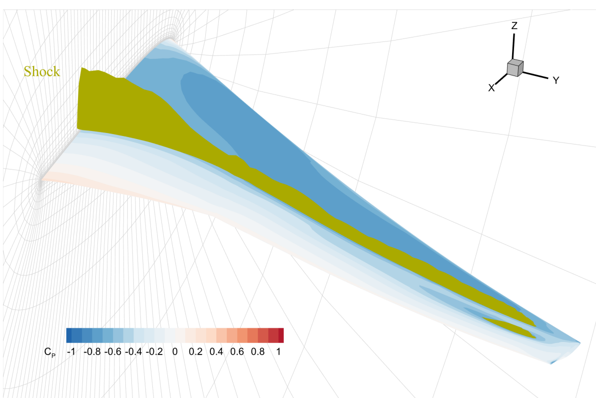
Despite its ability to capture overall aerodynamic trends, an optimization that relies solely on LF evaluations is unable to resolve fine-scale flow physics, as will be discussed below. In this study, we refer to the optimization that employs the CFD solver directly (as the aerodynamic estimator) as the *CFD-driven* approach, and the optimization that



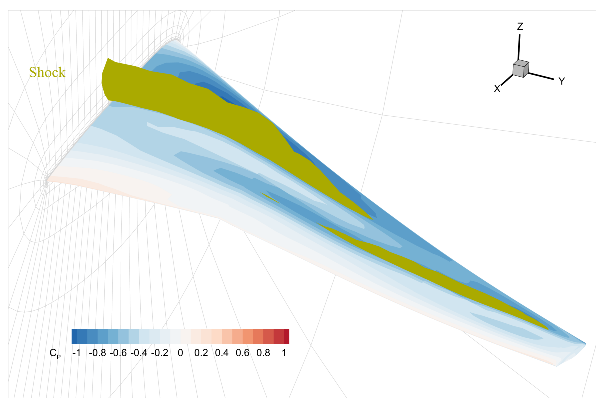
(a) CFD-driven optimized wing using the LF (L3) mesh.



(b) Data-driven optimized wing using LF data and the L3 mesh.



(c) The CFD analysis of the data-driven optimized wing shown in Figure 2b performed on the HF (L2) mesh.



(d) The CFD analysis of baseline wing on the HF (L2) mesh.

Fig. 2 Comparison of pressure coefficients (C_P) and shock regions for the baseline and optimized wings using CFD adjoint solver and data-driven approach.

relies on a surrogate model (pre-trained on CFD data) as the *data-driven* approach. Figure 2 compares the pressure distributions (C_P) and shock regions for four wings: (1) the CFD-driven wing optimized on a coarse (LF) mesh using steady Reynolds-averaged Navier–Stokes (RANS) solver (Figure 2a); (2) the data-driven optimized wing with the same LF data (Figure 2b); (3) the LF data-driven optimized wing (shown in Figure 2b) re-evaluated with CFD solver on a fine (HF) mesh (Figure 2c); (4) the baseline wing with fine (HF) mesh analyzed with CFD solver (Figure 2d). The optimization results shown in Figure 2 are obtained by solving the single-point ASO case, which will be presented in Section II.A. A yellow shock band appears whenever sub-optimal shapes accelerate the flow such that the local Mach number exceeds one, causing shock-induced abrupt pressure variations that increase wave drag. Figures 2a and 2b show that the data-driven approach attains aerodynamic performance comparable to the CFD-driven optimization on the LF mesh, with only minor differences in C_P and localized tip shocks. However, when the data-driven LF design (shown in Figure 2b) is re-analyzed on the HF mesh (Figure 2c), significant discrepancies emerge relative to the baseline (Figure 2d). This observation suggests that the LF-based optimization fails to capture the correct shock location and intensity, demonstrating a lack of physical consistency at higher fidelity. These results highlight the limitations of relying exclusively on LF optimization and motivate the incorporation of HF RANS data for accurate aerodynamic characterization.

The commonly used surrogate-based approach for solving multifidelity design optimization problems is the combination of gradient-free optimization algorithms and surrogate models, typically Gaussian Processes (GP) or Kriging, which has been investigated in previous studies [24–27]. However, this combination faces two major limitations, which stem from both the surrogate model and the optimization algorithm. From the surrogate model perspective, GP performs poorly in high-dimensional input spaces due to the well-known *curse of dimensionality*. Moreover, the computational cost of exact GP inference scales cubically with the number of data points, which makes them impractical for large-scale problems without approximations [28]. On the optimization side, gradient-free methods are generally less efficient than gradient-based ones when dealing with large-scale design optimization problems. Their inability to leverage gradient information limits both the representation and exploration of complex geometric features [29]. To address these two limitations, we propose a gradient-based optimization framework that leverages deep learning techniques in the surrogate model construction. Thanks to their scalability, deep neural networks are well-suited for large-scale, high-dimensional problems. The main advantage of this framework is its *generalizability*, which in this context is defined as the ability to reuse a single trained model across multiple design optimization problems under varying flight conditions. This capability is not achievable with traditional case-specific surrogate-based design optimization approaches.

Previous studies on neural network-based multifidelity modeling have primarily focused on network concatenation and composition strategies. These include the composite physics-informed neural networks (PINNs) proposed by Meng and Karniadakis [30], which encode partial differential equations (PDEs); and the multi-step neural networks developed by Guo *et al.* [31], which integrate the strengths of earlier approaches to reduce the computational cost of low-fidelity data when solving PDEs. In the context of ASO, Zhang *et al.* [32] employed the framework of composite neural networks to optimize the DLR-F4 configuration with 30 design variables, using a gradient-free particle swarm optimizer. However, it is worth noting that under such network composition, both physical field predictions and design optimization can become increasingly complex. The associated computational cost is exacerbated when more fidelity levels are considered in the framework. This relationship between hierarchy of fidelity and model complexity has been discussed in some previous studies [30, 33].

Compared to composite strategies, fine-tuning [34, 35] improves resource efficiency and lowers computational cost by reusing pre-trained models that already capture general features, thereby requiring significantly fewer data, epochs, and parameters to adapt to new tasks and avoiding the need to train deep networks from the start. In aerodynamics, recent research [36–38] has applied similar transfer learning techniques to model aerodynamic performance and flow fields, showcasing their versatility and effectiveness in addressing complex design and analysis challenges. In this work, we address the need for high-fidelity, high-accuracy modeling by introducing the concept of residual learning. Instead of learning the target values directly, we fine-tune the model by learning the residuals. Similar ideas have been applied in other domains, such as large language model [39] and robotics [40–42], where the models combine physical simulations with experimental data.

In this paper, we propose a generalizable, differentiable framework for solving surrogate-based multifidelity design optimization problems built upon the previous data-driven ASO architectures developed by one of the authors [43, 44]. By employing a reduced-order, compact shape parameterization, we achieve efficient control over high-dimensional design spaces. Fine-tuning neural networks using residual learning enables building a scalable multifidelity surrogate model capable of providing high-fidelity aerodynamic estimations across geometric variations. Finally, we use a gradient-based optimization algorithm to efficiently solve the resulting high-dimensional design problem. The framework development is primarily aimed at assisting the design exploration stage, where different candidate designs are evaluated under different scenarios. Hence, we prioritize the generalizability of the method and its ability to replicate the “trend” of high-fidelity models, which is important in gradient-based optimization. Despite the benefits of using computationally inexpensive models at the design exploration stage, the use of high-fidelity models is advised at the final design stage.

The remaining of this paper is organized as follows. Section II describes the problem of interest, including the ASO formulation and multifidelity data used in the study. Section III introduces the methodology, including the data-driven gradient-based ASO for handling high-dimensional geometry and surrogate model fine-tuning with residual learning for multifidelity data fusion. Section IV presents the results, covering the surrogate model performance and wing shape design optimization. Lastly, we conclude our work in Section V.

II. Problem statement

This study is focused on performing ASO for a commercial aircraft wing, aiming to minimize drag while satisfying lift, moment, and geometric constraints. The optimization process involves evaluating aerodynamic performance metrics—including lift coefficient (C_L), drag coefficient (C_D), and moment coefficient (C_M)—using surrogate models trained on multifidelity data. Specifically, the data are generated using CFD simulations at two fidelity levels, involving a computationally inexpensive coarse (LF) mesh and a more accurate but resource-intensive fine (HF) mesh. In this study, the same CFD solver and setup are used to obtain both LF and HF data. In Section II.A, we present the ASO problem formulation, which encompasses the single-point and multipoint cases performed in this study. In Section II.B, we describe the multifidelity data used in this study.

A. ASO Problem

This study focuses on the ASO of the National Aeronautics and Space Administration (NASA) Common Research Model (CRM) wing geometry [45] with the primary objective of drag reduction. The CRM wing is a standardized aerodynamic test platform widely used in aerospace research to study transonic flow characteristics [46, 47], validate CFD models [48, 49], and investigate design optimizations [17, 50]. Following the setup of NASA Aerodynamic Design Optimization Discussion Group (ADODG)*, we formulate a weighted drag reduction ASO problem as shown in Table 1. Here, we follow the same formulation as the one from our previous work [5].

	Function/variable	Description	Bounds	Quantity
Minimize	$f_{\text{obj}} = \sum_{k=1}^{K} w_k C_{D_k}$	Weighted-average drag coefficient	-	-
With respect to	α_k	Angle of attack, deg	$[1.0, 3.5]$	K
	$\boldsymbol{\lambda}$	Coefficients of wing shape modes	$[\boldsymbol{\lambda}_{\text{lower}}, \boldsymbol{\lambda}_{\text{upper}}]$	50
	$\boldsymbol{\alpha}_{\text{twist}}$	Wing twists, deg	$[-1.0, 1.0]$	7
		Total design variables		$57 + K$
Subject to	$C_{L_k} - C^*_{L_k} = 0$	Lift constraints	-	K
	$C_M \geq -0.17$	Moment constraint at nominal condition	-	1
	$\mathcal{V} \geq \mathcal{V}_{\text{initial}}$	Volume constraint	-	1
	$t \geq 0.98 \times t_{\text{initial}}$	Thickness constraints	-	750
		Total constraints		$752 + K$

Table 1 ASO problem formulation.

The objective is to minimize the weighted average drag coefficient across K flight conditions, where w_k represents the weight for the k-th condition. The weights sum to one, $\sum_{k=1}^{K} w_k = 1$, such that each w_k indicates the relative importance of each flight condition. In this work, we conduct both single-point optimization ($K = 1$) and multipoint optimization ($K > 1$). The implementation of single-point optimization is described in Section IV.B.1, following the formulation of the ADODG case 4.1†. For multipoint optimizations, we consider both three-point (Section IV.B.2) and

*NASA ADODG: `https://sites.google.com/view/mcgill-computational-aerogroup/adodg` (last accessed on 7 March 2025)

†ADODG Case 4.1: `https://drive.google.com/open?id=1O7HHAUioOdiyFG_e63_S6R3TwlwsVM47` (last accessed on 7 March 2025)

nine-point (Section IV.B.3) settings, which are based on ADODG cases 4.4[‡] and 4.5[§], respectively.

The design variables consist of the angle of attack α_k (one value for each flight condition), a vector of wing shape mode coefficients $\boldsymbol{\lambda}$ (from the compact modal parameterization, to be described briefly below), and a vector of wing twists $\boldsymbol{\alpha}_{\text{twist}}$. To reduce the computational burden of such complex optimization cases, we use the compact modal parameterization method originally proposed by Li and Zhang [44] for the wing geometry. The compact modal parameterization approach is derived based on singular value decomposition to efficiently capture the global representation of wing shapes (the method is described in more details in Appendix A). This approach is compatible with `pyGeo`[¶] [51] and its effectiveness has been validated in a series of ASO studies [5, 43, 44]. For the optimization cases presented in this paper, we use 50 modes to represent the wing shapes, with a coefficient assigned to each mode; hence we have $\boldsymbol{\lambda} \in \mathbb{R}^{50}$. We define the mode perturbation coefficient bounds as $\boldsymbol{\lambda}_{\text{lower}} \in \mathbb{R}^{50}$ and $\boldsymbol{\lambda}_{\text{upper}} \in \mathbb{R}^{50}$, with fixed leading-edge and trailing-edge constraints inherently imposed by the modal parameterization. For more details on the determination of the mode coefficient bounds, the choice of 50 modes, and other derivations, readers are referred to the original work by Li *et al.* [52]. Wing twist is controlled by rotating several wing cross-sections around a reference axis, instead of shearing them vertically. In particular, we use seven equally spaced sectional airfoils along the wing span (including the tip section but excluding the root section, which is fixed)—this setting has been commonly adopted in other ASO works, including our previous work [5]. As such, there are seven wing twist design variables ($\boldsymbol{\alpha}_{\text{twist}} \in \mathbb{R}^7$). The design variable bounds for $\alpha_k \forall k$ and $\boldsymbol{\alpha}_{\text{twist}}$ are shown in Table 1.

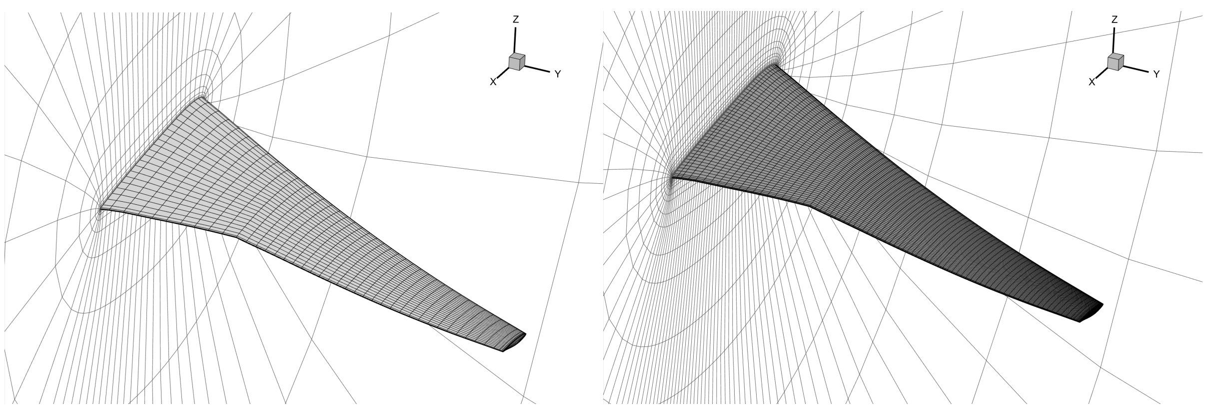
The design constraints include lift, moment, and geometric constraints. There are K equality constraints for C_L (one for each flight condition) to ensure that the target lift is achieved. Considering steady flight, we impose a C_M constraint under nominal condition (Mach number 0.85 and flight altitude 11,740 m). Geometric constraints, including thickness and volume constraints, are adopted from ADODG cases, particularly case 4.5. The lower bound on volume preserves the CRM wing baseline, while a slightly relaxed upper bound allows for a broader design space. There are 750 thickness constraints, defined at 25 chordwise locations (from 1% to 99% of the local chord) and 30 spanwise positions (from root to tip). The lower bound is set at 98% of the baseline thickness so as to approximately maintain the wing's initial structural strength.

B. Multifidelity data

We perform CFD analyses on `ADflow`[‖] [10] to compute the state variables, objective functions, and constraints, where their gradients are evaluated using the adjoint method [13]. `ADflow` solves the RANS equations on multi-block meshes and supports various numerical schemes, including Runge-Kutta, diagonally dominant alternating direction implicit

[‡] ADODG Case 4.4: `https://drive.google.com/open?id=1lqC3wtc5iGWP52c1kpFSp_r1c2P1W0sW` (last accessed on 7 March 2025)
[§] ADODG Case 4.5: `https://drive.google.com/open?id=11o0961oodFGlee9iU2G8HgOUX5TU1TLd` (last accessed on 7 March 2025)
[¶] Repository available online: `https://github.com/mdolab/pygeo.git` (last accessed on 7 March 2025)
[‖] Repository available online `https://github.com/mdolab/adflow.git` (last accessed on 7 March 2025)

(DDADI), and Newton-Krylov methods [53]. Among its built-in turbulence models, we select the Spalart-Allmaras model [54], which is suitable for attached and moderately separated flows. To accelerate computation, `ADflow` employs an approximate Newton–Krylov method in the early iterations to reduce cost, and once the residuals drop below a specified threshold, it switches to the full Newton–Krylov method with explicit Jacobian evaluation to ensure stable convergence [55]. We use `pyHyp`[**] for mesh generation and `IDWarp`[††] for mesh warping. In this study, both LF and HF datasets are generated using the same CFD solver `ADflow` with consistent mesh refinement, ensuring that the differences between fidelity levels are solely due to mesh resolution. By setting different sizes of mesh cells, we generate training datasets with two fidelity levels: a coarse L3 mesh with 56,320 cells for LF data and a finer L2 mesh with 450,560 cells for HF data as shown in Figure 3. A detailed mesh convergence study involving (from fine to coarse) L00, L0, L1, and L2 meshes has been presented by Lyu *et al.* [17]. The L2 mesh has been further investigated in a previous study by one of the authors [44]. Although L3 mesh was not included in the original paper [17], it has been commonly used in works involving computationally intensive optimizations [43, 44, 56] including our previous work [5]. With this consideration, the choice of L2 and L3 meshes is deemed suitable for the multifidelity setup presented in this paper.



(a) L3 mesh (b) L2 mesh

Fig. 3 The LF (L3) and HF (L2) meshes used in the multifidelity framework.

Table 2 summarizes the ranges of key parameters for the flight conditions considered in this work, including Mach number (M), angle of attack, wing twist, and flight altitude (h) to ensure that the derived surrogate models are robust across varying operating conditions. These ranges of flight condition values are evaluated to enable a generalized ASO framework capable of handling multipoint design scenarios. Table 3 outlines the corresponding data used to train and test the surrogate models, including the number of samples, the computational costs of CFD evaluations, and the ranges of aerodynamic coefficients (C_L, C_D, C_M).

The training process involves multiple rounds of design of experiments (DoE) to systematically generate data points. We use the Latin Hypercube Sampling (LHS) to populate the design space, generating 183,075 L3 mesh samples as LF

[**]Repository available online: `https://github.com/mdolab/pyhyp.git`(last accessed on 7 March 2025)
[††]Repository available online: `https://github.com/mdolab/idwarp.git` (last accessed on 7 March 2025)

Variable	Lower	Upper
Mach number	0.82	0.88
Angle of attack [deg]	0.5	3.5
Wing twist [deg]	−1.0	1.0
Altitude [m]	9,000	12,000

Table 2 Value ranges for key flight condition parameters.

Variable	L3 mesh	L2 mesh
Mesh size	56,320	450,560
CPU time / wing [core hour] (CFD)	0.4	6.0
CPU time / wing [core second] (data-driven)	1.1×10^{-4}	5.0×10^{-2}
Training set size	135,108	11,923
Testing set size	47,967	2,981
C_L range	[0.155, 0.728]	[0.218, 0.665]
C_D range	[0.0131, 0.0525]	[0.0135, 0.0466]
C_M range	[-0.298, -0.024]	[-0.265, -0.0466]
L/D range	[9.45, 22.7]	[10.95, 22.14]

Table 3 Low- and high-fidelity model information and setup for the CRM benchmark geometry.

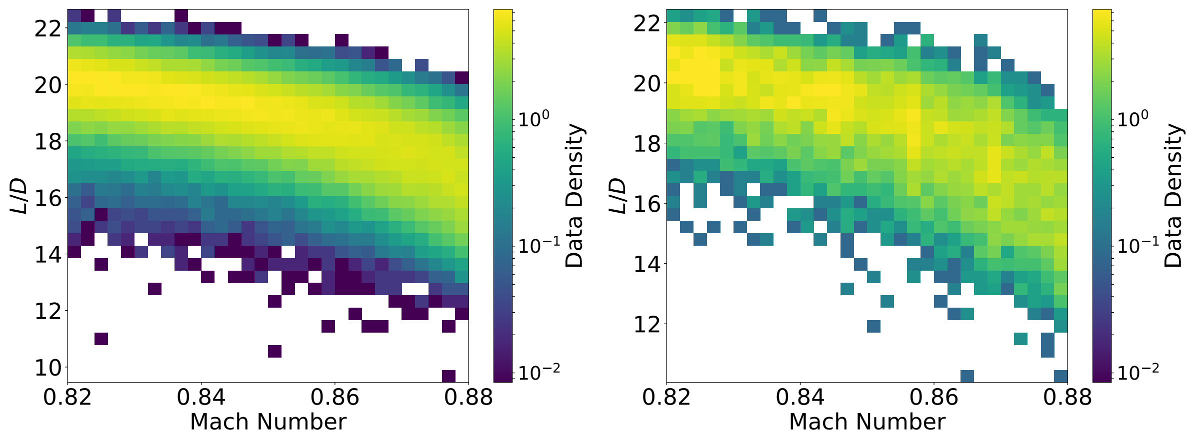
data and 14,904 L2 mesh samples as HF data. These samples are split into training and testing data sets, as listed in Table 3. We follow a previous work by one of the authors [43] to determine the number of LF samples, whereas the number of HF samples is determined empirically by running some numerical experiments. Considering the importance of the number of samples in determining the estimation accuracy and computational efficiency of the multifidelity approach, we perform a systematic study on the model performance under different HF/LF sampling ratio, which will be presented in Section IV.C. Figure 4 shows the distributions of the low- and high-fidelity training data (in log-normalized scale) demonstrating the diversity and coverage of the training data, ensuring the surrogate models can capture the full range of aerodynamic behavior. This comprehensive data set is critical for constructing accurate multifidelity surrogate models, to support efficient and robust ASO.

III. Methodology

In this section, we present the methodology for solving the optimization problem described in Section II. Specifically, Section III.A introduces the data-driven gradient-based ASO framework that is suitable for high-dimensional design problems. In Section III.B, we introduce the developed multifidelity data fusion approach, which relies on a residual learning-based fine-tuned surrogate model.

A. Data-driven gradient-based ASO

A surrogate-based design optimization procedure typically involves physical or numerical experiments, surrogate model construction, infill criteria for updating the sample set, and the evaluations of objective function and constraints [57]. Surrogate-based approach rapidly performs the computational analysis of the physical system and then solves the design



(a) Low-fidelity data

(b) High-fidelity data

Fig. 4 Log-normalized distributions of low- and high-fidelity data.

optimization problem with gradient-based or gradient-free algorithms. To ensure the accuracy and reliability of estimates when constructing and refining the surrogate model, sampling of the full function evaluations should span the design space. In the context of ASO, surrogate-based approaches provide an effective alternative to traditional CFD-driven design optimization methods by constructing predictive models that approximate aerodynamic performance, enabling cost-effective optimization. These methods have been successfully applied to airfoil and wing shape optimization problems [58–60]. Building upon this foundation, data-driven ASO extends surrogate-based methods by leveraging large datasets and advanced machine learning techniques. This allows for further reductions in optimization time while maintaining accuracy, making it a powerful tool for high-dimensional design problems [21, 61–63].

We use the MACH-Aero framework[‡‡] developed by the Multidisciplinary Design Optimization (MDO) Laboratory at the University of Michigan as the main framework to perform ASO cases. MACH-Aero is a gradient-based framework that integrates the CFD solver, geometry parameterization, mesh perturbation, optimizer, and some other base classes relevant to ASO. The tool suite shows good performance in the ADODG standard cases [17, 64, 65] and other state-of-the-art ASO research [66–68]. As mentioned in Section II.B, `ADflow` is used as the CFD solver to analyze the aerodynamic performance of each CRM wing geometry sample under various flight conditions in the design space.

Using the method described above, we perform CFD analyses on LF and HF wing samples derived from global wing shape modes, where the level of fidelity is determined by the mesh quality (i.e., coarse or fine mesh). These samples are then used as the training set for our multifidelity regression models. The proposed framework is illustrated in Figure 5, presented using the eXtended Design Structure Matrix (XDSM) diagram [69][§§].

Gradient-based optimization has proven effective for high-dimensional design problems [70]. After training the multifidelity surrogate network with the LF and HF CFD samples, the design loop becomes fully data-driven: the

[‡‡] MACH-Aero Framework: `https://github.com/mdolab/MACH-Aero` (last accessed on 7 March 2025)
[§§] XDSM: `https://mdolab.engin.umich.edu/wiki/xdsm-overview` (last accessed on 7 March 2025)

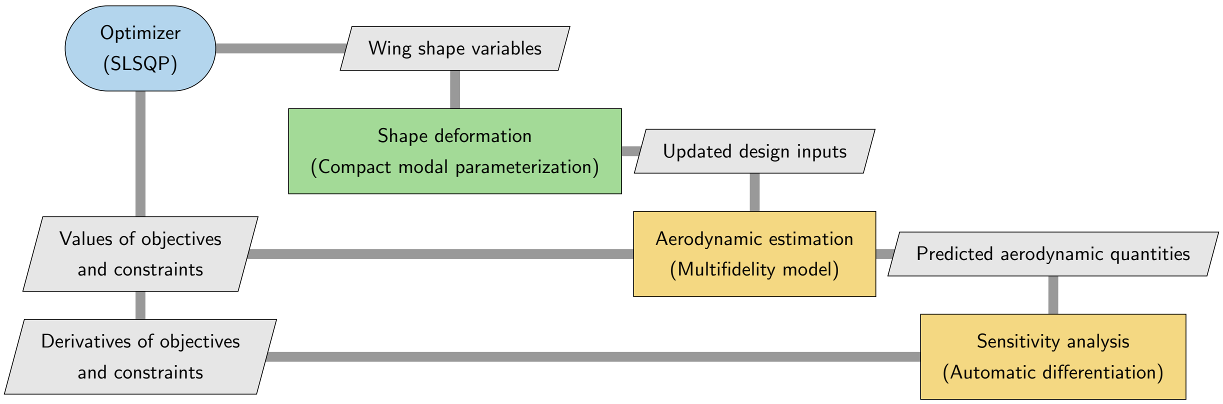


Fig. 5 Data-driven multifidelity gradient-based ASO framework.

surrogate supplies both predictions and their analytic sensitivities, so no additional CFD differentiation is required during optimization. The sensitivity of each predicted coefficient with respect to every design variable x is obtained through automatic differentiation in `PyTorch`[¶¶] or `TensorFlow`[***]. These surrogate-based gradients are then passed to the sequential least squares programming (SLSQP) optimizer available in `pyOptSparse`[†††]. SLSQP efficiently solves constrained nonlinear problems by iteratively minimizing a quadratic sub-problem while enforcing both equality and inequality constraints.

B. Multifidelity model fine-tuning via residual learning

Fine-tuning, as one of transfer learning techniques, provides a flexible multifidelity framework for progressively integrating different fidelity levels, maximizing the utility of low-fidelity data for information provision while leveraging high-fidelity data for correction [71–73]. Residual learning [39] aims to capture the differences between datasets of varying fidelity by learning the differences between high- and low-fidelity information, denoted as Δ, rather than the data themselves. This approach enhances the fine-tuned model's accuracy in fitting high-fidelity data. Note that in this paper, the proposed approach is only applied to multifidelity cases where the LF and HF data are generated using the same CFD solver with consistent mesh refinements. Further investigations and adjustments might be needed for implementations with different multifidelity data sets.

Figure 6 introduces the multifidelity framework presented in this paper. The modal parameterization inset shown in this figure is extracted from Figure 2 of Li and Zhang's paper [43], which is also shown in Figure 17 (Appendix A). Even though only two levels of fidelity are shown here (and used in the optimization cases presented herein), the framework can be extended to include multiple fidelity levels, $i = 1, 2, 3, \dots, n$, where n denotes the highest level of fidelity. The explanation provided below applies to multiple fidelity level cases. Based on the computational effort associated with

[¶¶] `PyTorch`: `https://pytorch.org/` (last accessed on 7 March 2025).
[***] `TensorFlow`: `https://www.tensorflow.org/` (last accessed on 7 March 2025).
[†††] `pyOptSparse` repository: `https://github.com/mdolab/pyoptsparse.git` (last accessed on 7 March 2025).

generating the corresponding aerodynamic data, the data volume N_i should decrease following $N_1 \gg N_2 ... \gg N_n$. We consider a general multifidelity learning problem that involves multiple fidelity levels of data with different fidelity levels of target function. We denote $\mathbf{y}^{(i)}(\mathbf{x})$ as the ground truth (we take CFD solution as ground truth value in this study) and $\hat{\mathbf{y}}^{(i)}(\mathbf{x})$ as the prediction of model $f^{(i)}(\mathbf{x}; \theta^i)$, where θ represents the neural network parameters. The variable $\mathbf{x}$ depends solely on the geometry and is independent of mesh size. The superscript in $\mathbf{x}^i$ serves only to denote that the input is used for training at fidelity level i.

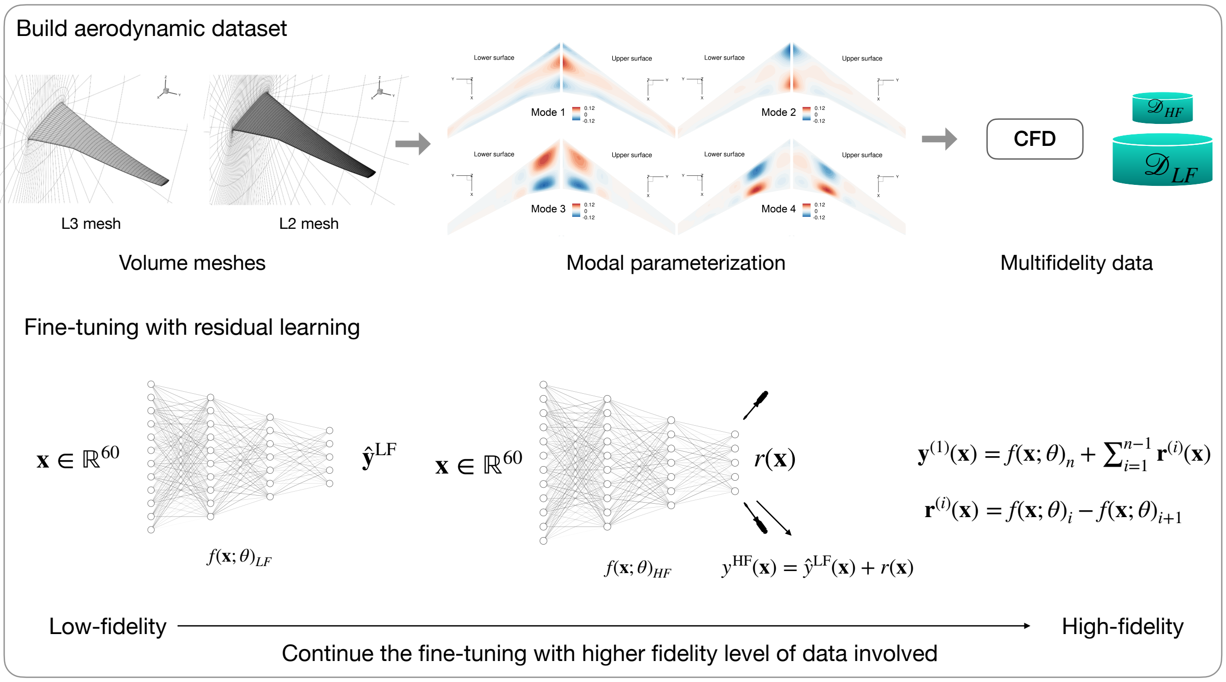


Fig. 6 Surrogate model fine-tuning via residual learning and data generation.

The fine-tuning process is a sequential one starting from the lowest fidelity level, i.e., $i = 1$, and has no limit in terms of the number of fidelity levels involved. The process (for an arbitrary level i) is described as follows. Consider a model $f^i(\mathbf{x}^i; \theta^i)$ trained on the i-th fidelity level of data with training inputs $\mathbf{x}^i$ and CFD estimation ground truth outputs $\mathbf{y}^i(\mathbf{x}^i)$. The loss function for this model is defined as the mean squared error (MSE) between the ground truth and the model predictions:

$$\mathcal{L}^i = \frac{1}{N^i} \sum_{k=1}^{N^i} \left\| \mathbf{y}^i\left(\mathbf{x}_k^i\right) - \hat{\mathbf{y}}^i\left(\mathbf{x}_k^i\right) \right\|_2^2, \tag{1}$$

where $\hat{\mathbf{y}}^i(\mathbf{x}_k^i)$ is the prediction of the model $f^i(\mathbf{x}^i; \theta^i)$ at fidelity level i for the k-th sample, and N^i is the number of data points at the i-th fidelity level. We then use the $(i+1)$-th fidelity level of data to fine-tune this trained model. Instead of learning the prediction outputs, we learn the residual

$$\mathbf{r}^{i \to i+1} = \mathbf{y}^{i+1}\left(\mathbf{x}^{i+1}\right) - \hat{\mathbf{y}}^i\left(\mathbf{x}^{i+1}\right). \tag{2}$$

Given the collected inputs at the $(i+1)$-th fidelity level, $\mathbf{x}^{i+1}$, the corresponding ground truth is denoted by $\mathbf{y}^{i+1}\left(\mathbf{x}^{i+1}\right)$, while the prediction from the i-th fidelity pre-trained model $f^i\left(\mathbf{x}^{i+1};\theta^i\right)$ is denoted by $\hat{\mathbf{y}}^i\left(\mathbf{x}^{i+1}\right)$. Accordingly, the residual $\mathbf{r}^{i\to i+1}$ represents the difference between the lower-fidelity prediction $\hat{\mathbf{y}}^i\left(\mathbf{x}^{i+1}\right)$ and the higher-fidelity ground truth $\mathbf{y}^{i+1}\left(\mathbf{x}^{i+1}\right)$ evaluated at the $(i+1)$-th fidelity inputs.

In fine-tuning with residual learning, the high-fidelity model learns to correct the low-fidelity predictions by modeling the residual between the high-fidelity and low-fidelity outputs. The high-fidelity model's MSE loss function is therefore defined on the residuals,

$$\mathcal{L}^{i+1} = \frac{1}{N^{i+1}}\sum_{k=1}^{N^{i+1}}\left\|\mathbf{r}_k^{i\to i+1} - f^{i+1}\left(\mathbf{x}_k^{i+1};\theta^{i+1}\right)\right\|_2^2, \tag{3}$$

where $\mathbf{r}_k^{i\to i+1} = \mathbf{y}^{i+1}\left(\mathbf{x}_k^{i+1}\right) - \hat{\mathbf{y}}^i\left(\mathbf{x}_k^{i+1}\right)$.

Following this sequential learning procedure, given any design input $\mathbf{x}$, the highest fidelity estimation $\hat{\mathbf{y}}^{(n)}$ can be computed as

$$\hat{\mathbf{y}}^{(n)}(\mathbf{x}) = \hat{\mathbf{y}}^{(1)}(\mathbf{x}) + \sum_{i=1}^{n-1}\hat{\mathbf{r}}^{(i\to i+1)}(\mathbf{x}), \tag{4}$$

with given pre-trained model $f^1\left(\mathbf{x};\theta^1\right),\ldots,f^{n-1}\left(\mathbf{x};\theta^{n-1}\right)$, and the n-th fidelity input samples $\mathbf{x}^n$.

In this paper, we only consider two fidelity levels of data, due to the prohibitive computational cost involved with performing CFD analyses with finer mesh. We differentiate the fidelity levels as 'LF' and 'HF', with the following notations:

1) A small L2 mesh evaluation dataset of size N_{HF} for the high-fidelity function:

$$\mathcal{D}_{HF} = \left\{\left(\mathbf{x}_k, \mathbf{y}_{HF\,k}\right) : 1 \le k \le N_{HF}\right\}, \tag{5}$$

2) A large L3 mesh evaluation dataset of size N_{LF} for the low-fidelity function:

$$\mathcal{D}_{LF} = \left\{\left(\mathbf{x}_k, \mathbf{y}_{LF\,k}\right) : 1 \le k \le N_{LF}\right\}, \tag{6}$$

where $N_{HF} \ll N_{LF}$. Algorithm 1 outlines the fine-tuning procedure via residual learning for a two-fidelity setting, where we denote 'LF' as fidelity level 1 and 'HF' as fidelity level 2. It is worth noting that the input $\mathbf{x}$ is independent of the fidelity level; the notations $\mathbf{x}^{HF}$ and $\mathbf{x}^{LF}$ are solely used to indicate that the inputs are drawn from the HF and LF datasets, respectively, for the purpose of model training.

Algorithm 1 Fine-tuning via residual learning (with two fidelity levels)

1. **Train LF model** $f^{LF}\left(\mathbf{x}^{LF};\theta^{LF}\right)$ on $\mathcal{D}_{LF}$

2. **Compute residuals** $\mathbf{r}^{LF\rightarrow HF} = \mathbf{y}^{HF}\left(\mathbf{x}^{HF}\right) - \hat{\mathbf{y}}^{LF}\left(\mathbf{x}^{HF}\right)$, where $\hat{\mathbf{y}}^{LF}\left(\mathbf{x}^{HF}\right)$ is the prediction using pre-trained model $f^{LF}\left(\mathbf{x}^{HF};\theta^{LF}\right)$ and $\mathbf{y}^{HF}\left(\mathbf{x}^{HF}\right)$ is the ground truth

3. **Train HF model** $f^{HF}\left(\mathbf{x}^{HF};\theta^{HF}\right)$ to predict $\mathbf{r}^{LF\rightarrow HF}$

4. **Final prediction, given any x:** $\hat{\mathbf{y}}^{HF}(\mathbf{x}) = \hat{\mathbf{y}}^{LF}\left(\mathbf{x}^{HF}\right) + \mathbf{r}^{LF\rightarrow HF}$, where $\hat{\mathbf{y}}^{LF}\left(\mathbf{x}^{HF}\right)$ is estimated using model $f^{LF}\left(\mathbf{x};\theta^{LF}\right)$ and $\mathbf{r}^{LF\rightarrow HF}$ is estimated using model $f^{HF}\left(\mathbf{x};\theta^{HF}\right)$

IV. Results and Discussion

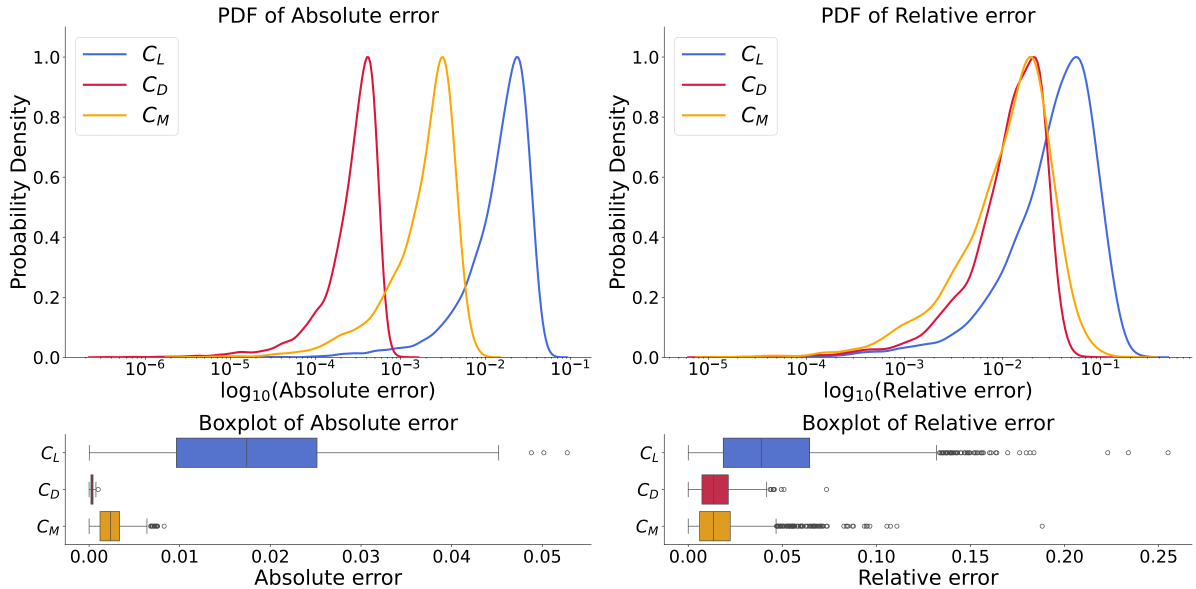
In this section, we present the design optimization results and corresponding discussions, upon presenting the multifidelity model accuracy in Section IV.A. In Section IV.B, we apply the proposed multifidelity data-driven framework to a series of wing shape design optimization cases, including single-point and multipoint optimizations. In Section IV.C, we further analyze the computational cost, performance, and physical consistency of the multifidelity surrogate model, providing insights into its effectiveness and limitations.

A. Multifidelity model accuracy

In this section, we assess the regression accuracy of the surrogate model using aerodynamic data introduced in Section II.B. Independent surrogate models are trained with inputs $\mathbf{x} := \{M, h, \alpha, \alpha_{\text{twist}}, \lambda\}$ and outputs $\mathbf{y} \in \{C_L, C_D, C_M\}$. The multifidelity surrogate model construction is detailed below.

The baseline model, $f(\mathbf{x};\theta)_{LF}$, is a five-layer fully connected multilayer perceptron (MLP) with 100 neurons per layer, trained on LF samples. The value ranges for flight conditions listed in Table 2 are used for the inputs. Figure 7 presents the probability density functions (PDFs) of absolute and relative errors for the model predictions. The PDF illustrates error distributions and concentration, while the box plots summarize central tendency, spread, and outliers. The error distributions shown here are deemed acceptable and consistent with a previous study [43], thereby supporting the surrogate model's applicability to LF aerodynamic modeling.

Building on this model, we refine it with HF data using the method described in Section III.B. Residual learning captures the discrepancies between LF and HF predictions, allowing the model to be adjusted using HF information. Separate fine-tuning models are trained for C_L, C_D, and C_M, with inputs and outputs normalized to $[-1, 1]$ for better numerical conditioning. The LF-trained model first generates predictions using HF inputs, and the residuals are then



(a) Absolute error distribution

(b) Relative error distribution

Fig. 7 Error distributions for C_L, C_D, C_M in LF surrogates.

used as correction targets. This ensures consistency despite differences in LF and HF aerodynamic characteristics.

The training process employs the Adam optimizer [74] with a learning rate of 10^{-4}, which minimizes the mean absolute error (MAE) while monitoring both MAE and MSE. We use early stopping to prevent overfitting by monitoring the validation MAE. Training stops if the validation MAE does not improve by at least 10^{-4} for 10 consecutive epochs. A learning rate scheduler reduces the rate by half after five epochs without improvement, with a lower bound of 10^{-6}. Dropout layers (rate = 0.3) are incorporated to mitigate overfitting.

After training, the model produces HF predictions by integrating learned residuals into LF predictions. The relative error consistently falls within the range of 2% to 4%, which is consistent with established benchmarks for residual-based fine-tuning. This result underscores the model's capability to accurately capture and correct discrepancies between LF and HF data.

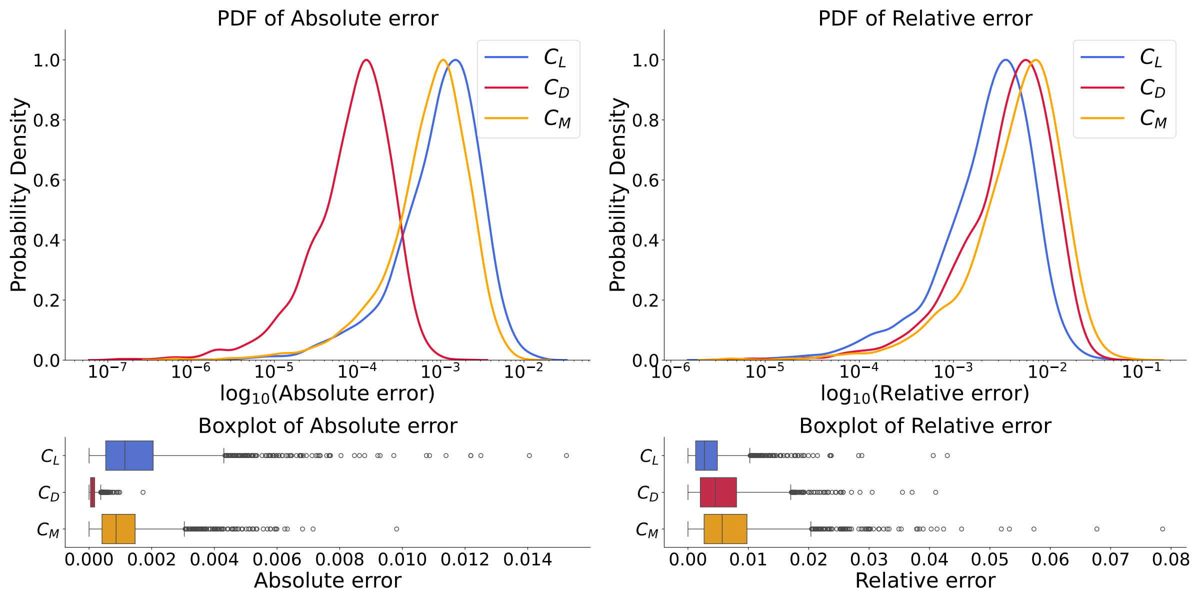
To evaluate model accuracy relative to data variability, we compute the ratio of MAE to standard deviation (denoted as η),

$$\eta = \frac{\text{MAE}}{\sigma} = \frac{\frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|}{\sqrt{\frac{\sum_{i=1}^{N}(y_i-\mu)^2}{N}}}, \tag{7}$$

where μ and σ denote the mean and standard deviation, respectively, and N is the training size. For this metric, lower η values indicate better surrogate model performance. The computed η values for the HF surrogate model predictions are 0.0204, 0.0254, and 0.0343 for C_L, C_D, and C_M, respectively.

We present the PDF and box plots of absolute and relative errors for HF estimations of C_L, C_D, and C_M in

Figure 8. The mean absolute errors are $\left[\delta_{C_L}, \delta_{C_D}, \delta_{C_M}\right] = [0.0037, 0.00024, 0.0019]$, and the mean relative errors are $\left[\epsilon_{C_L}, \epsilon_{C_D}, \epsilon_{C_M}\right] = [0.0091, 0.0102, 0.0181]$. These values indicate a high level of accuracy, well within the typical acceptable thresholds for aerodynamic predictions. The reported errors—0.9% for C_L, 1.0% for C_D, and 1.8% for C_M—demonstrate that the surrogate model effectively approximates high-fidelity CFD values. This accuracy ensures reliability in design optimization workflows, providing a computationally efficient alternative to full CFD evaluations.



(a) Absolute error distribution

(b) Relative error distribution

Fig. 8 Error distributions for C_L, C_D, C_M in HF surrogates.

B. Wing shape design optimization

In this section, we demonstrate the proposed data-driven multifidelity learning approach within the gradient-based ASO framework introduced in Section III to optimize the CRM wing design with both single-point and multipoint objective functions. We conduct three cases to verify the generalizability of the data-driven approach, with all cases optimized using the same surrogate model. All CFD-driven optimizations presented in this section are performed on L2 mesh.

1. Single-point optimization results

We begin with presenting single-point design optimization results, focusing on a wing optimized for the nominal flight condition of a wide-body commercial aircraft at Mach 0.85 and an altitude of 11,740 m. As outlined in Table 1, the design objective is to minimize the drag coefficient C_D with $K = 1$ under this flight condition. This problem formulation follows the ADODG case 4.1, as mentioned in Section II.A.

Figure 9 shows the comparison between the optimized wings obtained using CFD-driven and our data-driven MF approach. In these plots, all C_D values are reported as drag counts ($C_{D_{\text{counts}}} = C_D \times 10^4$). The planform view shown on the upper left-hand side of Figure 9 shows the C_P contours of two optimized wings, along with the corresponding C_L, C_D, and C_M values. On the right, we display the C_P distributions and sectional airfoil variations for selected wing sections at six spanwise locations. These locations, marked as A to F, represent the positions at 2.35%, 26.7%, 55.7%, 69.5%, 82.8%, and 94.4% spanwise distances from the wing root, respectively. These results show that at Mach 0.85, the CFD-driven optimized wing exhibits a higher lift-to-drag (L/D) ratio of 23.15 compared to 22.94 obtained with the data-driven MF optimized wing. We place an isosurface at the shock sensor value of one to visualize the shock region. A front view with the shock surface visualization and the physical thickness distribution of the wing is shown below the planform view. The lower left-hand side of Figure 9 illustrates the spanwise distributions of lift, twist, and thickness-to-chord ratio (t/c), along with the reference elliptical lift distribution (in gray). Elliptical lift distribution tends to minimize induced drag and enhance lift efficiency, and both optimized wings show close resemblance to this reference. The benefit of using compact modal parameterization method can be observed in the smooth variations of twist and t/c distribution along the span, because this method inherently smooths the trade-offs between volume and drag by distributing thickness changes more gradually, avoiding abrupt transitions at discrete locations (e.g., the Yehudi break). Overall, these results are consistent with those of our previous work on an operation-aware multipoint ASO [5], which has similar setup and formulation.

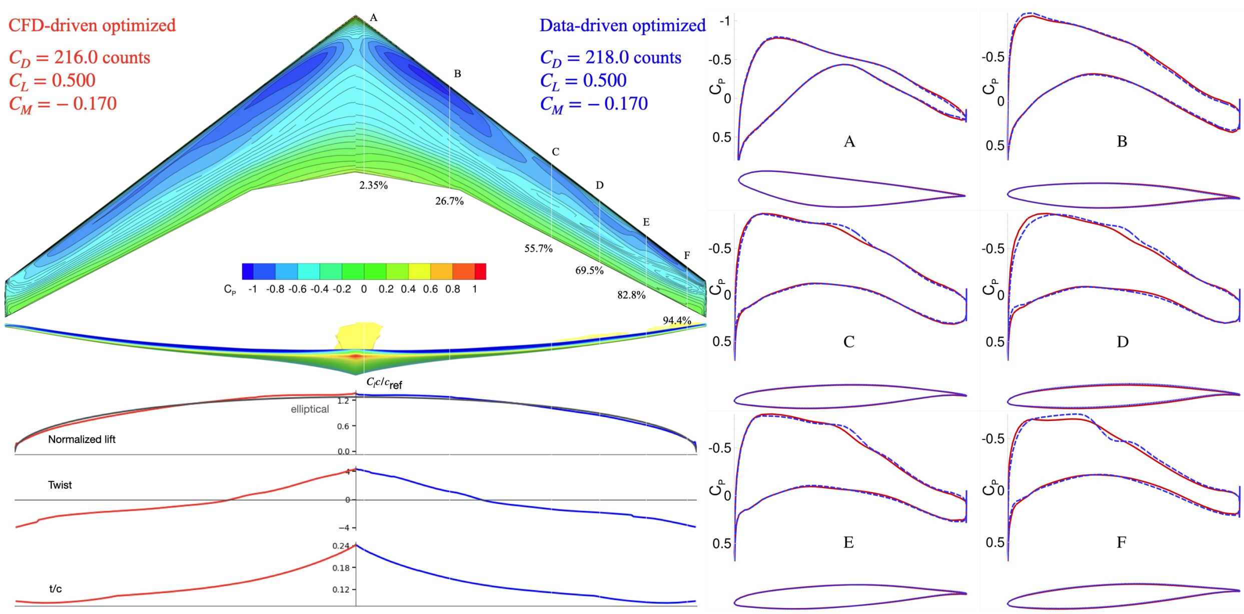


Fig. 9 Design comparison for the single-point optimization case: CFD-driven vs data-driven optimized wings.

2. Three-point optimization results

Next, we solve a three-point wing optimization problem, following the framework of ADODG case 4.4. The objective is to minimize the weighted average of C_D over a range of Mach numbers at a fixed aircraft weight, with altitude adjusted to maintain a constant C_L across the specified Mach number range. This optimization case is performed to validate the capability of our method in handling multipoint optimization under varying Mach numbers and altitudes. The flight conditions and their corresponding weights (for a multipoint objective function with $K = 3$) are listed in Table 4.

Design point	Weight	Mach number	C_L constraint	Flight altitude (m)
1	0.25	0.82	0.50	11,285
2	0.50	0.85	0.50	11,740
3	0.25	0.88	0.50	12,180

Table 4 Flight conditions and their corresponding weights for the three-point optimization.

We present the optimized-wing comparison results in Figure 10; the visualization style matches that of Figure 9. Both optimized wings are evaluated at the nominal condition. The data-driven MF approach attains a drag level of 224.2 counts, only 2.9 counts higher than the CFD-optimized wing. Because multipoint optimization seeks to balance performance over a range of conditions, complete shock suppression at any single evaluation point is not necessarily anticipated. With respect to lift span-loading, wing twist, and t/c, the two designs are largely similar, demonstrating improved aerodynamic efficiency while remaining within the prescribed thickness constraints that act as a proxy for structural requirements, and thus exhibit comparable overall performance. The sectional C_P distributions are broadly similar, although the data-driven MF wing shows a slightly weaker suction peak near the leading edge and localized pressure variations around mid-span.

3. Nine-point optimization results

We define our nine-point design optimization problem based on ADODG case 4.5 to assess our method's performance in a large-scale optimization case. Our problem setup follows the same structure as outlined in Table 1, with $K = 9$. The flight conditions and corresponding weights for the nine design points are summarized in Table 5.

The comparison between the optimized wing designs obtained through CFD-driven and data-driven MF approaches for this nine-point design optimization is presented in Figure 11, with a similar format with that of Figures 9 and 10. The two wings are evaluated under the nominal condition with $C_L = 0.5$. The wing optimized using the data-driven MF approach achieves a drag count of 224.8, which is 3.6 counts higher than the result from the CFD-driven optimization. Although the pitching moment coefficient ($C_M = -0.166$) of the data-driven MF optimized wing satisfies the constraint of $C_M > -0.17$, there remains room for further improvement, considering the objective function of the optimization is formulated in a Lagrangian framework. The inability of the data-driven MF approach to reach an

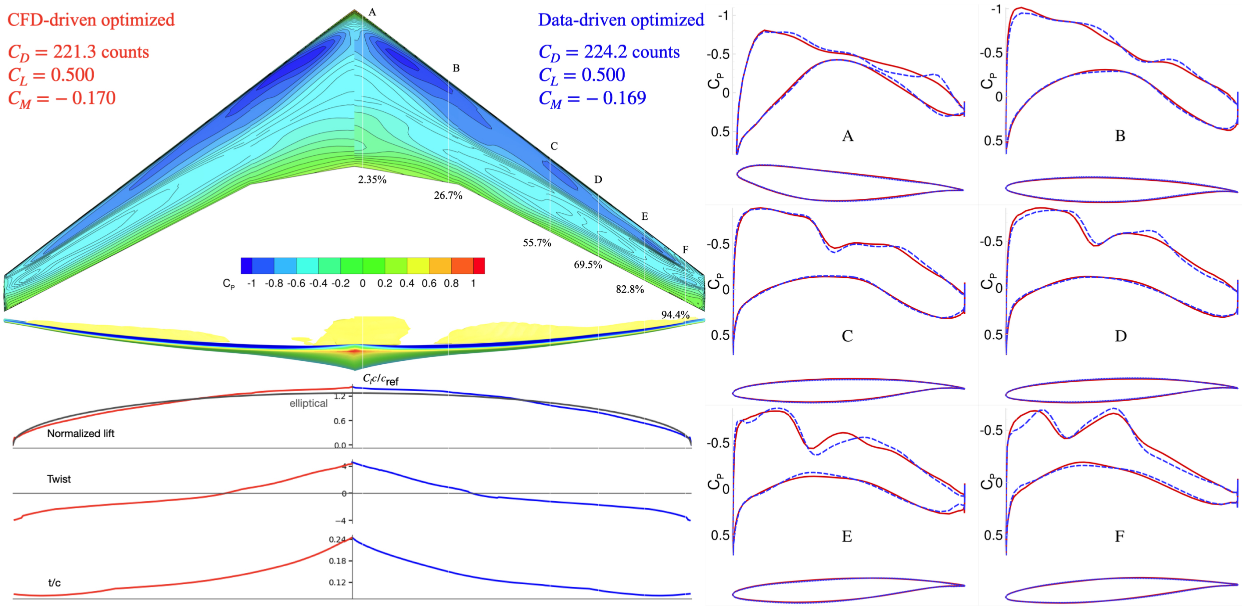


Fig. 10 Design comparison for the three-point optimization case: CFD-driven vs data-driven optimized wings.

Design point	Weight	Mach number	C_L constraint	Flight altitude (m)
1	0.125	0.82	0.483	11,740
2	0.250	0.82	0.537	11,740
3	0.125	0.82	0.591	11,740
4	0.250	0.85	0.450	11,740
5	0.500	0.85	0.500	11,740
6	0.250	0.85	0.550	11,740
7	0.125	0.88	0.420	11,740
8	0.250	0.88	0.466	11,740
9	0.125	0.88	0.513	11,740

Table 5 Flight conditions and their corresponding weights for the nine-point optimization.

optimal solution may be due to the excessively high dimensionality of the surrogate model, causing the optimization search to become trapped in local optima. Both optimized wings exhibit similar spanwise distributions of lift, twist, and t/c, indicating that the data-driven MF approach achieves mechanical characteristics comparable to the CFD-driven optimized wing. Concerning wing shocks, as discussed in Section IV.B.2 for the three-point optimization, the nine-point optimization prioritizes overall drag reduction and thus cannot match single-point optimization performance at any specific aerodynamic condition. Regarding the C_P distribution across six slices, the data-driven MF and CFD-driven optimized wings show generally similar chordwise trends. The data-driven MF wing has slightly higher magnitude of negative pressure near the leading edge but overall exhibits weaker negative pressure compared to the CFD-driven optimized wing. This difference is also evident across the entire wing surface, where the CFD-driven optimized wing consistently achieves greater negative pressure regions.

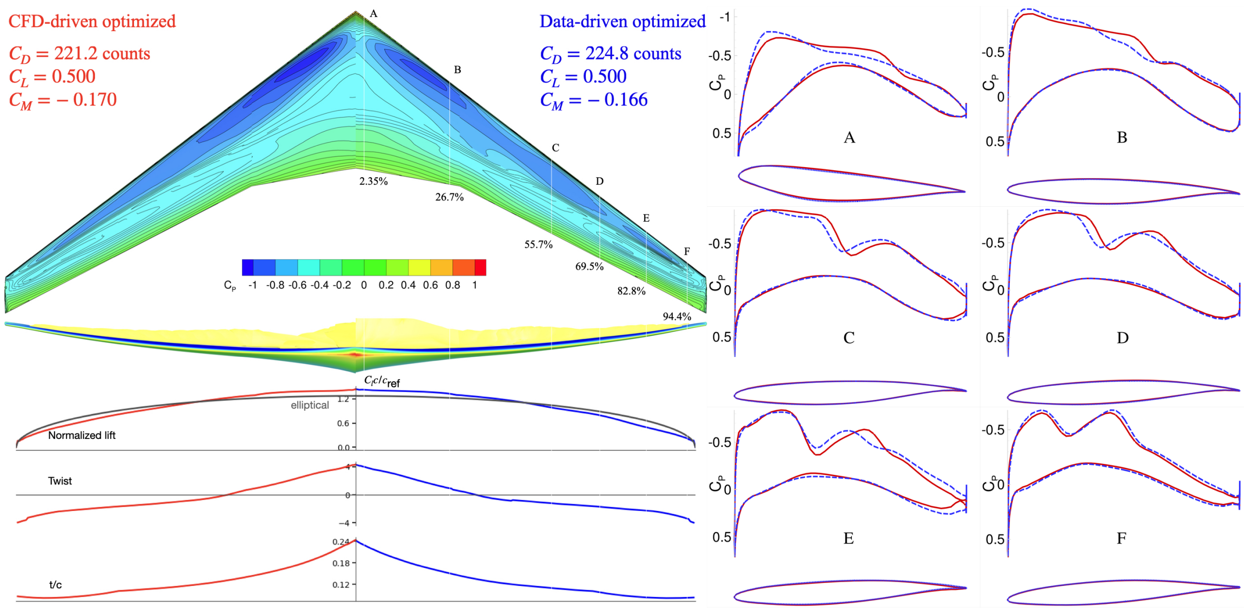


Fig. 11 Design comparison for the nine-point optimization case: CFD-driven vs data-driven optimized wings.

4. Optimization performance discussion

In this section, we briefly discuss the overall performance of the three optimization cases presented above, particularly in terms of how well the data-driven MF approach replicates the full CFD-driven optimization and the computational efficiency improvement it offers. Table 6 summarizes the CFD-driven and data-driven optimization results. The ΔC_D values shown in this table refer to the difference in $C_{D_{\text{counts}}}$ between the baseline and optimized designs (with negative values indicating C_D reduction in the optimized wings, which is desirable). The single-point case is optimized at the nominal design condition, while the three- and nine-point cases use the design points listed in Tables 4 and 5, respectively. All aerodynamic performance metrics presented in this table are evaluated at the nominal design condition of Mach 0.85

and altitude 11,740 m. The design objective function, defined as the weighted drag summation in drag counts, is 229.1 for the CFD-driven three-point optimization and 234.2 for the data-driven approach. For the nine-point optimization, it is 229.95 for the CFD-driven approach and 234.90 for the data-driven approach. Comparing these results with the drag comparison at the nominal condition, we observe the same pattern—that the CFD-driven optimization results generally yield lower drag. All data-driven MF optimizations presented in Sections IV.B.1 to IV.B.3 employ the same surrogate model using the approach introduced in Section III.B. Overall, the results show that data-driven optimization consistently achieves drag reduction performance comparable to CFD-driven methods, across both single-point and multipoint design problems.

Mesh	**Approach**	C_L	$C_{D_{\text{counts}}}$ (ΔC_D)	C_M
L3	Baseline wing	0.500	234.0	-0.190
L3	Single-fidelity CFD-driven optimized wing	0.500	229.0 (-5.0)	-0.170
L3	Single-fidelity data-driven optimized wing	0.498	229.0 (-5.0)	-0.170
L2	Data-driven optimized L3 wing, re-analyzed on L2 mesh	0.497	233.0	-0.144
L2	Baseline wing	0.502	228.0	-0.182
L2	Single-point data-driven MF optimized wing	0.500	218.0 (-10.0)	-0.170
L2	Single-point CFD-driven optimized wing	0.500	216.0 (-12.0)	-0.170
L2	Three-point data-driven MF optimized wing	0.500	224.2 (-3.8)	-0.169
L2	Three-point CFD-driven optimized wing	0.500	221.3 (-6.7)	-0.170
L2	Nine-point data-driven MF optimized wing	0.500	224.8 (-3.2)	-0.166
L2	Nine-point CFD-driven optimized wing	0.500	221.2 (-6.8)	-0.170

Table 6 Aerodynamic performance summary across mesh fidelities and optimization problem setups.

All CFD-driven and data-driven MF optimizations are executed on a CPU server. Each CPU is configured as an Intel Core i7-10700K processor, featuring a 3.8 GHz base clock speed, eight cores, and 16 MB of cache. For CFD-driven design optimization, we employ parallel computing utilizing 64 cores, resulting in computation times of 6.3 hours for single-point, 11.33 hours for three-point, and 50.87 hours for nine-point optimization. For data-driven MF design optimization, we use a single core, achieving computation times of 2.1 minutes for single-point, 5.3 minutes for three-point, and 6.2 minutes for nine-point optimization.

While it is acknowledged that constructing a surrogate model requires additional computational effort compared to solving a single problem using a CFD adjoint solver, this upfront cost is justified by the model's versatility across a wide range of design scenarios. In the cases presented above, for instance, the same surrogate model is used to solve both single-point and multipoint optimization problems. This highlights the generalizability of our method, where a single trained model can be reused across different tasks with significantly reduced computational cost. Considering the importance of computational cost (both during training and optimization) in assessing the effectiveness of a surrogate model, we systematically evaluate how the upfront cost of data generation scales with problem complexity and affects performance; the detailed discussion is presented in Section IV.C. Additionally, although CFD adjoint-based solvers are powerful and effective, implementing adjoint computations often demands considerable development effort. Our study shows that, even in the absence of adjoint information, surrogate models can still deliver competitive optimization results.

C. Computational cost and model performance

This subsection investigates the computational efficiency of the surrogate model by analyzing its performance as the proportions of HF and LF data in the training set are systematically varied. The goal is to evaluate how effectively the model leverages data of different fidelity to balance computational efficiency and predictive accuracy.

Figure 12 illustrates the relationship between the mean relative error (MRE) of aerodynamic coefficients ((by comparing MF predictions against HF data) and HF data and the associated computational cost. The figure shows that as the computational cost increases, the MRE of MF prediction continues to decrease. However, beyond a certain threshold of computational cost, the rate of reduction in MRE gradually diminishes. To further investigate the relationship between the rate of MRE reduction and computational cost, we introduce the computational cost efficiency ρ, defined as the rate

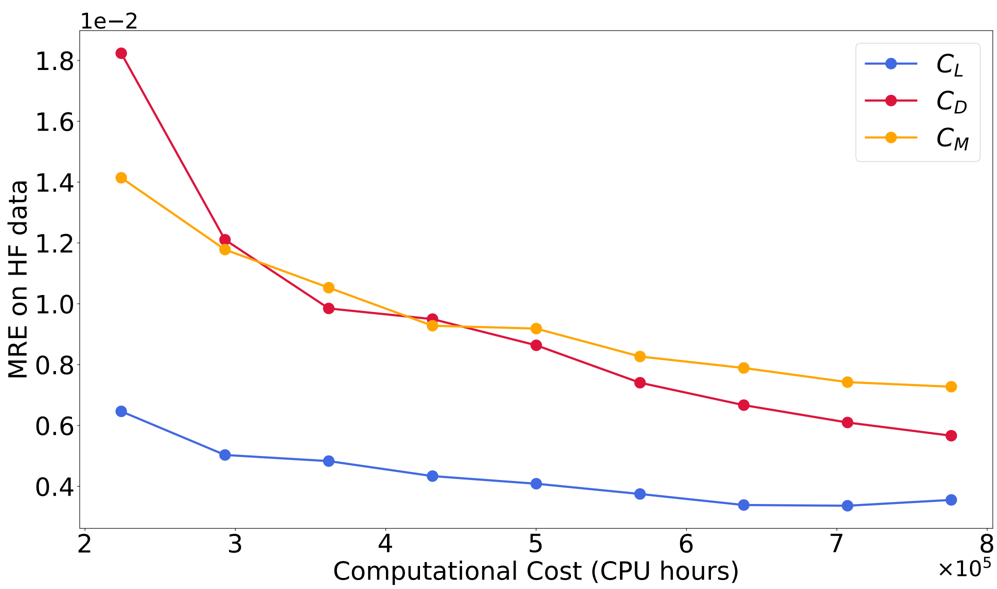


Fig. 12 The MRE of MF prediction with respect to high-fidelity data for C_L, C_D and C_M versus computational cost.

of change in MRE per unit of computational cost:

$$\rho = \frac{\Delta \text{MRE}}{\text{Computational cost}}, \tag{8}$$

which quantifies the trade-off between improvements in accuracy and the computational resources required. The relationship between ρ and computational cost is shown in Figure 13, which shows that ρ decreases with increasing computational cost, indicating a higher computational cost is required to achieve a unit reduction in MRE. Beyond a computational cost of 4×10^5 CPU hours, ρ approaches a plateau at a low value, implying that further increases in computational resources yield diminishing returns in reducing MRE. Therefore, additional data generation under such conditions is neither efficient nor justified.

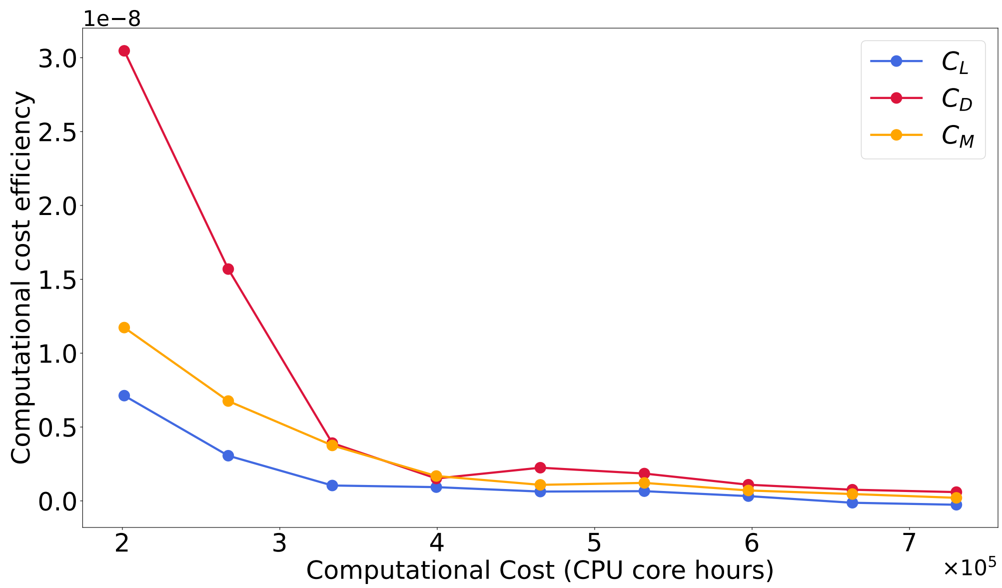


Fig. 13 The computational cost efficiency ρ for C_L, C_D and C_M predictions versus computational cost.

To have a comprehensive analysis, we compare the MRE of MF predictions with respect to HF data with different data demand ratios by conducting a series of comparative experiments, as shown in Figure 14. In particular, we train the model using different ratios of LF and HF data, following the approach outlined in Section II.B. As an example, we can train the model with 30% LF data and 70% HF data, where the percentage is taken with respect to the training dataset size (for the respective fidelity level) shown in Table 3. As such, with 135,108 LF training data and 11,923 HF training data, 30% LF data and 70% HF data correspond to 40,532 LF data and 8,346 HF data, respectively. In Figure 14, we denote the "Mixed HF/LF Region" as the configuration where both HF and LF data are used in varying proportions. The "Full LF with Different HF Region" refers to the setting where the entire LF dataset is employed and the HF data ratio is varied for fine-tuning purposes. From Figure 14, we can observe that using mostly LF data effectively determines the upper bound of the MRE. Increasing the proportion of HF data enables further reduction in MRE, indicating that while LF data control the baseline error level, HF data contribute to improving model accuracy. The value of LF data primarily lies in its ability to provide "structural guidance" when HF data are scarce, offering crucial priors and trend information in the absence or limited availability of HF data.

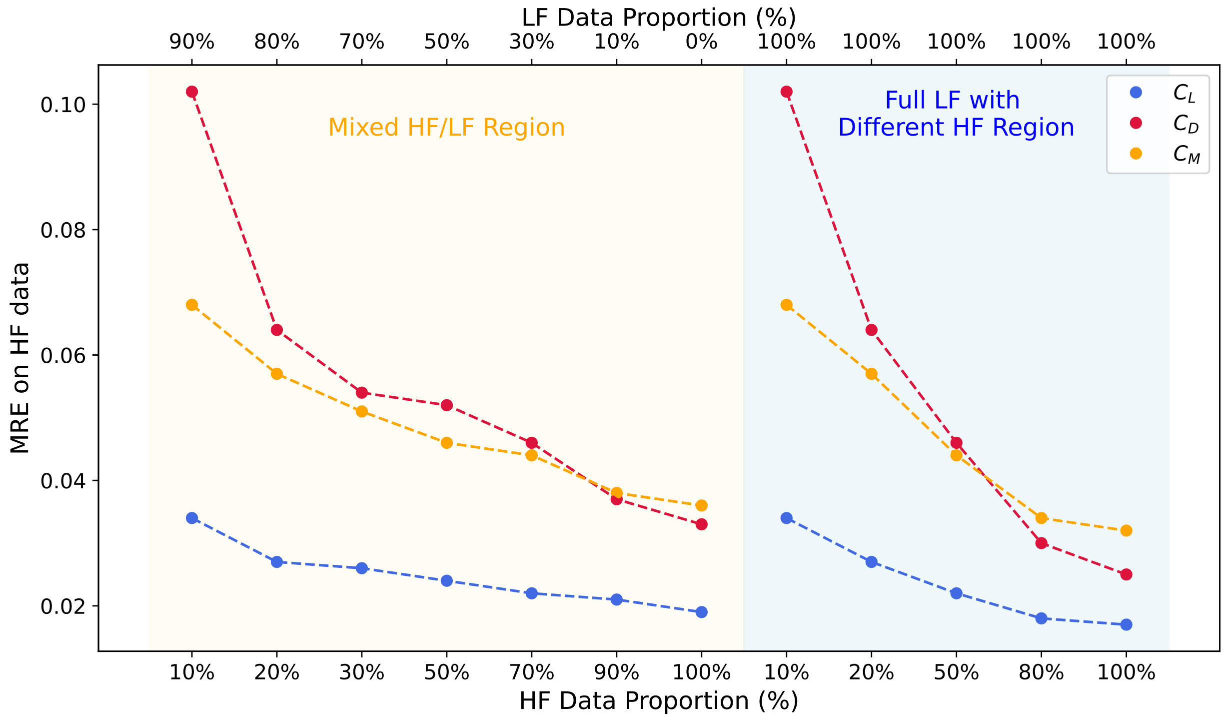


Fig. 14 The MRE for C_L, C_D and C_M with different LF and HF data demand ratios.

We plot Figure 15 to further investigate the model performance when all LF data are used, but with different HF demand ratio (corresponding to the light-blue region in the right-hand side of Figure 14). We observe that continuously increasing the proportion of HF data indeed leads to improved model accuracy; however, the rate of improvement diminishes relative to the growing computational cost required to obtain HF data. These results highlight the need to carefully balance model accuracy against the simulation cost required for data generation. Specifically, the figure illustrates that beyond 80% HF data, the accuracy gains become marginal.

Nevertheless, in the context of the high-fidelity ASO problem considered in this work, even minor improvements in model accuracy can translate into meaningful aerodynamic performance gains. Therefore, we choose to utilize all available data to perform the data-driven MF-based optimization presented in this paper. For other optimization scenarios, however, practitioners should carefully consider the trade-off between model accuracy and the computational cost of data generation.

Given the flexibility of the pre-training and fine-tuning processes, simulation data do not need to be generated upfront. Instead, they can be acquired progressively in response to the evolving needs of the model. This enables a more adaptive and efficient data acquisition process, where HF samples are selectively added as higher accuracy becomes necessary. Strategies such as active learning [75] can provide theoretical foundations for this approach, which we aim to incorporate in future works. It is important to note, however, that the specific approach presented in this paper is derived

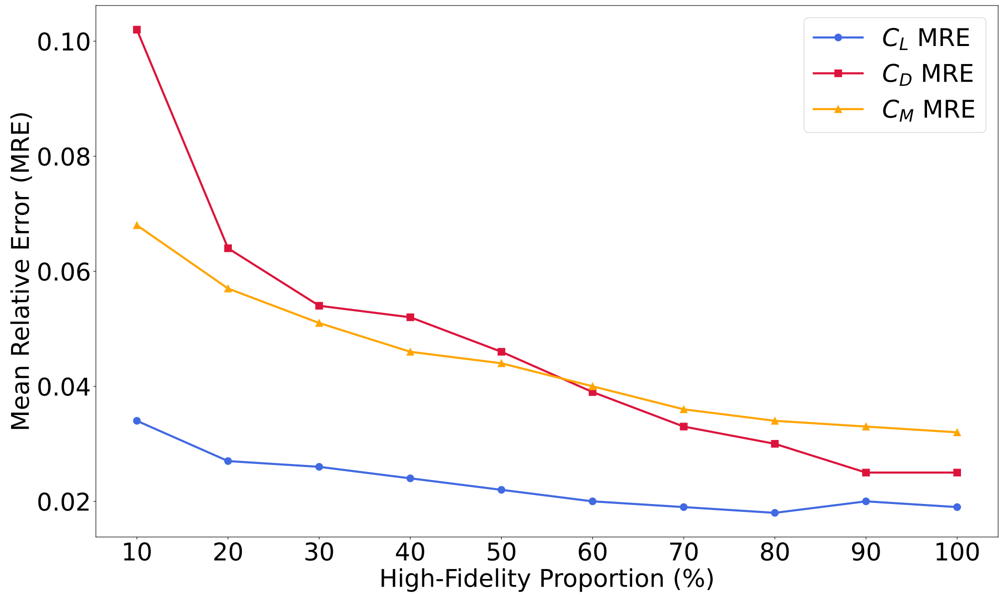


Fig. 15 The MRE with different HF ratio, with full LF data.

and validated for data obtained using the same CFD solver, geometry parameterization, and structured mesh distribution that ensure consistency across fidelity levels. As mentioned in Section III.B, further investigations would be needed to extend the framework's capability to admit other multifidelity data sources.

Lastly, we analyze the break-even point (BEP) to assess the effectiveness of data-driven MF models, compared to the CFD-driven ones, with different levels of model complexity, by borrowing the concept from the economic theory [76]. For this purpose, we define a dimensionless problem complexity C as the sum over individual optimization tasks:

$$C = \sum_{i=1}^{N_{\text{opt}}} N_{\text{points}}^{(i)} \cdot \log\left(N_{\text{design\_iters}}^{(i)}\right), \tag{9}$$

where $N_{\text{points}}^{(i)}$ and $N_{\text{design\_iters}}^{(i)}$ denote the number of operating points and design iterations associated with the i-th optimization problem, respectively. N_{opt} signifies the total number of optimization cases that need to be performed using the same model. We visualize the relationship between problem complexity C and computational cost (including model training time) in Figure 16, from which we can identify the BEP where adopting data-driven MF methods becomes more computationally efficient than running full high-fidelity CFD-driven optimizations. When using data-driven MF models, most of the computational cost investment occurs during the training process, whereas performing each optimization only takes a small fraction of the total computational time. As such, the MF line (in blue) appears almost horizontal in Figure 16. The CFD-driven approach, on the other hand, does not benefit from such gentle scaling, due to the substantial computational investment required at every optimization run. With this consideration, we can then define the areas of "profit" and "loss" from adopting the data-driven MF approach. "Profit" denotes the region where data-driven MF requires lower computational cost than high-fidelity CFD-driven optimization, indicating a net gain in computational efficiency. Conversely, "loss" denotes the region where MF methods are more computationally expensive, i.e., where the number of full CFD runs required for MF training is higher than that of the full CFD-driven optimization. To illustrate this "profit" and "loss" concept, we consider running a number of ASOs, each with a different optimization problem formulation (e.g., by considering different design missions or constraints), at the early aircraft design exploration stage. In particular, we take a practical example from an operation-aware multipoint ASO case, where it is common to involve a dozen or even up to 30 design points in order to account for a wide range of operational flight conditions [5]. When performing 12 optimization problems each with 30 design points and approximately 50 design iterations, for instance, the resulting complexity is $C \approx 1{,}408$, which lies slightly below the BEP shown in Figure 16. By contrast, with 15 optimization problems of the same size, the complexity increases to $C \approx 1{,}759$, which surpasses the BEP. These examples illustrate that, for large-scale and operation-aware optimization tasks, the proposed multifidelity data-driven method becomes competitive with, and eventually more cost-effective than, adjoint- or gradient-based approaches. As the scale of the optimization problem increases, the MF data-driven approach gradually demonstrates its advantage. This observation echoes our initial motivation: when the problem dimensionality grows and more optimization tasks are considered, the computational efficiency of the MF data-driven method becomes increasingly valuable.

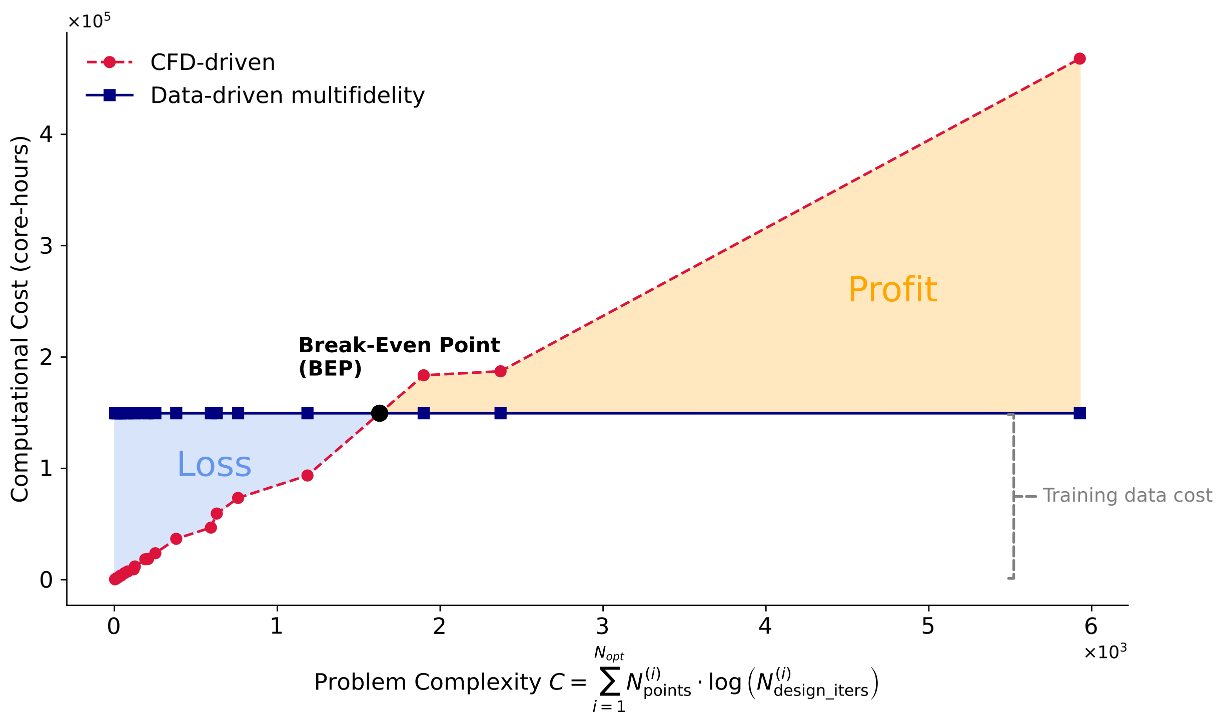


Fig. 16 Break-even analysis to determine the benefit of performing MF-based optimizations, compared to the CFD-driven ones.

V. Conclusion

This work introduces a scalable multifidelity design optimization framework tailored for addressing high-dimensional, high-fidelity ASO challenges. The key aim is to enable rapid design exploration, where a number of design optimizations need to be performed under different scenarios. Hence, we put an emphasis on computational efficiency and generalizability in the framework development, while maintaining a high level of accuracy. The framework combines a deep-learning-based surrogate model with a gradient-based nonlinear programming solver to effectively handle complex high-dimensional geometries. The surrogate model leverages fine-tuning through a residual learning approach to achieve robust and scalable multifidelity data fusion. Experimental results reveal that incorporating an appropriate amount of data significantly enhances the aerodynamic performance of the optimized wing. Furthermore, the incremental integration of high-fidelity data demonstrates substantial potential for further improving the design quality.

With 183,075 low-fidelity data points and 14,904 high-fidelity data points, the MF surrogate model's regression performance for C_L, C_D and C_M achieved an η value of 2.04%, 2.54% and 3.43%, respectively, where η represents the ratio of MAE to the standard deviation of the CFD simulation data. We showed that a single fine-tuned MF surrogate model exhibited strong generalization capabilities by effectively solving multiple high-fidelity design optimization problems without the need for retraining. Once trained, this surrogate-based infrastructure enabled ultra-fast design optimization, completing each evaluation within 5.0×10^{-2} seconds for the L2 wing mesh. This performance was achieved on a standard desktop equipped with an Intel Core i7-10700K processor (3.8 GHz base clock, 8 cores, 16 MB cache), demonstrating the model's efficiency and practicality. These results highlighted the surrogate model's potential as a scalable, real-time solution for high-fidelity aerodynamic design optimization.

The results presented in this paper also illustrated the trade-off of the data-driven MF approach. Whilst this approach reduced the wall-clock time required for design optimization, it yielded slightly higher final drag than the CFD-driven optimization, with differences of 2.0, 2.9, and 3.6 drag counts for the single-, three-, and nine-point problems evaluated at nominal condition, respectively. This level of discrepancy is deemed sufficient for the purpose of design exploration. As the design converges towards the final one, high-fidelity models can play a bigger role for higher accuracy.

The computational benefit of using this MF surrogate model in design optimization was assessed by evaluating the break-even point against using the full CFD model, where it was evident that the benefit increased with problem complexity. In particular, the BEP analysis showed that the surrogate model started to offer a net advantage over conventional solvers at a problem complexity (in terms of the number of points, optimization problems, and design iterations) of approximately 1.5×10^3. Beyond this threshold, the surrogate model demonstrated increasing computational advantages, making it particularly suitable for large-scale or multipoint design optimization tasks.

It is worth to mention that the adjoint-based approach for design optimization has demonstrated significant success in solving various transonic and RANS-based computational problems. Furthermore, its influence continues to expand to various applications, as it remains highly effective in addressing high-dimensional design optimization challenges, consistently achieving optimal solutions with outstanding performance. However, there may be cases where the adjoint-based approach cannot be rapidly adapted to new problems or the CFD solver does not support adjoint computation. In such cases, our data-driven MF method will offer advantages. This is because our method ensures that efficient design optimization can still be achieved through sampling of multifidelity data even in the absence of design priors, when rapid deployment of adjoint methods is not feasible, or when the CFD solver and design modules are not well-integrated.

The significance of this work lies in providing a new perspective for modern industrial design: by leveraging cumulative data, design optimization can be achieved, and with increased data volume and the inclusion of higher-fidelity data, further enhancements in aerodynamic performance can be realized. Although generating such data may incur additional computational costs, this highly parallelized data generation approach is undoubtedly more efficient and streamlined than traditional serial gradient-based optimization methods.

VI. Acknowledgment

This work was supported by the Hong Kong Research Grants Council General Research Fund (Project No. 16207523). The computational fluid dynamics simulations were performed on the Beijing Paratera supercomputing platform.

A. Compact modal parameterization

Compact modal parameterization achieves shape control of the entire wing by managing the reduced-order modeling modes of the sectional airfoil shapes. The wing shape modes are derived using singular value decomposition (SVD). Suppose the wing sampling process perturbs the z coordinates of n sectional airfoil locations in m wing samples. The perturbation of wing samples based on original CRM wing's n sectional slices can be denoted as $\boldsymbol{G} = \left\{z_j^{(k)} - z_j^{(CRM)} : 1 \leq j \leq n, 1 \leq k \leq m\right\}$, as shown in Equation 10. Here, j and k denote the index of sectional airfoil location and wing sample, respectively. We then decompose $\boldsymbol{G} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}$, in which the columns of $\boldsymbol{U}$ represents the wing shape modes and $\boldsymbol{\Sigma}$ contains the diagonal entries in a descending order. $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthogonal matrices with a transposed relationship. With this setup, we can use the columns of $\boldsymbol{U}$ as wing shape modes. Since these modes are sorted in terms of their importance, a reduced order model can be achieved by using the first few modes instead of the full matrix, depending on the problem requirement and computational budget. Figure 17 shows the first eight wing shape modes derived from the compact modal parameterization method, which are the same as the one used in our previous work [5].

$$
\boldsymbol{G} = \begin{bmatrix} z_1^1 - z_1^{\mathrm{CRM}} & z_1^2 - z_1^{\mathrm{CRM}} & \cdots & z_1^m - z_1^{\mathrm{CRM}} \\ z_2^1 - z_2^{\mathrm{CRM}} & z_2^2 - z_2^{\mathrm{CRM}} & \cdots & z_2^m - z_2^{\mathrm{CRM}} \\ \vdots & \vdots & \ddots & \vdots \\ z_n^1 - z_n^{\mathrm{CRM}} & z_n^2 - z_n^{\mathrm{CRM}} & \cdots & z_n^m - z_n^{\mathrm{CRM}} \end{bmatrix} \tag{10}
$$

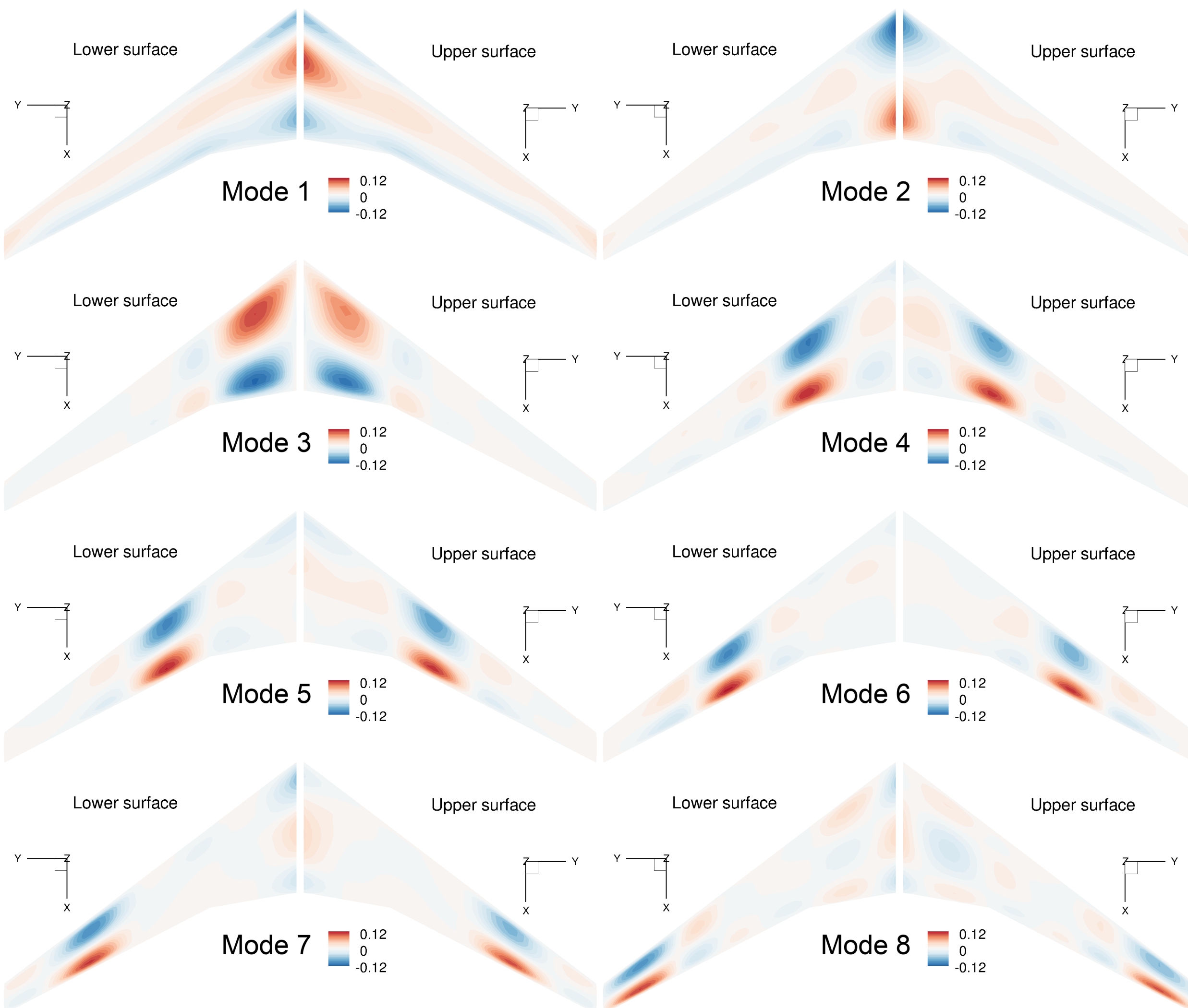


Fig. 17 The first eight wing shape modes obtained from the compact modal parameterization of the CRM wing.

References

[1] Martins, J. R. R. A., "Perspectives on aerodynamic design optimization," *AIAA Scitech 2020 Forum. Special Session: Aerodynamic Design Optimization Discussion Group I.*, Orlando, FL, 2020.

[2] Guo, J., Zhao, N., Sun, L., and Zhang, S., "Modular based flexible digital twin for factory design," *Journal of Ambient Intelligence and Humanized Computing*, Vol. 10, 2019, pp. 1189–1200. https://doi.org/10.1007/s12652-018-0953-6.

[3] Liem, R. P., Kenway, G. K. W., and Martins, J. R. R. A., "Multimission Aircraft Fuel Burn Minimization via Multipoint Aerostructural Optimization," *AIAA Journal*, Vol. 53, No. 1, 2015, pp. 104–122. https://doi.org/10.2514/1.J052940.

[4] Liem, R. P., Mader, C. A., and Martins, J. R. R. A., "Surrogate models and mixtures of experts in aerodynamic performance prediction for mission analysis," *Aerospace Science and Technology*, Vol. 43, 2015, pp. 126–151. https://doi.org/10.1016/j.ast.2015.02.019.

[5] Yang, A., Lyu, Y., Li, J., and Liem, R. P., "Operation-Aware Aircraft Wing Design Using Cluster-Based Multipoint Aerodynamic Shape Optimization," *Journal of Aircraft*, Vol. 62, No. 6, 2025, pp. 1531–1547. https://doi.org/10.2514/1.C038291.

[6] Giles, M. B., and Pierce, N. A., "An Introduction to the Adjoint Approach to Design," *Flow, Turbulence and Combustion*, Vol. 65, 2000, pp. 393–415. https://doi.org/10.1023/A:1011430410075.

[7] Queipo, N. V., Haftka, R. T., Shyy, W., Goel, T., Vaidyanathan, R., and Tucker, P. K., "Surrogate-based analysis and optimization," *Progress in Aerospace Sciences*, Vol. 41, No. 1, 2005, pp. 1–28. https://doi.org/10.1016/j.paerosci.2005.02.001.

[8] Economon, T. D., Palacios, F., Copeland, S. R., Lukaczyk, T. W., and Alonso, J. J., "SU2: An Open-Source Suite for Multiphysics Simulation and Design," *AIAA Journal*, Vol. 54, No. 3, 2016, pp. 828–846. https://doi.org/10.2514/1.J053813.

[9] Jasak, H., "OpenFOAM: Open source CFD in research and industry," *International Journal of Naval Architecture and Ocean Engineering*, Vol. 1, No. 2, 2009, pp. 89–94. https://doi.org/10.2478/IJNAOE-2013-0011.

[10] Mader, C. A., Kenway, G. K. W., Yildirim, A., and Martins, J. R. R. A., "ADflow—An Open-Source Computational Fluid Dynamics Solver for Aerodynamic and Multidisciplinary Optimization," *Journal of Aerospace Information Systems*, Vol. 17, No. 9, 2020, pp. 508–527. https://doi.org/10.2514/1.I010796.

[11] Jameson, A., Vassberg, J. C., and Shankaran, S., "Aerodynamic-Structural Design Studies of Low-Sweep Transonic Wings," *Journal of Aircraft*, Vol. 47, No. 2, 2010, pp. 505–514. https://doi.org/10.2514/1.42775.

[12] Nadarajah, S. K., "The Discrete Adjoint Approach to Aerodynamic Shape Optimization," Ph.D. thesis, Stanford University, Stanford, CA, 2003.

[13] Kenway, G. K. W., Mader, C. A., He, P., and Martins, J. R. R. A., "Effective adjoint approaches for computational fluid dynamics," *Progress in Aerospace Sciences*, Vol. 110, 2019, p. 100542. https://doi.org/10.1016/j.paerosci.2019.05.002.

[14] He, P., Mader, C. A., Martins, J. R. R. A., and Maki, K. J., "An aerodynamic design optimization framework using a discrete adjoint approach with OpenFOAM," *Computers & Fluids*, Vol. 168, 2018, pp. 285–303. https://doi.org/10.1016/j.compfluid.2018.04.012.

[15] Economon, T. D., Alonso, J. J., Albring, T. A., and Gauger, N. R., "Adjoint Formulation Investigations of Benchmark Aerodynamic Design Cases in SU2," *35th AIAA Applied Aerodynamics Conference*, 2017. https://doi.org/10.2514/6.2017-4363.

[16] Reuther, J., Jameson, A., Farmer, J., Martinelli, L., and Saunders, D., "Aerodynamic Shape Optimization of Complex Aircraft Configurations via an Adjoint Formulation," *34th AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, USA, 1996. https://doi.org/10.2514/6.1996-94.

[17] Lyu, Z., and Martins, J. R. R. A., "Aerodynamic Design Optimization Studies of a Blended-Wing-Body Aircraft," *Journal of Aircraft*, Vol. 51, No. 5, 2014, pp. 1604–1617. https://doi.org/10.2514/1.C032491.

[18] Abergo, L., Morelli, M., and Guardone, A., "Aerodynamic shape optimization based on discrete adjoint and RBF," *Journal of Computational Physics*, Vol. 477, 2023, p. 111951. https://doi.org/10.1016/j.jcp.2023.111951.

[19] Fernández-Godino, M. G., Park, C., Kim, N. H., and Haftka, R. T., "Issues in Deciding Whether to Use Multifidelity Surrogates," *AIAA Journal*, Vol. 57, No. 5, 2019, pp. 2039–2053. https://doi.org/10.2514/1.J057750.

[20] Forrester, A. I. J., Keane, A. J., and Bressloff, N. W., "Design and Analysis of "Noisy" Computer Experiments," *AIAA Journal*, Vol. 44, No. 10, 2006, pp. 2331–2339. https://doi.org/10.2514/1.20068.

[21] Li, J., Du, X., and Martins, J. R. R. A., "Machine learning in aerodynamic shape optimization," *Progress in Aerospace Sciences*, Vol. 134, 2022, p. 100849. https://doi.org/10.1016/j.paerosci.2022.100849.

[22] Han, Z.-H., and Görtz, S., "Hierarchical kriging model for variable-fidelity surrogate modeling," *AIAA Journal*, Vol. 50, No. 9, 2012, pp. 1885–1896. https://doi.org/10.2514/1.J051354.

[23] Peherstorfer, B., Willcox, K., and Gunzburger, M., "Survey of Multifidelity Methods in Uncertainty Propagation, Inference, and Optimization," *SIAM Review*, Vol. 60, No. 3, 2018, pp. 550–591. https://doi.org/10.1137/16M1082469.

[24] Raul, V., and Leifsson, L., "Surrogate-based aerodynamic shape optimization for delaying airfoil dynamic stall using kriging regression and infill criteria," *Aerospace Science and Technology*, Vol. 111, 2021, p. 106555. https://doi.org/10.1016/j.ast.2021.106555.

[25] Bartoli, N., Lefebvre, T., Dubreuil, S., Olivanti, R., Priem, R., Bons, N., Martins, J. R. R. A., and Morlier, J., "Adaptive modeling strategy for constrained global optimization with application to aerodynamic wing design," *Aerospace Science and Technology*, Vol. 90, 2019, pp. 85–102. https://doi.org/10.1016/j.ast.2019.03.041.

[26] Raissi, M., and Karniadakis, G., "Deep Multi-fidelity Gaussian Processes," *arXiv preprint arXiv:1604.07484*, 2016.

[27] Bonfiglio, L., Perdikaris, P., Brizzolara, S., and Karniadakis, G. E., "Multi-fidelity optimization of super-cavitating hydrofoils," *Computer Methods in Applied Mechanics and Engineering*, Vol. 332, 2018, pp. 63–85. https://doi.org/10.1016/j.cma.2017.12.009.

[28] Kanagawa, M., Hennig, P., Sejdinovic, D., and Sriperumbudur, B., "Gaussian Processes and Kernel Methods: A Review on Connections and Equivalences," 2018. https://doi.org/10.48550/arXiv.1807.02582, arXiv (Cornell Univ.), Ithaca, NY, USA, arXiv:1807.02582 [stat.ML], submitted July 6, 2018.

[29] Martins, J. R. R. A., "Aerodynamic design optimization: challenges and perspectives," *Computers and Fluids*, Vol. 239, 2022, p. 105391. https://doi.org/10.1016/j.compfluid.2022.105391.

[30] Meng, X., and Karniadakis, G. E., "A composite neural network that learns from multi-fidelity data: application to function approximation and inverse PDE problems," *Journal of Computational Physics*, Vol. 401, 2020, p. 109020. https://doi.org/10.1016/j.jcp.2019.109020.

[31] Guo, M., Manzoni, A., Amendt, M., Conti, P., and Hesthaven, J. S., "Multi-fidelity regression using artificial neural networks: efficient approximation of parameter-dependent output quantities," *Computer Methods in Applied Mechanics and Engineering*, Vol. 389, 2022, p. 114378. https://doi.org/10.1016/j.cma.2021.114378.

[32] Zhang, X., Xie, F., Ji, T., Zhu, Z., and Zheng, Y., "Multi-fidelity deep neural network surrogate model for aerodynamic shape optimization," *Computer Methods in Applied Mechanics and Engineering*, Vol. 373, 2021, p. 113485. https://doi.org/10.1016/j.cma.2020.113485.

[33] Meng, X., Babaee, H., and Karniadakis, G. E., "Multi-fidelity bayesian neural networks: algorithms and applications," *Journal of Computational Physics*, Vol. 438, 2021, p. 110361. https://doi.org/10.1016/j.jcp.2021.110361.

[34] Pratt, L. Y., "Discriminability-based transfer between neural networks," *Proceedings of the 6th International Conference on Neural Information Processing Systems*, Morgan Kaufmann Publishers, San Francisco, CA, USA, 1992, pp. 204–211. https://doi.org/10.5555/2987061.2987087.

[35] Erhan, D., Bengio, Y., Courville, A., Manzagol, P.-A., Vincent, P., and Bengio, S., "Why Does Unsupervised Pre-Training Help Deep Learning?" *Journal of Machine Learning Research*, Vol. 11, 2010, pp. 625–660. URL https://dl.acm.org/doi/10.5555/1756006.1756025.

[36] Kou, J., Ning, C., and Zhang, W., "Transfer Learning for Flow Reconstruction Based on Multifidelity Data," *AIAA Journal*, Vol. 60, No. 10, 2022, pp. 5993–6006. https://doi.org/10.2514/1.J061647.

[37] Wang, Z., Liu, X., Yu, J., Wu, H., and Lyu, H., "A General Deep Transfer Learning Framework for Predicting the Flow Field of Airfoils with Small Data," *Computers and Fluids*, Vol. 251, 2023, p. 105738. https://doi.org/10.1016/j.compfluid.2022.105738.

[38] Shen, Y., Patel, H. C., Xu, Z., and Alonso, J. J., "Application of Multi-Fidelity Transfer Learning with Autoencoders for Efficient Construction of Surrogate Models," *AIAA Scitech 2024 Forum*, American Institute of Aeronautics and Astronautics, 2024. https://doi.org/10.2514/6.2024-0013, session: Special Session: Applied Surrogate Modeling I.

[39] Xia, W., Qin, C., and Hazan, E., "Chain of lora: Efficient fine-tuning of language models via residual learning," *arXiv preprint arXiv:2401.04151*, 2024.

[40] Julian, R., Swanson, B., Sukhatme, G. S., Levine, S., Finn, C., and Hausman, K., "Never stop learning: The effectiveness of fine-tuning in robotic reinforcement learning," *arXiv preprint arXiv:2004.10190*, 2020.

[41] Gao, J., Michelis, M. Y., Spielberg, A., and Katzschmann, R. K., "Sim-to-Real of Soft Robots with Learned Residual Physics," *IEEE Robotics and Automation Letters*, Vol. 9, No. 10, 2024, pp. 8523–8530. https://doi.org/10.1109/LRA.2024.3446287.

[42] Jiang, Y., Wang, C., Zhang, R., Wu, J., and Fei-Fei, L., "Transic: Sim-to-real policy transfer by learning from online correction," *arXiv preprint arXiv:2405.10315*, 2024.

[43] Li, J., and Zhang, M., "Data-Based Approach for Wing Shape Design Optimization," *Aerospace Science and Technology*, Vol. 112, 2021, p. 106639. https://doi.org/10.1016/j.ast.2021.106639.

[44] Li, J., and Zhang, M., "Adjoint-Free Aerodynamic Shape Optimization of the Common Research Model Wing," *AIAA Journal*, Vol. 59, No. 6, 2021, pp. 1990–2000. https://doi.org/10.2514/1.J059921.

[45] Vassberg, J., Dehaan, M., Rivers, M., and Wahls, R., "Development of a Common Research Model for Applied CFD Validation Studies," *26th AIAA Applied Aerodynamics Conference*, Honolulu, Hawaii, USA, 2008, p. 6919. https://doi.org/10.2514/6.2008-6919, session: APA-19, AIAA CFD Drag Prediction Workshop Updates.

[46] Rivers, M., and Dittberner, A., "Experimental Investigations of the NASA Common Research Model in the NASA Langley National Transonic Facility and NASA Ames 11-Ft Transonic Wind Tunnel," *49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, Orlando, Florida, USA, 2011. https://doi.org/10.2514/6.2011-1126, invited paper.

[47] Togiti, V., Eisfeld, B., and Brodersen, O., "Turbulence Model Study for the Flow Around the NASA Common Research Model," *Journal of Aircraft*, Vol. 51, No. 4, 2014. https://doi.org/10.2514/1.C032609.

[48] Clark, A. M., Slotnick, J. P., Taylor, N. J., and Rumsey, C. L., "Requirements and Challenges for CFD Validation within the High-Lift Common Research Model Ecosystem," *AIAA AVIATION 2020 Forum*, Virtual Event, 2020. https://doi.org/10.2514/6.2020-2772, session: Special Session—High Lift Common Research Model Ecosystem.

[49] Vassberg, J. C., DeHaan, M. A., Rivers, M. S., and Wahls, R. A., "Retrospective on the Common Research Model for Computational Fluid Dynamics Validation Studies," *Journal of Aircraft*, Vol. 55, No. 4, 2018. https://doi.org/10.2514/1.C034906.

[50] Palacios, F., Economon, T. D., and Alonso, J. J., "Large-Scale Aircraft Design Using SU2," *53rd AIAA Aerospace Sciences Meeting*, Kissimmee, Florida, USA, 2015. https://doi.org/10.2514/6.2015-1946, session: Special Session—Aerodynamic Design Optimization of Benchmark Cases III.

[51] Hajdik, H. M., Yildirim, A., Wu, E., Brelje, B. J., Seraj, S., Mangano, M., Anibal, J. L., Jonsson, E., Adler, E. J., Mader, C. A., Kenway, G. K. W., and Martins, J. R. R. A., "pyGeo: A geometry package for multidisciplinary design optimization," *Journal of Open Source Software*, Vol. 8, No. 87, 2023, p. 5319. https://doi.org/10.21105/joss.05319.

[52] Li, J., Zhang, M., Martins, J. R., and Shu, C., "Efficient aerodynamic shape optimization with deep-learning-based geometric filtering," *AIAA Journal*, Vol. 58, No. 10, 2020, pp. 4243–4259. https://doi.org/10.2514/1.J059254.

[53] Kenway, G. K., Kennedy, G. J., and Martins, J. R., "Scalable parallel approach for high-fidelity steady-state aeroelastic analysis and adjoint derivative computations," *AIAA Journal*, Vol. 52, No. 5, 2014, pp. 935–951. https://doi.org/10.2514/1.J052255.

[54] Spalart, P., and Allmaras, S., "A One-Equation Turbulence Model for Aerodynamic Flows," *30th Aerospace Sciences Meeting and Exhibit*, Reno, NV, USA, 1992. https://doi.org/10.2514/6.1992-439.

[55] Yildirim, A., Kenway, G. K., Mader, C. A., and Martins, J. R. R. A., "A Jacobian-free approximate Newton–Krylov startup strategy for RANS simulations," *Journal of Computational Physics*, Vol. 397, 2019, p. 108741. https://doi.org/10.1016/j.jcp.2019.06.018.

[56] Hwang, J. T., Jasa, J. P., and Martins, J. J. R. R. A., "High-fidelity Design-allocation Optimization of a Commercial Aircraft Maximizing Airline Profit," *Journal of Aircraft*, Vol. 56, No. 3, 2019, pp. 1164–1178. https://doi.org/10.2514/1.C035082.

[57] Forrester, A. I. J., and Keane, A. J., "Recent advances in surrogate-based optimization," *Progress in Aerospace Sciences*, Vol. 45, No. 1–3, 2009, pp. 50–79. https://doi.org/10.1016/j.paerosci.2008.11.001.

[58] Han, Z., Xu, C., Zhang, L., Zhang, Y., Zhang, K., and Song, W., "Efficient Aerodynamic Shape Optimization Using Variable-Fidelity Surrogate Models and Multilevel Computational Grids," *Chinese Journal of Aeronautics*, Vol. 33, No. 1, 2020, pp. 31–47. https://doi.org/10.1016/j.cja.2019.05.001.

[59] Li, J., Bouhlel, M. A., and Martins, J. R. R. A., "Data-based approach for fast airfoil analysis and optimization," *AIAA Journal*, Vol. 57, No. 2, 2019, pp. 581–596. https://doi.org/10.2514/1.J057129.

[60] Han, Z.-H., Chen, J., Zhang, K.-S., Xu, Z.-M., Zhu, Z., and Song, W.-P., "Aerodynamic Shape Optimization of Natural-Laminar-Flow Wing Using Surrogate-Based Approach," *AIAA Journal*, Vol. 56, No. 7, 2018, pp. 2579–2593. https://doi.org/10.2514/1.J056661.

[61] Li, J., He, S., Zhang, M., Martins, J. R. R. A., and Khoo, B. C., "Physics-Based Data-Driven Buffet-Onset Constraint for Aerodynamic Shape Optimization," *AIAA Journal*, Vol. 60, No. 8, 2022, pp. 4775–4788. https://doi.org/10.2514/1.J061519.

[62] Magalhães Júnior, J. M., Halila, G. L. O., Kim, Y., Khamvilai, T., and Vamvoudakis, K. G., "Intelligent data-driven aerodynamic analysis and optimization of morphing configurations," *Aerospace Science and Technology*, Vol. 121, 2022, p. 107388. https://doi.org/10.1016/j.ast.2022.107388.

[63] Moni, A., Yao, W., and Malekmohamadi, H., "Data-Driven Reduced Order Modelling for Aerodynamic Shape Optimisation," *AIAA SCITECH 2023 Forum*, National Harbor, MD, USA and Online, 2023. https://doi.org/10.2514/6.2023-2201, session: Metamodeling, Reduced Order Models, and Approximation Methods IIII.

[64] Bons, N. P., He, X., Mader, C. A., and Martins, J. R. R. A., "Multimodality in Aerodynamic Wing Design Optimization," *AIAA Journal*, Vol. 57, No. 3, 2019, pp. 1004–1018. https://doi.org/10.2514/1.J057294.

[65] He, X., Li, J., Mader, C. A., Yildirim, A., and Martins, J. R. R. A., "Robust aerodynamic shape optimization—from a circle to an airfoil," *Aerospace Science and Technology*, Vol. 87, 2019, pp. 48–61. https://doi.org/10.1016/j.ast.2019.01.051.

[66] Lyu, Z., Kenway, G. K. W., and Martins, J. R. R. A., "Aerodynamic Shape Optimization Investigations of the Common Research Model Wing Benchmark," *AIAA Journal*, Vol. 53, No. 4, 2015, pp. 968–985. https://doi.org/10.2514/1.J053318.

[67] Secco, N. R., Jasa, J. P., Kenway, G. K. W., and Martins, J. R. R. A., "Component-Based Geometry Manipulation for Aerodynamic Shape Optimization with Overset Meshes," *AIAA Journal*, Vol. 56, No. 9, 2018, pp. 3667–3679. https://doi.org/10.2514/1.J056550.

[68] Kenway, G. K. W., and Martins, J. R. R. A., "Buffet-Onset Constraint Formulation for Aerodynamic Shape Optimization," *AIAA Journal*, Vol. 55, No. 6, 2017, pp. 1930–1947. https://doi.org/10.2514/1.J055172.

[69] Lambe, A. B., and Martins, J. R. R. A., "Extensions to the design structure matrix for the description of multidisciplinary design, analysis, and optimization processes," *Structural and Multidisciplinary Optimization*, Vol. 46, No. 2, 2012, pp. 273–284. https://doi.org/10.1007/s00158-012-0763-y.

[70] Martins, J. R. R. A., and Ning, A., *Engineering Design Optimization*, Cambridge University Press, Cambridge, UK, 2021. https://doi.org/10.1017/9781108980647, (Section 1.4.1).

[71] Song, D. H., and Tartakovsky, D. M., "Transfer Learning on Multifidelity Data," *Journal of Machine Learning for Modeling and Computing*, Vol. 3, No. 1, 2022, pp. 31–47. https://doi.org/10.1615/JMachLearnModelComput.2021038925.

[72] Zhang, Z., Ye, Q., Yang, D., Wang, N., and Meng, G., "A multi-fidelity transfer learning strategy based on multi-channel fusion," *Journal of Computational Physics*, Vol. 506, 2024, p. 112952. https://doi.org/10.1016/j.jcp.2024.112952.

[73] Jiang, S., and Durlofsky, L. J., "Use of multifidelity training data and transfer learning for efficient construction of subsurface flow surrogate models," *Journal of Computational Physics*, Vol. 474, 2023, p. 111800. https://doi.org/10.1016/j.jcp.2022.111800.

[74] Kingma, D. P., and Ba, J. L., "Adam: A Method for Stochastic Optimization," *International Conference on Learning Representations (ICLR)*, 2015.

[75] Settles, B., "Active Learning Literature Survey," Computer Sciences Technical Report 1648, University of Wisconsin–Madison, Madison, WI, 2009.

[76] Cafferky, M. E., and Wentworth, J., *Breakeven Analysis: The Definitive Guide to Cost-Volume-Profit Analysis*, 2^{nd} ed., Business Expert Press, New York, 2014. (Chapter 2, p. 9).