

REPORT DOCUMENTATION PAGE			Form Approved OMB NO. 0704-0188		
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 03-01-2020		2. REPORT TYPE Final Report		3. DATES COVERED (From - To) 11-Sep-2015 - 10-Mar-2019	
4. TITLE AND SUBTITLE Final Report: Scalable Environment for Quantification of Uncertainty and Optimization in Industrial Applications (SEQUOIA)			5a. CONTRACT NUMBER W911NF-15-2-0116		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHORS			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAMES AND ADDRESSES Stanford University 3160 Porter Drive Suite 100 Stanford, CA 94304 -8445			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211			10. SPONSOR/MONITOR'S ACRONYM(S) ARO		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) 68061-MA-DRP.1		
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Juan Alonso
a. REPORT UU	b. ABSTRACT UU	c. THIS PAGE UU			19b. TELEPHONE NUMBER 650-723-9954

RPPR Final Report

as of 06-Jan-2020

Agency Code:

Proposal Number: 68061MADRP
INVESTIGATOR(S):

Agreement Number: W911NF-15-2-0116

Name: Juan J. Alonso
Email: jjalonso@stanford.edu
Phone Number: 6507239954
Principal: Y

Organization: **Stanford University**

Address: 3160 Porter Drive, Stanford, CA 943048445

Country: USA

DUNS Number: 009214214

EIN: 941156365

Report Date: 10-Jun-2019

Date Received: 03-Jan-2020

Final Report for Period Beginning 11-Sep-2015 and Ending 10-Mar-2019

Title: Scalable Environment for Quantification of Uncertainty and Optimization in Industrial Applications (SEQUOIA)

Begin Performance Period: 11-Sep-2015

End Performance Period: 10-Mar-2019

Report Term: 0-Other

Submitted By: Juan Alonso

Email: jjalonso@stanford.edu

Phone: (650) 723-9954

Distribution Statement: 1-Approved for public release; distribution is unlimited.

STEM Degrees:

STEM Participants:

Major Goals: The “Scalable Environment for Quantification of Uncertainty and Optimization in Industrial Applications” (SE- QUOIA) project provides an integrated plan for performing uncertainty quantification (UQ) and design under uncertainty (DUU) that aggressively pursues new frontiers in scale and complexity. In particular, the coordinated investments in this effort will create advancements in scalable forward and inverse UQ algorithms and the rigorous quantification of model inadequacy, providing the primary foundation for the development of generalized stochastic design approaches that address the robustness and reliability of complex multi-disciplinary systems. This project will demonstrate new UQ methods on high-performance aircraft nozzle analysis and design problems that are simultaneously designed for aerodynamic performance, thermal and pressure loads, and fatigue, while subject to geometric constraints to be fully integrated with complex vehicle shapes. The ability to handle all kinds of uncertainty at very large scale will enable the design of future components and vehicles that can have a substantial impact on DARPA’s mission.

Within Thrust Area 1, we will combine scalable polynomial chaos expansions with explicit dimension reduction in order to accurately resolve the most influential subspace within large-scale parameter domains. These approaches will be tailored and applied within both forward and inverse UQ contexts, as we seek to discover low-dimensional structure within high-dimensional domains, and will enable the use of very large numbers of uncertain parameters in our UQ and DUU efforts.

In Thrust Area 2, we will address the critical challenge of model inadequacies and develop methodologies for determining the source and extent of model-form discrepancies in simulation models by using i) observational/high-fidelity simulation data, and ii) Physics constraints. These methods enable accurate predictions with quantified uncertainty, despite the presence of imperfect sub-model empiricisms and closures and are constrained by the underlying physics. A novel formulation for embedded discrepancy is pursued, combining scalable inversion techniques in Thrust Area 1 along with large-scale machine learning techniques. Multifidelity algorithms for forward and inverse UQ will leverage the predictive capability of less expensive models through sparse, adaptive recovery of model discrepancy.

In Thrust Area 3, we build on the foundation provided by the previous thrust areas and incorporate rigorous, scalable UQ within the context of specific design problems, optimizing for robustness and reliability in the presence of variability, lack of knowledge, and model inadequacy. We employ a trust-region model management framework (TRMM) for DUU that exploits synergies within a hierarchy of statistical fidelity, provably achieving high-fidelity results at much lower computational cost.

RPPR Final Report

as of 06-Jan-2020

Moreover, we also focus on the application of advances in these three research areas to the problem of stochastic design of advanced aircraft nozzle systems. For this purpose we construct a hierarchy of simulation capabilities of increasing fidelity and cost, so that research contributions can be tested early (using low-fidelity models) in relevant environments, and so that the different simulation capabilities can be combined into multi-fidelity techniques using common parameterization and stochasticity.

Fundamental research breakthroughs in the efficacy of UQ and DUU methods for complex multiphysics systems will enable the deployment of UQ algorithms to higher-fidelity computational simulations operating in more complex high-dimensional uncertain environments, and paving the way for greater robustness and reliability in advanced vehicle designs of interest to DARPA and extending the predictive capabilities of computational models into new regimes.

Accomplishments: See uploaded PDF file.

Training Opportunities: Nothing to Report

Results Dissemination: See publications in uploaded PDF file

Honors and Awards: Nothing to Report

Protocol Activity Status:

Technology Transfer: Nothing to Report

Scalable Environment for Quantification of Uncertainty and Optimization in Industrial Applications (SEQUOIA)

Juan J. Alonso*, Charbel Farhat†, Gianluca Iaccarino‡, Rick Fenrich, Victorien Menier, Aashwin Mishra
Stanford University, Stanford, CA 94305, USA

Michael S. Eldred§, Gianluca Geraci¶, and John D. Jakeman||
Sandia National Laboratories, Albuquerque, NM 87185, USA

Paul A. Constantine** and Jeff Hokanson
University of Colorado, Boulder, CO 80309, USA

Karthik Duraisamy††, Alex A. Gorodetsky‡‡
University of Michigan, Ann Arbor, MI 48109, USA

The “Scalable Environment for Quantification of Uncertainty and Optimization in Industrial Applications” (SEQUOIA) project was funded as part of the DARPA “Efficient Quantification of Uncertainty in Physical Systems” (EQUIPS) program. It focused on performing uncertainty quantification (UQ) and design under uncertainty (DUU) while aggressively pursuing new frontiers in scale and complexity. In particular, the project focused on advancements in scalable forward and inverse UQ algorithms and the rigorous quantification of model inadequacy, providing the primary foundation for the development of stochastic design approaches that address the robustness and reliability of complex multi-disciplinary systems. The project was divided into three technical thrust areas and an application deployment thrust. Thrust area one was focused on scalable forward and inverse UQ methods, thrust area two was focused on multifidelity and model form, and thrust area three was focused on optimization under uncertainty. Each of these thrust areas focused on deployment to our application demonstration problem: high-performance aircraft nozzle analysis and design, for which the nozzle is designed for aerodynamic performance, thermal and pressure loads, and fatigue, while subject to the geometric constraints of being fully integrated within a complex vehicle shape. This report summarizes the highlights of what was achieved by our multi-institution team over the course of 3 years. It is necessarily a summary: details of the various topics discussed herein are included in the over 45 publications, referenced in the bibliography, that were produced as part of this effort. The Co-PIs and the entire SEQUOIA team thank DARPA and the DARPA program manager, Dr. Fariba Fahroo, for their generous support of this program of research.

I. Introduction

The “Scalable Environment for Quantification of Uncertainty and Optimization in Industrial Applications” (SEQUOIA) project provides an integrated plan for performing uncertainty quantification (UQ) and design under uncertainty (DUU) that aggressively pursues new frontiers in scale and complexity. In particular, the coordinated investments in this effort will create advancements in scalable forward and inverse UQ algorithms and the rigorous quantification of

*Vance D. and Arlene C. Coffman Professor, Department of Aeronautics & Astronautics, Co-PI

†Vivian Church Hoff Professor of Aircraft Structures, Department of Aeronautics & Astronautics

‡Professor, Department of Mechanical Engineering

§Distinguished Member of Technical Staff, Optimization and Uncertainty Quantification Department, P.O. Box 5800, MS 1318, Albuquerque, NM 87185, Co-PI

¶Interim Technical Staff, Optimization and Uncertainty Estimation, P.O. Box 5800, MS 1318, Albuquerque, NM 87185

||Senior Member of Technical Staff, Optimization and Uncertainty Quantification, Sandia National Laboratories, P.O. Box 5800, MS 1318, Albuquerque, NM 87185.

**Assistant Professor, Department of Computer Science

††Associate Professor, Department of Aerospace Engineering

‡‡Assistant Professor, Department of Aerospace Engineering, University of Michigan

model inadequacy, providing the primary foundation for the development of generalized stochastic design approaches that address the robustness and reliability of complex multi-disciplinary systems. This project will demonstrate new UQ methods on high-performance aircraft nozzle analysis and design problems that are simultaneously designed for aerodynamic performance, thermal and pressure loads, and fatigue, while subject to geometric constraints to be fully integrated with complex vehicle shapes. The ability to handle all kinds of uncertainty at very large scale will enable the design of future components and vehicles that can have a substantial impact on DARPA’s mission.

Within Thrust Area 1, we will combine scalable polynomial chaos expansions with explicit dimension reduction in order to accurately resolve the most influential subspace within large-scale parameter domains. These approaches will be tailored and applied within both forward and inverse UQ contexts, as we seek to discover low-dimensional structure within high-dimensional domains, and will enable the use of very large numbers of uncertain parameters in our UQ and DUU efforts.

In Thrust Area 2, we will address the critical challenge of model inadequacies and develop methodologies for determining the source and extent of model-form discrepancies in simulation models by using i) observational/high-fidelity simulation data, and ii) Physics constraints. These methods enable accurate predictions with quantified uncertainty, despite the presence of imperfect sub-model empiricisms and closures and are constrained by the underlying physics. A novel formulation for embedded discrepancy is pursued, combining scalable inversion techniques in Thrust Area 1 along with large-scale machine learning techniques. Multifidelity algorithms for forward and inverse UQ will leverage the predictive capability of less expensive models through sparse, adaptive recovery of model discrepancy.

In Thrust Area 3, we build on the foundation provided by the previous thrust areas and incorporate rigorous, scalable UQ within the context of specific design problems, optimizing for robustness and reliability in the presence of variability, lack of knowledge, and model inadequacy. We employ a trust-region model management framework (TRMM) for DUU that exploits synergies within a hierarchy of statistical fidelity, provably achieving high-fidelity results at much lower computational cost.

Moreover, we also focus on the application of advances in these three research areas to the problem of stochastic design of advanced aircraft nozzle systems. For this purpose we construct a hierarchy of simulation capabilities of increasing fidelity and cost, so that research contributions can be tested early (using low-fidelity models) in relevant environments, and so that the different simulation capabilities can be combined into multi-fidelity techniques using common parameterization and stochasticity.

Fundamental research breakthroughs in the efficacy of UQ and DUU methods for complex multiphysics systems will enable the deployment of UQ algorithms to higher-fidelity computational simulations operating in more complex high-dimensional uncertain environments, and paving the way for greater robustness and reliability in advanced vehicle designs of interest to DARPA and extending the predictive capabilities of computational models into new regimes.

The remainder of the report is structured as follows. Algorithmic highlights are presented for thrust area 1 in §V, for thrust area 2 in §VI and §VII, and for thrust area 3 in §VIII.

II. Contributions to the broader community

The contributions of the DARPA EQUiPS program and the work of the SEQUOIA team can be summarized in the following:

- *Model-form UQ*: development of new data-free and data-enabled model error approaches to quantify the inadequacy of our modeling assumptions
 - These new approaches are deployed within our DUU process
 - Implementations are available to the broader CFD community through the v6 release of SU2
- *Scalability*: development of new UQ approaches that exploit special structure including sparsity, low rank, and discovery of low-dimensional subspaces.
 - UQ problems are now limited by the true intrinsic complexity of the quantities of interest, rather than cursed by the original dimensionality
 - Idea of active subspaces has been demonstrated for dimensionality reduction of a broad range of engineering applications
- *Multifidelity*: development of new algorithms that exploit hierarchies of model form and discretization to render UQ/DUU tractable for complex, high-fidelity models.
 - By fusing information across a modeling hierarchy, use of the most expensive models can be reduced by orders of magnitude
- *Comprehensive DUU*: Incorporation of both parametric and model-form uncertainties within the design process,

- Enabling the design of vehicles that are more robust and reliable than industry-standard safety-factor approaches.
- Potential to eliminate dependence on historical data based approaches and safety factors to result in higher-performance systems

III. Software

In order to pursue our research in uncertainty quantification and design under uncertainty, multiple pieces of software were developed during the EQUIPS program. The following is a summary of those software packages, which are available from the performing team:

- *Dakota*: The Dakota software (<https://dakota.sandia.gov>) has been significantly extended in the areas of sparse and low rank approximation, dimensionality reduction with active subspaces, multilevel-multifidelity algorithms for forward and inverse UQ, and recursive trust region model management.
- *MULTI-F*: The MULTI-F python framework (<https://github.com/vmenier/MULTIF>) has been entirely developed for this project. According to parameters specified by the user through a single configuration file (design variables, fidelity levels definition), it performs a complex aero-thermal-structural analysis of a supersonic nozzle and returns scalar values of the desired quantities of interest (thrust, stresses, strains etc.). Several levels of fidelity are available, from a quasi-1D nozzle analysis to 3D non-axisymmetric RANS CFD. Significant efforts were devoted to ensuring the good behavior of the entire model: verification & validation, smoothness of the QoIs with respect to input parameter variations, robustness.
- *SU2*: New capabilities have been implemented in the SU2 code (<https://su2code.github.io>) in support of this project, including a model-form uncertainty estimation module for the SST turbulence model (using perturbations to the eigenvalues, eigenvectors and the amplitude of the Reynolds stress tensor), robustness improvement techniques, gradient computation for thrust, and an anisotropic mesh adaptation process. SU2 has just recently gone through the release of v6.0 (Falcon), the most successful ever launch of a new version of the software.
- *AERO-S*: The general-purpose finite element structural analyzer AERO-S (<https://frg.bitbucket.io/aero-s/index.html>) has been updated to output quantities of interest required for this project, such as the mass of specified groups of elements (corresponding to certain components of the model such as the baffles or stringers) and the nodal stresses and strains averaged across specified groups of elements. Furthermore stresses and strains can optionally be outputted in the material coordinate frame used to define the properties of the orthotropic layers of the load-bearing sandwich panel (i.e. the fiber direction) to facilitate the evaluation of the preferred failure criteria.
- *ANSYS Fluent*: ANSYS, Inc.'s Fluent software is the most widely used commercial CFD solver in academia and the industry. During the project, we have developed a *turbulence model uncertainty estimation module* in Fluent. This has been tested extensively and numerous results from its application have been reported in literature. This UQ module is being actively used at Stanford and other research institutions, such as Centrum Wiskunde & Informatica, Netherlands.
- *C³*: The Compressed Continuous Computation (*C³*) library (<https://github.com/goroda/Compressed-Continuous-Computation>) enables computation with high-dimensional multivariate functions with up to hundreds of input dimensions through the exploitation of low-multilinear-rank structure using continuous analogues of tensor decompositions. Many multilinear algebra computations with such functions are available; they include addition, multiplication, integration, differentiation, and approximation of multivariate functions. More expansive capabilities for uncertainty quantification and the solution of high-dimensional partial differential equations are also available. As part of this project we have implemented a new capability for regression of high-dimensional multivariate functions from scattered data, this improvement now allows for extensive analysis and post-processing of high-dimensional scattered data. It is also a state-of-the-art capability not existing in many other other libraries.
- *Python Active-subspaces Utility Library*: The Python Active-subspaces Utility Library is hosted on github at https://github.com/paulcon/active_subspaces and the software was published in the Journal of Open Source Software (<https://doi.org/10.21105/joss.00079>). The Library contains Python utilities for working with active subspaces in a given model. Given either (i) a function handle to the model's output and gradients as a function of the model's inputs or (ii) a set of previously generated input/output pairs, the Utility Library provides several methods for estimating the active subspaces—along with several diagnostics for the

quality of the estimated active subspace. With an active subspace in hand, the Utility Library contains tools for (i) constructing a polynomial or Gaussian radial basis response surface, (ii) estimating the minimum value of the function, and (iii) estimating the average of the function—all of which exploit the low-dimensional active subspace for increased efficiency. The Library also contains several convenience functions for plotting that may reveal simple relationships between the model’s active variables (linear combinations of the original variables) and its output.

IV. Application: UCAV Nozzle

In the context of the DARPA funded project SEQUOIA [1] we are interested in the design under uncertainties of devices that involve aero-structural-thermal interactions in advanced aerospace vehicles. One problem of interest is the analysis and design of a jet engine nozzle for small unmanned military aircrafts as, for instance, the Northrop Grumman UCAS X-47B. As in many other complex engineering systems, many sources of uncertainty are present in this problem definition, for instance the material properties, the atmospheric and flight conditions might not be known *a priori*. All these source of uncertainty need to be characterized and propagated through the numerical code in order to understand their impact on the final performance of the device and, ultimately, provide the basis for a design optimization under uncertainty.

A. Aero-thermo-structural model of a jet engine nozzle

Our aero-thermo-elastic problem is inspired by the Pratt & Whitney F100-PW-220 turbofan engine without after-burning capabilities, that powers the Northrop Grumman UCAS X-47B military unmanned aircraft, for which we want to design its nozzle. During normal operative conditions the nozzle structure is exposed to both thermal and mechanical loading. Therefore it needs to include a two-layers array, namely a thermal and a loading layer, placed internally and externally, respectively.

Evaluating the performance of this device requires solving a multi-physics problem, and this solution consists of the following steps which are illustrated in the diagram in Figure 1. An engine model simulator [2–4] provides the inlet conditions for the nozzle. These conditions are subsequently used as input for the computational fluid dynamics (CFD) solver SU2 [5]. This solver computes the solution for the internal flow and the temperature and pressure profile along the nozzle internal walls. The temperature and pressure loading are used by a Finite Element (FEM) solver AERO-S [6] which computes several quantities regarding the thermal and mechanical stresses inside the structure. More details on the computational approach adopted are provided in [7, 8].

The intrinsic multiphysics nature of this problem provides a natural stage for the development and test of several model alternatives. For instance, a 1D ideal solver can be adopted to replace the SU2 CFD solver and consequently decrease the overall computational cost. In a previous work we demonstrated that a similar strategy might be exploited in a multifidelity forward UQ analysis [9]. In the same spirit, in this contribution we deploy an array of model fidelities that ranges from the 1D ideal solver to a more accurate 3D Reynolds-Averaged Navier Stokes (RANS) model for the CFD. For the FEM solvers, we relied on several meshes with increasing spatial resolution and even axi-symmetric mesh configurations to avoid the cost associated to the cost of a fully three-dimensional solution. Figure 2 shows the coarse Euler mesh used by SU2 and a medium FEM mesh used by AERO-S.

1. Nozzle structure

Figure 3 shows a schematic of the nozzle structure. The structure of nozzle’s walls is comprised of the following elements. An inner thermal layer (1) and an outer structural layer (2-4) which are separated by an air gap. The material for the thermal layer (1) is a ceramic matrix composite that insulates the structural layer from high temperatures. The structural layer is constituted by a composite sandwich material composed of a titanium honeycomb (3) between two layers (2,4) of graphite-bismaleimide (Gr/BMI) composite material. The air gap between the layers adds additional insulation from the thermal loading for the structural layer. Additional supports between these two layers are not modeled, however the nozzle structure also includes six composite sandwich baffles (6-10) to attach the nozzle to the surrounding structure and stingers (5) that provide additional support in the axial direction.

2. Operative conditions

The operative conditions for the nozzle are inspired by a typical reconnaissance mission for a small unmanned high-subsonic military aircraft. In particular, we decided to focus on the most critical conditions which occurs for the

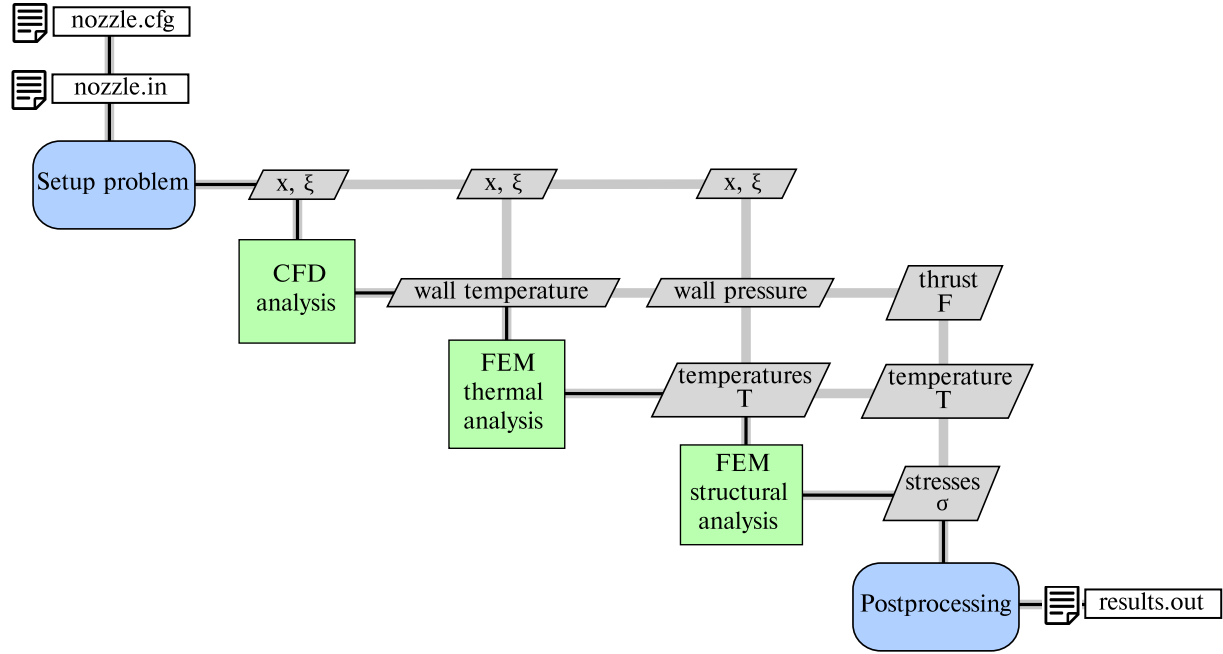


Fig. 1 The coupled multiphysics simulation used to analyze the supersonic nozzle and the associated process flow for deterministic design variables x and random parameters ξ . Three coupled analyses are run in order to calculate the thrust of the nozzle and temperatures and stresses/strains in the nozzle structure. A variety of fidelities (quasi-1D area-averaged flow/Euler/RANS for the CFD analysis) and mesh resolutions (for both CFD and FEM meshes) are available.

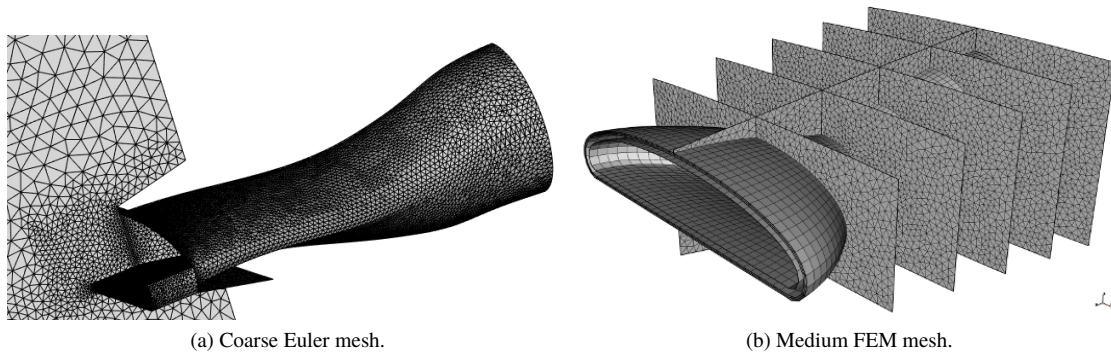


Fig. 2 A coarse resolution Euler mesh used by SU2 for the aerodynamic analysis and a medium resolution FEM mesh used by AERO-S for the structural analysis. A variety of mesh resolutions, including for RANS aerodynamic calculations are available.

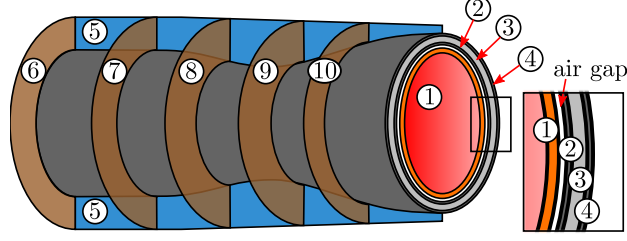


Fig. 3 Schematic of nozzle components.

top-of-climb condition when the required thrust is 21,500 N at an altitude of approximately 40,000 ft and Mach 0.511 [10].

3. Uncertain parameters

Following the numerical setup in [10] we use 40 uncertain parameters that are divided in 35 material properties variables, 2 atmospheric conditions variables, 2 inlet conditions and 1 heat transfer. All the variables are considered to be independent. More details on the bounds and distributions of the 40 uncertain parameters are reported in [10] however, we mention here that for this problem we deal with a set of uniformly and log-normally distributed variables.

V. Technical thrust area: scalable methods that exploit special structure

Our efforts for TA1 have focused on R&D of foundational methods for scalable uncertainty quantification, including both forward uncertainty propagation and statistical inversion. The emphasis on scalability addresses challenges related to the number of random input variables (n_ε), the number of output QoI (n_q), and the efficiency of parallel implementations on extreme-scale computers. Many of the scalable techniques developed in TA1 have been directly utilized in TA2 and TA3.

Our main phase 1 and phase 2 TA1 forward UQ accomplishments include:

- **Efficient sampling strategies for sparse PCE** [Jakeman, Narayan, Zhou, 2017]. Constructing a polynomial surrogate is a popular approach to decrease the number of samples required to accurately estimate statistics of expensive computer models. Typically the simulation output is approximated by a generalized Polynomial Chaos Expansion (PCE). Compressive sampling has been shown to be an effective method for approximating PCE coefficients from small number of function samples. However, the efficacy of sparse polynomial approximation is heavily reliant on the sampling strategy. We developed a sampling scheme that significantly improves the accuracy of sparse PCE.
- **Low-rank regression** [Gorodetsky, Jakeman, 2017]. Every method that scales to high dimensions must exploit some low-dimensional structure. Sparsity, discussed above, is one form of structure, low rank or separable structure is another. We developed a novel algorithm for low-rank regression which unlike existing low rank methods can be applied to archived unstructured data and be used with both linear and non-linear basis functions. This new flexibility allows surrogates to be built for a wider range of functions with much fewer samples.
- **Quadrature for correlated densities** [Jakeman, Narayan, 2017]. We developed a method for numerically generating polynomial quadrature rules, which unlike existing methods can be applied to correlated random variables and target problem structure. Our approach can generate quadrature rules for posterior densities generated by Bayesian inversion. Our quadrature rules exhibit exponential convergence and can be used for efficiently estimating integrals for data-informed robust design under uncertainty. Current methods for data-informed integration only exhibit Monte Carlo type convergence. Targeting problem structure also allows us to significantly increase the dimensionality for which polynomial quadrature rules can be successfully used. For example we have constructed quadrature rules on active subspaces to compute moments of QoI from our nozzle application.
- **Polyomial chaos expansions for dependent random variables** [Jakeman, Franzelin, Eldred, Narayan, Pflüger] Polynomial chaos expansions (PCE) are well-suited to quantifying uncertainty in models parameterized by independent random variables. The assumption of independence leads to simple strategies for building multivariate orthonormal bases and for sampling strategies to evaluate PCE coefficients. In contrast, the application of PCE to models of dependent variables is much more challenging. We developed a novel method that uses Gram-Schmidt

orthogonalization (GSO) to numerically compute orthonormal polynomials for the dependent random variables and used these basis to generate Leja sequences to build accurate polynomial interpolants. Polynomials constructed using Leja sequences focus sampling in regions of high-probability while mitigating ill-conditioning.

- **Fault-tolerant sparse collocation** [Adcock, Bao, Jakeman, Narayan, 2017]. An emerging challenge in UQ is the development of fault-tolerant or resilient algorithms. Such algorithms have become necessary in computational science as large-scale simulation models can fail in unpredictable and sometimes undetectable ways. We developed an algorithm for sparse that can detect the location of any corruptions amongst a collection of function samples as well as the corresponding corruption values.
- **Gradient-free active subspace discovery via ridge approximation** [Constantine, del Rosario, Iaccarino, 2016; Constantine, Eftekhari, Ward, 2016; Constantine, Hokanson, 2017] Active subspaces are defined by the gradient of the model’s output quantity of interest with respect to the model’s input parameters. However, many large-scale computational models do not have subroutines that compute the necessary gradients. We developed and analyzed optimization-based heuristics for estimating active subspaces when only input/output pairs are available.

Our main phase 1 and phase 2 TA1 inverse UQ accomplishments include:

- **Push forward based inference** [Butler, Jakeman, Wildey, 2017]. Inferring information about the parameters of a model from observations is a common goal in scientific modelling and engineering design. We formulated a numerical method for solving, an inverse problem for inferring parameters of a deterministic model from stochastic observational data (quantities of interest). This differs from traditional Bayesian inference which is used for stochastic maps and fixed data. Our approach for finding posterior measures, is simple and only requires the computation of the push-forward probability measure induced by the combination of a prior probability measure and the deterministic model. This fact allows us to directly leverage our advances in forward UQ. Indeed active-subspaces dimension reduction was combined with push-forward inference to infer 100 parameters describing the heat profile along the inner wall of our nozzle application. We describe this procedure in detail later in this report.
- Push forward based inference using surrogates [Butler, Jakeman, Wildey 2018] Push forward based inference (PBI) can require an intractably large number of forward simulations. To overcome this challenge we showed that PBI can be carried out using surrogates. We showed that the resulting push-forward and posterior densities are stable with respect to small errors and will converge, as the accuracy of the surrogate is increased, to the true solution which would be obtained without using a surrogate
- **Optimal experimental Design (OED)** [Walsh, Wildey, Jakeman, 2017] [Butler, Jakeman, Pisolov, Walsh, Wildey (2019) submitted]. The collection of experimental data can be costly and time consuming. In such situations we can only afford to gather some limited number of experimental data, however not all experiments provide the same amount of information about the processes they are helping inform. Consequently, it is important to design experiments in an optimal way, i.e., to choose some limited number of experimental data to maximize the value of each experiment. We developed an algorithm based upon our approach for push-forward based inference to construct optimal experimental designs which can be done in real time. Unlike existing approaches based upon traditional statistical Bayesian inference, we can perform all expensive model simulations offline and then update the posterior densities, at virtually no cost, when data becomes available during an online mode. We also developed an approach for OED which only requires jacobians of the parameter to observation map. This approach is extremely simple to apply and the designs generated can be used for many different inverse problems.

A. Highlight

Many of the developments listed above were directly used to impact TA2 and TA3. Here we outline how we were able to successfully infer a 100 uncertain variables describing the temperature distribution along the inner nozzle wall, using both active subspace dimension reduction and push-forward based inference. Our inference problem consisted of inferring 100 coefficients of Karhunen Loeve expansion (KLE) generated from an exponential covariance kernel with a correlation length of 0.1. To infer these parameters we used data generated from a baseline temperature profile (Figure 5), specifically thrust, wall pressure at 6 locations and pressure at 4 locations using our low fidelity model.

Push forward based inference computes a probability density on the model output and then seeks to determine a probability measure (posterior) on the space of model parameters, that is *consistent* with both the model and the observed data. By consistent we mean that the push-forward measure of the prior determined by the computational model matches the probability measure on the observations. The computational procedure simply requires evaluating the model at samples from the prior and building a density estimate of the push forward of those samples. The convergence

rate of the estimate of the posterior density is **independent of the parameter dimension** and only dependent on the number of QoI being used as part of the inference procedure. The formulation we have constructed is a **form of dimension reduction** as typically the dimension of the model output space is often much smaller than the dimension of the parameter space.

Unlike traditional Bayesian inference which requires the specification of an error model typically a mean zero Gaussian, our approach requires the specification of the observed densities of the QoI we are trying to match. The values of these QoI generated from the baseline configuration were used as the mean value of the observed densities which were assumed to be Gaussian with variance equal to 0.001 of the variance of the push-forward density. Given these observed densities we computed the push-forward densities of the QoI using Gaussian kernel density estimation (GKDE) based upon samples drawn randomly from the prior. The push forward for thrust and wall pressure at location 1 is shown in Figure 4 (c). Although the convergence rate of push forward based inference is independent of parameter dimension, the convergence rate obtained using GKDE is slow. To reduce the number of samples we needed to compute posterior densities we used active subspaces to reduce the parameter dimension and then built surrogates over the active subspace. We found a one-dimensional subspace for all QoI. In Figure 4 we plot thrust against the active variable and also plot the active subspace eigenvector that represents the relative contribution of each of the 100 KLE variables to that active variable.

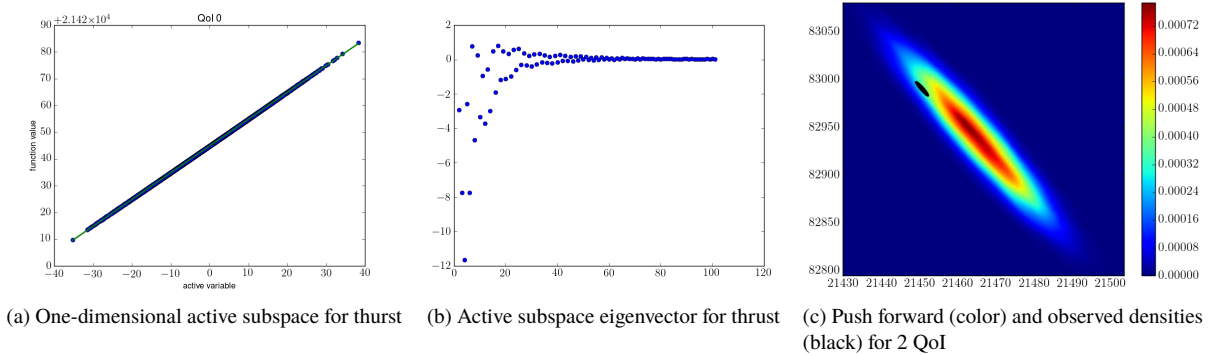


Fig. 4 Active subspaces of QoI are one-dimensional. Push forward of prior for 2 QoI compared against observed density which we want to match.

Figure 5 compares the prior distribution of the temperature profile with the posterior distributions obtain using 2 QoI and 3 QoI. Specifically we plot a set of realizations from each density against the ‘truth’ field used to generate the data which was used in the inference. Clearly the posterior produces realizations which are much more like the truth. Realizations from the posterior distribution ‘closer’ to the truth as the amount of data (number of QoI) used in the inference increases, however increasing the dimensionality of the QoI makes sampling from the posterior more difficult (Figure 5 (c)). We used rejection sampling here, but we are investigating more efficient means of sampling from the posterior.

VI. Technical thrust area: Model form uncertainty

With the recognition that the major source of modeling uncertainty arises from structural rather than parametric effects, we have developed different approaches to address non-parametric uncertainties:

- *Data-based techniques*, which use observational data to quantify uncertainties in the functional form of models, and
- *Physics-based techniques*, which use realizability constraints from physics to inform perturbations to the model output.

A. Model Form uncertainty: Physics-based

Turbulence models are constitutive relations that attempt to express quantities of interest that may be expensive to compute exactly, in terms of lower order quantities. A vast majority of Computational Fluid Dynamics (CFD) studies rely on eddy-viscosity based closures, such as the $k - \epsilon$ and $k - \omega$ models. Due to assumptions and simplifications

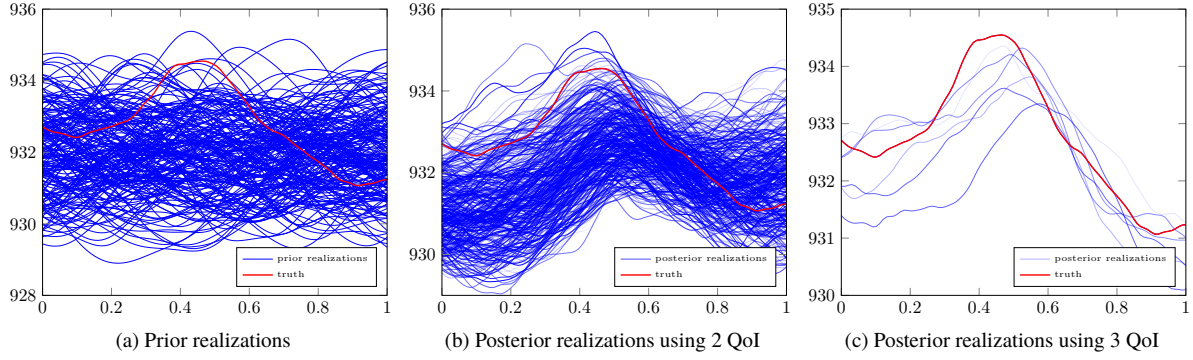


Fig. 5 Comparison of the wall temperature prior and posterior distributions. Shown are realizations from the respective densities.

introduced in the formulation of such Reynolds Averaged Navier Stokes (RANS) models, these models are limited in the characteristics of turbulent flows that they can replicate and their overall fidelity. Such simplifications include coarse graining where it is assumed that the Reynolds stress tensor can adequately describe the turbulent flow field, the eddy viscosity hypothesis that assumes the instantaneous Reynolds stress anisotropy to be proportional to the instantaneous mean rate of strain, the application of the gradient diffusion hypothesis to model turbulent transport, the severe limitations in the modeled evolution equations for the turbulent dissipation, besides others. These simplifications severely limit the performance of such models and introduce structural uncertainty in the predictions. Due to these, the performance of such RANS closures is severely limited under conditions such as swirl, mean rotation, flow separation and reattachment, etc.

To address such model form uncertainties in simulation, we focused on a two-pronged strategy: the development of efficient methodologies to provide reliable estimates of the model uncertainty and error, along with the integration of computational tools based on this methodology into CFD software. To this end, we developed the Eigenspace perturbation framework for uncertainty estimation due to turbulence models. This was integrated into the SU2 CFD software suite as the EQUiPS library.

The Eigenspace perturbation methodology utilizes physics based perturbations to the shape, the alignment and the size of the modeled Reynolds stress tensor. Explicitly, the spectral representation of the modeled Reynolds stress tensor is perturbed, during CFD solution iterations. This perturbation is expressed as: $R_{ij}^* = 2k^*(\frac{\delta_{ij}}{3} + v_{nl}^* \Lambda_{nl}^* v_{lj}^*)$, where R_{ij} and k represent the Reynolds stresses and turbulent kinetic energy, v_{nl} is the matrix of orthonormal eigenvectors, Λ_{nl} is the traceless diagonal matrix of eigenvalues λ_k and $*$ represent the perturbed quantities. The perturbations to the eigenvalues, Λ , correspond to varying the componentiality of the flow (or the shape of the Reynolds stress ellipsoid). Similarly, the perturbations to the eigenvectors and the turbulent kinetic energy vary the orientation and amplitude of the Reynolds stress ellipsoid. These perturbations are sequentially applied to the modeled Reynolds stress tensor.

The process of deriving uncertainty estimates from perturbed simulations is schematically reported in Fig. 6. The central panel of Fig. 6 outlines the unperturbed, baseline CFD solution, giving a unique flow field realization in the domain and a single profile for the QoI at the highlighted region (here, the mean velocity at $x/H = 24$ in the diffuser). The upper and lower panels of the figure outline perturbed solutions (only two of the perturbations are shown for clarity). Each perturbation leads to a different realization of the flow field, leading to a different profile for the QoI. The uncertainty estimates on the profiles of a QoI at a location are engendered by the union of all the states lying in the profiles from this set of perturbed RANS simulations. This is illustrated by the gray shaded zone in Fig. 6 panel C. This methodology has been tested for a variety of canonical turbulent flows[11], as well as complex turbulent flow cases[12].

In spite of the well accepted need to quantify the uncertainties and errors due to turbulence models, there were no strategies or tools available to estimate these. To address this need, the Eigenspace perturbation methodology was integrated into SU2 CFD suite as the EQUiPS library[13]. Before the release of this UQ tool, it was extensively tested for a bevy of turbulent flow cases. Some of the results for the turbulent flow over a NACA0012 airfoil at different angles of attack are outlined below. We outline the variation in the coefficient of lift with the angle of attack α . At low angles of attack, there is almost no discernible difference between the RANS predictions and the experimental data. Accordingly, here the uncertainty bounds are negligible. At higher angles of attack closer to stall, there is substantial discrepancy

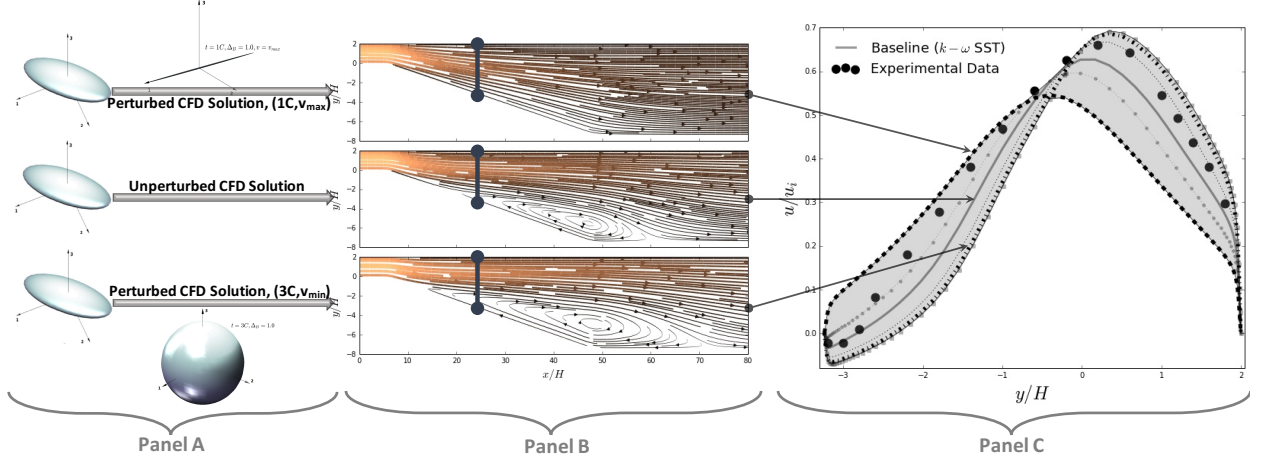


Fig. 6 Schematic outlining the composition of uncertainty estimates.

between the RANS predictions and the high fidelity data. For these values of α , the uncertainty bounds are substantial as well. At all values of α , the uncertainty bounds envelope the experimental data. In addition to the research project, this library is being used by researchers at INRIA (France), the University of Cambridge, Massachusetts Institute of Technology, Delft University of Technology, besides others.

B. Model Form uncertainty: Data-informed

In the present work, we focus on embedded model form uncertainty quantification methods that introduce functional calibration terms inside the model using a combination of statistical inference and machine learning methods. The propagation of the relevant stochastic information that defines the calibrated model endows the prediction with uncertainty that represents the ability of the calibrated model to represent the data. Model inadequacy is calibrated via a data-driven procedure which extracts information from experiments or higher-fidelity simulation and embeds them in a predictive setting. In comparison to output-based [14] model error quantification techniques, embedded model form techniques are capable of yielding physically realizable solutions. The key difference between the embedded UQ technique of Najm et al. [15] is that we are focused on extracting the functional form of the discrepancy rather than on parametric uncertainties.

A schematic of the methodology is shown in Figure 9. Starting with an existing model $\mathcal{M}$, a model correction field $\beta(\mathbf{x}, \xi)$ is introduced into the governing transport equations, where $\mathbf{x}$ represents physical space and ξ denotes the sample space. Starting with some high fidelity or experimental data G_d , Bayesian inverse problems are solved to extract $p(\beta|\mathcal{M}, G_d)$. Note that the intent is not to infer a limited set of parameters that characterizes the correction, but *rather to extract the correction field itself*. In other words, β is sought at every discrete location in the computational domain and inserted into the transport equation. *Thus we are directly addressing structural errors and uncertainties in the model.*

The correction fields will be obtained for an ensemble of problems (on different data sets $\mathbf{G}_d^1, \mathbf{G}_d^2, \dots$ as shown in the figure) representative of the physical phenomena that is to be modeled. Unlike in parameter estimation, in which the inferred parameters can be used directly in simulations, the $\beta(\mathbf{x})$ field by itself is of limited use in quantitative modeling. To be useful in predictive modeling, the spatial dependence of $\beta(\mathbf{x})$ has to be transformed into feature space using machine learning. Machine learning effectively constructs a functional form $\beta(\boldsymbol{\eta})$, where $\boldsymbol{\eta}(\mathbf{U}(\mathbf{x}))$ are features (derived from field variables $\mathbf{U}$) that will be available during the solution. It has to be recognized that the inversion and machine learning are *off-line* steps as shown in the red box. During the predictive simulation, (at each time-step or solver iteration), the solver will pass feature vectors $\hat{\boldsymbol{\eta}}$ to the ML evaluation routine and receive appropriate model correction quantities $\hat{\beta}$ for injection into the data-augmented model.

Inference and learning A key challenge in TA2 is the inversion to obtain model discrepancy bounds. Since a non-parametric uncertainty field is desired, the distribution of the discrepancy is inferred at every grid point in the computational domain. To reduce the extreme dimensionality of the inverse problem, we employ active subspaces. As an example, Figure 8a shows the eigenvalues of the active subspace computed at the prior and the posterior states (the maximum a posteriori estimate assuming Gaussian distributions is used as a pre-conditioner). The decay in modal energy is clearly very high in the posterior. For the posterior sampling, we use the first posterior eigenvector as well as

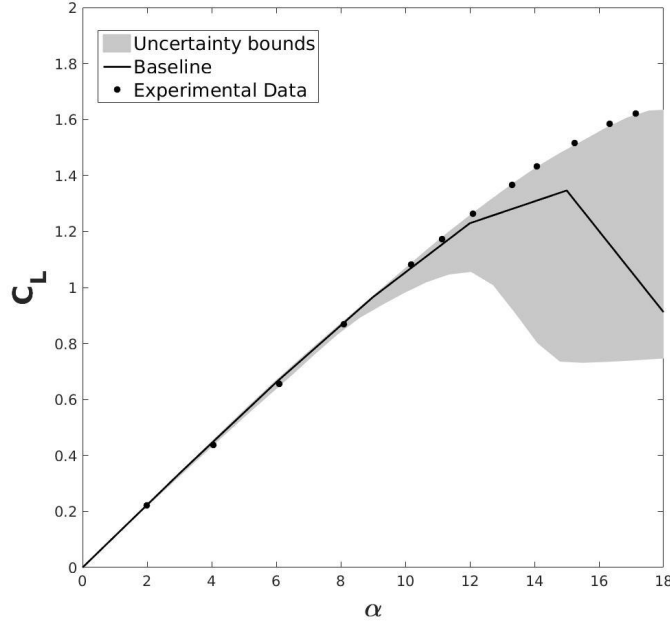


Fig. 7 Variation in the coefficient of lift C_L with the angle of attack α

its orthogonal subspace. Figure 8b shows the posterior realizations. For the 1D problem, we were able to limit the total number of MCMC samples to $O(10^3)$.

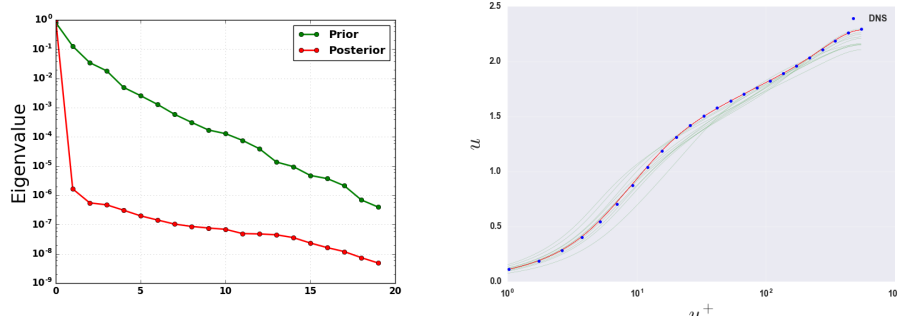


Fig. 8 Active subspaces calculated at the prior and the posterior for channel flow at $Re_\tau = 550$. **Right:** Prior and the posterior velocity using MCMC sampling and active subspace

As a representative example, machine learning augmented predictive simulations for a turbulent channel flow are shown in figure 10. Inverse problems were performed on a number of DNS datasets at different Reynolds numbers and Gaussian process (GP) regression was used for the machine learning. During the predictive cycle, envelopes of the output are obtained by sampling from the GP.

Application to Nozzle The axisymmetric version (medium-fidelity model) of the nozzle is considered for demonstration in Figure 11. The base model (which is the eddy viscosity-based Spalart Allmaras model [16]) results are shown as blue lines. Using high-fidelity data from a full Reynolds stress model, field inference and machine learning was used to reconstruct the form of the model discrepancy, introduced as a correction to the production of eddy viscosity. Predictive computations were then augmented with this machine learned ensembles and the envelope of the predictions is shown in the shaded red region, confirming the ability of the data-augmented approach to improve predictions (in the mean) as

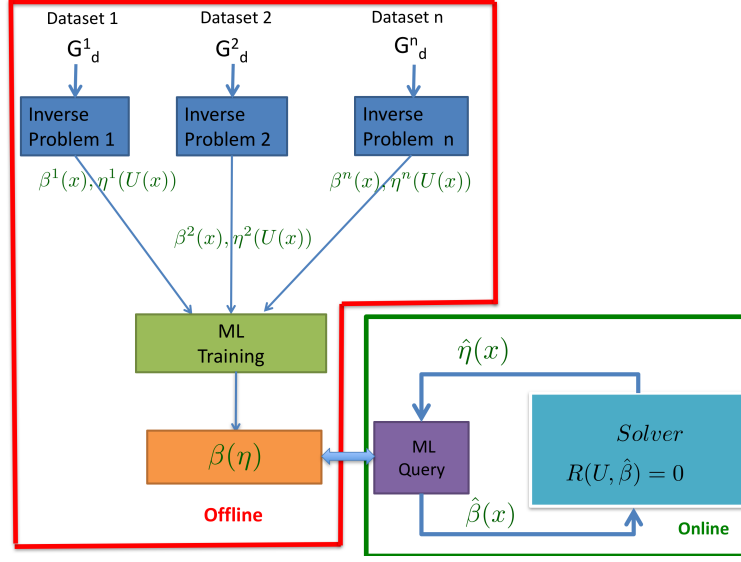


Fig. 9 Schematic of framework for data-based UQ method

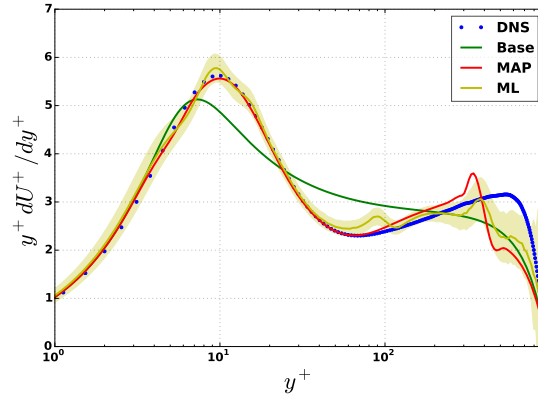


Fig. 10 Machine learning prediction for turbulent channel flow (shaded yellow, which represents the 95% interval). These results are compared to the baseline solution (green) and true solution (blue symbols). More details can be found in AIAA Scitech Paper titled “Augmentation of Turbulence Models Using Field Inversion and Machine Learning”.

well as quantify uncertainties.

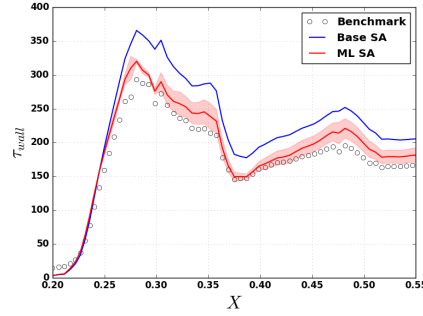


Fig. 11 Prediction (with data-based embedded model-form UQ) of skin friction over the nozzle wall.

C. Model Form uncertainty: a machine learning infused nonparametric probabilistic approach

To summarize again, model-form uncertainties can result either from the lack of knowledge of the true physics underlying the problem of interest, or the omission or truncation of modeling details, for reasons ranging from the lack of available information to the inability to discern in some circumstances between important and unimportant modeling details. Such uncertainties affect the ability of even a High-Dimensional Model (HDM) to deliver predictive results – that is, results that match reasonably well their experimental counterparts (which themselves can also be tainted by other types of uncertainties). Examples in computational structural dynamics include constitutive, multiscale, friction, homogenization, and free-play modeling errors. In the literature, the modeling, quantification, and propagation of parameter uncertainty has dominated the scene. By comparison, model-form uncertainty has been far less studied: therefore, its modeling and quantification are still in their infancies. This, despite the fact that once a model must be considered for any purpose, model-form uncertainties are unavoidable.

Indeed, many if not most of the methods developed so far for Uncertainty Quantification (UQ) are based on parametric probabilistic methods, and prior and posterior stochastic models of uncertain model parameters pertaining, for example, to geometry, boundary conditions, and material properties. Such methods are particularly well developed for modeling model parameter uncertainty, at least for a small number of parameters. Their computational feasibility has been demonstrated for small-size HDMs. For larger-scale HDMs, the underlying Monte Carlo simulations, or other type of stochastic computations, become so computationally intensive that model reduction becomes unavoidable. Unfortunately, model reduction introduces in the process its own model-form uncertainties which parametric probabilistic methods for UQ cannot take into account.

For example, Projection-based Model Order Reduction (PMOR) is one approach to achieve model order reduction that is grounded in mathematics. It is also the most prominent and perhaps the most rigorous approach for reducing the dimension of an HDM. It achieves this objective by projecting the HDM onto a subspace of lower dimension represented by a Reduced-Order Basis (ROB). This projection requires first constructing this ROB, then approximating the solution in the lower-dimensional subspace described by this ROB, and finally projecting some form of the governing equations associated with the HDM onto the same ROB or a related one. The outcome of PMOR is a Projection-based Reduced-Order Model (PROM). This reduced-order model distinguishes itself from simplified models that are also often called reduced-order models as follows: if the ROB is carefully constructed, the PROM can maintain most of the fidelity of the HDM, which a simplified model can rarely do. Unfortunately, a PROM inherits by construction any form of uncertainty in its underlying HDM. It also inherits the model-form uncertainties associated with the construction of the ROB onto which the HDM is projected to obtain this PROM. These uncertainties include at least:

- The uncertainty associated with finite sampling in the design of experiments type of procedure performed offline (and usually referred to as *training* or *learning*) to construct a global ROB whose domain of validity covers a large part of a parameter domain of interest.
- The uncertainty associated with the model truncation errors due to the truncation of the subspace of approximation underlying such a ROB via a data compression method such as, for example, SVD (Singular Value Decomposition) or POD (Proper Orthogonal Decomposition).
- The uncertainty arising from the approximation errors introduced by hyperreduction to achieve computational feasibility.

Hence, a PROM that is used to accelerate stochastic computations propagates in these computations the above model-form uncertainties which cannot be accounted for by parametric probabilistic methods for UQ. Considering that such methods are often designed for modeling model parameter uncertainty, it follows that at least in this case, modeling and quantifying model-form uncertainty is a pre-requisite for studying parameter uncertainty.

For structural dynamics analysis in general, nonparametric probabilistic methods for modeling uncertainties induced by general modeling errors have been developed, albeit for linear problems and a limited class of geometrically nonlinear structural analysis problems. In order to address this gap, we have developed under this research project a new, feasible, comprehensive, nonparametric, probabilistic method for modeling and quantifying model-form uncertainties in nonlinear structural dynamics problems [17] based on the concept of PMOR. When experimental data is available, this method also models and quantifies uncertainties in the associated HDMs. Its main underlying idea is two-fold. First, to substitute the deterministic ROB with a stochastic counterpart. Second, to construct the probability measure of the Stochastic Reduced-Order Basis (SROB) on a subset of a compact Stiefel manifold in order to preserve some important properties of a ROB. The stochastic modeling in this method is performed so that the probability distribution of the constructed SROB and associated Stochastic PROM (SPROM) depends on a small number of hyperparameters. These are determined by constructing and solving a reduced-order statistical inverse problem designed so that the mean value and statistical fluctuations of quantities of interest predicted using the SPROM match target values obtained from available data. Hence, our method can be interpreted as a probabilistic model learning or updating method. We have analyzed the mathematical properties of this novel approach for quantifying model uncertainties through both theoretical developments and numerical simulations. We have also demonstrated its potential for realistic problems through several example problems from nonlinear computational structural dynamics.

In [18], we have extended the scope of our method to generalized eigencomputations, as these are important in many fields including vibration analysis. We have also successfully demonstrated in [18] the potential of this extended method for the natural vibration analysis of a small-scale replica of an X-56 type aircraft made of a composite material, for which a small-scale Finite Element (FE) HDM and experimental data are available. In particular, we have shown in [18] that the extended, nonparametric, probabilistic method based on SPROMs is capable of modeling and quantifying reliably the uncertainties in the higher-modes of vibration, and in the modes of vibrations for which no experimental data is available.

During the research effort, we have also found that for some practical applications, the identification of the hyperparameters of our method for model-form UQ can be computationally intensive, for reasons that we have explained in [19]. Consequently, we have most recently redesigned this method in [19] around the concept of a Stochastic Hyperreduced PROM (SHPPROM) instead of that of an SPROM. We have also equipped in [20] the redesigned method with a faster, predictor-corrector approach for determining the appropriate value of the hyperparameters. This predictor-corrector approach is based on a probabilistic manifold learning approach.

Here, we highlight the ability of the most recent version of our method for modeling and quantifying uncertainties in structural dynamics problems by illustrating its application to what-if? vibration analysis scenarios associated with shape changes of the focus jet engine nozzle. For this purpose, we have built two shape-parametric structural HDMs: one with a higher fidelity to serve as a surrogate experimental system, and another one with a lesser fidelity to simulate model-form uncertainties. We have constructed a global SHPPROM for the latter HDM to accelerate the stochastic computations incurred by the proposed nonparametric probabilistic method for model-form UQ. All additional uncertainties introduced by this auxiliary SHPPROM, including those associated with the finite sampling of the domain shape-parameter domain – and the errors introduced by basis truncation and hyperreduction are accounted for. We have performed the UQ analysis for the eigenvalues of the modeled structural systems associated with two queried but unsampled parameter points.

1. Parametric jet engine nozzle and FE representations

The engine nozzle considered here consists of five cylindrical sections delimited by six baffles on the exterior of the nozzle, and four exterior stringers that run from the inlet to the outlet (see Fig. 12). Its shape is parameterized by a scaling factor l along the axis of symmetry (the x -axis in Fig. 12), and scaling factors r_y and r_z in the directions perpendicular to this axis. These three scaling factors define a shape-parameter domain $\Omega_\phi = \{(l, r_y, r_z)\} \subset \mathbb{R}^3$.

The nominal value of the main length of the nozzle is $L = 0.536$ m. The nominal value of its radius at the widest circular cross section is $R = 0.42$ m. Its thickness, which varies along the axis of axisymmetry but is constant in a cross section, decreases from 6.72×10^{-4} m at the outlet to 5.00×10^{-4} m in the middle section; then, it increases from its value at the middle section to the value of 1.32×10^{-3} m at the outlet. The inlet cross section of the nozzle is clamped.

For each considered shape configuration of this nozzle – that is, for each considered value of the parameter vector $\rho = (l, r_y, r_z)$ – two FE structural dynamics HDMs are constructed: each of them has 82,800 triangular anisotropic shell elements with 6 degrees of freedom (dofs) per node, 41,520 nodes, 247,320 unconstrained dofs – hence, $N = 247,320$ – and 1,800 constrained dofs where the boundary conditions are applied. The two HDMs differ only in how they represent the true thickness of the nozzle, which is discussed in the next section.

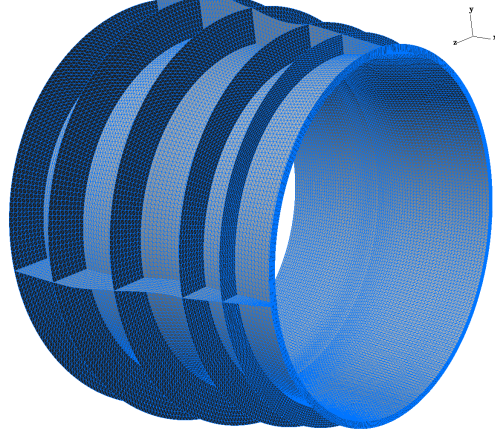


Fig. 12 Detailed FE structural model of a parametric jet engine nozzle with $N = 247,320$ unconstrained dofs.

2. Model-form uncertainties and what-if? scenarios associated with shape changes

One of the parametric HDMs mentioned above is equipped with an accurate representation of the variable thickness of the nozzle also described in that section. In the absence of true experimental data, this parametric HDM is used as a surrogate experimental system. On the other hand, the other parametric HDM mentioned above, from which all PROMs, HPROMs and SHPROMs are derived, is equipped with a constant representation of the thickness set here to 6.97×10^{-4} m in order to simulate a model-form uncertainty. Otherwise, both HDMs are parameterized in the parameter domain $\Omega_\rho = \{(l, r_y, r_z) \in [0.9, 1.1] \times [0.9, 1.1] \times [0.9, 1.1]\}$. Specifically, the main length of the nozzle and the semiminor and semimajor axes of its elliptical cross section at any point along its symmetry axis are varied by $\pm 10\%$ in the following what-if? vibration analysis scenarios.

In all considered scenarios, the QoIs are chosen to be the first six natural frequencies of the nozzle. Their surrogate experimental values and counterpart values predicted using the HDM tainted by model-form uncertainty are reported in Table 1, together with the relative discrepancies between these two sets of QoI values, at nine different points of the shape parameter domain Ω_ρ : the eight corners of this domain, and its center. The reported discrepancies are as high as 5.24%.

3. Global hyperreduced projection-based reduced-order model

A global HPROM of the jet engine nozzle is first constructed using: the parametric HDM tainted by model-form uncertainty described above; the greedy sampling algorithm applied in the parameter domain $\Omega_\rho = \{(l, r_y, r_z) \in [0.9, 1.1] \times [0.9, 1.1] \times [0.9, 1.1]\}$ and equipped with a $5 \times 5 \times 5$ uniform grid for pre-sampling at each iteration the candidate parameter points; and the ECSW hyperreduction method equipped with the non-negative LS optimization problem for determining the weight vector $\xi_{\tilde{\epsilon}}$ and the tolerance $\varepsilon = 3 \cdot 10^{-4}$. At each sampled parameter point, the following tasks are performed: the linear structural dynamics problem is solved using $f^{ext}(\rho, t) = 0$, and a broadband initial displacement u_0 and corresponding broadband initial velocity v_0 that excite all 6 natural modes of interest; 500 solution snapshots are collected in the time-interval $[0, 5 \times 10^{-3}]$ s at the sampling time-step $\Delta s = 1 \times 10^{-5}$ s; and these solution snapshots are accumulated in the snapshot matrix X which is compressed into a global deterministic ROB V using a truncated SVD that captures 99.97% of the energy of the snapshots.

The greedy sampling algorithm is initialized using $(l_1, r_{y1}, r_{z1}) = (0.95, 0.95, 0.95)$ as the first sampled parameter point. The residual-based error indicator is calibrated at the end of each iteration of this sampling algorithm by the

Table 1 Natural frequencies ($f_i = \omega_i/2\pi$ in Hz) computed using the surrogate experimental system and the deterministic FE-based HDM and relative discrepancies.

(l, r_y, r_z)	Surrogate experimental values (in Hz).					
	f_1	f_2	f_3	f_4	f_5	f_6
0.9, 0.9, 0.9	542.9	555.0	751.6	751.6	865.2	865.2
1.1, 0.9, 0.9	460.2	473.2	626.3	626.3	780.7	780.7
0.9, 1.1, 0.9	517.0	526.7	676.3	770.4	786.8	809.0
1.1, 1.1, 0.9	439.4	450.3	569.9	662.7	715.4	719.6
0.9, 0.9, 1.1	517.0	526.7	676.3	770.4	786.8	809.0
1.1, 0.9, 1.1	439.4	450.3	569.9	662.7	715.4	719.6
0.9, 1.1, 1.1	511.2	521.7	713.7	713.7	754.6	754.6
1.1, 1.1, 1.1	437.4	447.4	609.5	609.5	692.0	692.0
1.0, 1.0, 1.0	484.7	495.5	673.4	673.4	769.3	769.3
(l, r_y, r_z)	Numerically predicted values (in Hz) using the FE-based HDM.					
	f_1	f_2	f_3	f_4	f_5	f_6
0.9, 0.9, 0.9	554.2	563.5	768.2	768.2	902.2	902.2
1.1, 0.9, 0.9	470.5	481.1	640.5	640.5	821.6	821.6
0.9, 1.1, 0.9	527.6	534.6	690.2	797.0	816.0	831.9
1.1, 1.1, 0.9	449.0	457.6	582.2	679.3	748.5	753.8
0.9, 0.9, 1.1	527.6	534.6	690.2	797.0	816.0	831.9
1.1, 0.9, 1.1	449.0	457.6	582.2	679.3	748.5	753.8
0.9, 1.1, 1.1	521.5	529.4	728.8	728.8	779.9	779.9
1.1, 1.1, 1.1	446.7	454.3	622.7	622.7	722.4	722.4
1.0, 1.0, 1.0	494.9	503.2	688.2	688.2	802.9	802.9
(l, r_y, r_z)	Relative discrepancies (%).					
	f_1	f_2	f_3	f_4	f_5	f_6
0.9, 0.9, 0.9	2.08	1.53	2.21	2.21	4.28	4.28
1.1, 0.9, 0.9	2.24	1.67	2.27	2.27	5.24	5.24
0.9, 1.1, 0.9	2.05	1.50	2.06	3.45	3.71	2.83
1.1, 1.1, 0.9	2.18	1.62	2.16	2.50	4.63	4.75
0.9, 0.9, 1.1	2.05	1.50	2.05	3.45	3.71	2.83
1.1, 0.9, 1.1	2.18	1.62	2.16	2.50	4.63	4.75
0.9, 1.1, 1.1	2.01	1.48	2.12	2.12	3.35	3.35
1.1, 1.1, 1.1	2.13	1.54	2.17	2.17	4.39	4.39
1.0, 1.0, 1.0	2.10	1.55	2.20	2.20	4.37	4.37

averaged sum of the true relative errors of the HPROM for all 6 QoIs – that is, by

$$\frac{1}{m_u} \sum_{i=1}^{m_u} \frac{|\bar{f}_i - f_i|}{f_i} \quad (1)$$

where $m_u = 6$, and $\bar{f}_i = \bar{\omega}_i/2\pi$ and $f_i = \omega_i/2\pi$ are the i -th lowest natural frequencies predicted using the HPROM and the HDM representing the surrogate experimental system, respectively. The greedy algorithm is executed until the calibrated residual-based error indicator falls below the value of $\epsilon = 0.1$. In this case, the algorithm samples a total of five parameter points in Ω_ϱ , and produces an HPROM of dimension $n = 36$ that it computes on a reduced mesh $\tilde{\mathcal{E}}$ with 677 elements, 1,980 nodes, and $N_{\tilde{\mathcal{E}}} = 11,724$ unconstrained dofs (see Fig. 13).

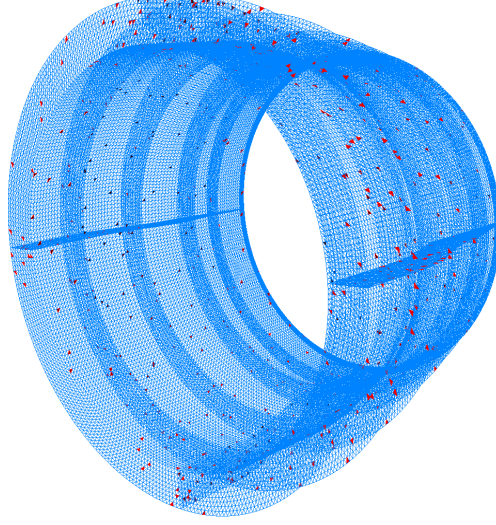


Fig. 13 Reduced mesh with $N_{\tilde{\mathcal{E}}} = 11,724$ unconstrained dofs constructed using the hyperreduction method ECSW (sampled elements are highlighted in red color).

To *verify* the performance of the greedy algorithm, and only for this purpose, the true average error (1) is computed on the aforementioned $5 \times 5 \times 5$ grid. Fig. 14 shows the progression of this true error during the five iterations performed by the algorithm. In this figure, each row is associated with one iteration of the greedy algorithm identified at the left, and each consecutive column corresponds to a cross-section of the nozzle located at $l = 0.90, 0.95, 1.00, 1.05$, and 1.10 , in this order. At each intersection of a row and a column of this figure, a contourplot of the true average error (1) is displayed in the (r_y, r_z) plane. The black dots in this figure reveal the parameter points that are sampled by the greedy algorithm. The reader can observe that: for this problem, the residual-based error indicator leads to sampling the parameter points where the true average error happens to be the largest; and after a parameter point is sampled, the true average error in the nearby region is minimized.

After completion of the final iteration of the greedy sampling algorithm, the averaged sum at any of the sampled parameter points of the relative true errors of the QoIs predicted using the constructed HPRM does not exceed 4.3%, and the relative true error for any QoI predicted using this HPRM at any of the sampled parameter points is less than 9%.

4. Global stochastic hyperreduced projection-based reduced-order model

Following the construction of the global HPRM of dimension $n = 36$, its underlying ROB V is randomized on the reduced mesh $\tilde{\mathcal{E}}$ of this HPRM, and the corresponding SHPRM is constructed. To identify the hyperparameter vector, the decoupled form of the governing minimization problem is solved using the following settings:

- $n_{rho} = 9$, $\varrho_1^{\text{ref}} = (0.9, 0.9, 0.9)$, $\varrho_2^{\text{ref}} = (0.9, 0.9, 1.1)$, $\varrho_3^{\text{ref}} = (0.9, 1.1, 0.9)$, $\varrho_4^{\text{ref}} = (0.9, 1.1, 1.1)$, $\varrho_5^{\text{ref}} = (1.1, 0.9, 0.9)$, $\varrho_6^{\text{ref}} = (1.1, 0.9, 1.1)$, $\varrho_7^{\text{ref}} = (1.1, 1.1, 0.9)$, $\varrho_8^{\text{ref}} = (1.1, 1.1, 1.1)$, $\varrho_9^{\text{ref}} = (1.0, 1.0, 1.0)$, and $m_u = m_o = 6$. It is noted here that none of these parameter points where surrogate experimental (reference) data is assumed to be available coincides with any of the parameter points sampled by the greedy algorithm during the construction of the HPRM.
- $\sigma^{\text{ref}} = \sigma^{\text{exp}}$, and $\sigma^{\text{exp}} = (f_1^{\text{exp}}, f_2^{\text{exp}}, f_3^{\text{exp}}, \dots, f_{m_o}^{\text{exp}})$, where f_j^{exp} , $j = 1, \dots, m_o$, denotes the j -th natural frequency computed using the surrogate experimental system (HDM with the true variable thickness of the jet engine nozzle).
- v_j^{ref} , $j = 1, \dots, m_o$, defined as follows

$$v_j^{\text{ref}} = \gamma |o_j^{\text{ref}} - o_j|, \quad j = 1, \dots, m_o$$

where $o_j = \omega_j/2\pi$, $j = 1, \dots, m_o$, and γ is set to $\gamma = 1.2$.

The weight w_J is set to 0.99 so that the target for J_{mean} is preponderant during the identification of the hyperparameter vector α . The solution of the coupled form of the minimization problem governing the hyperparameters is computed

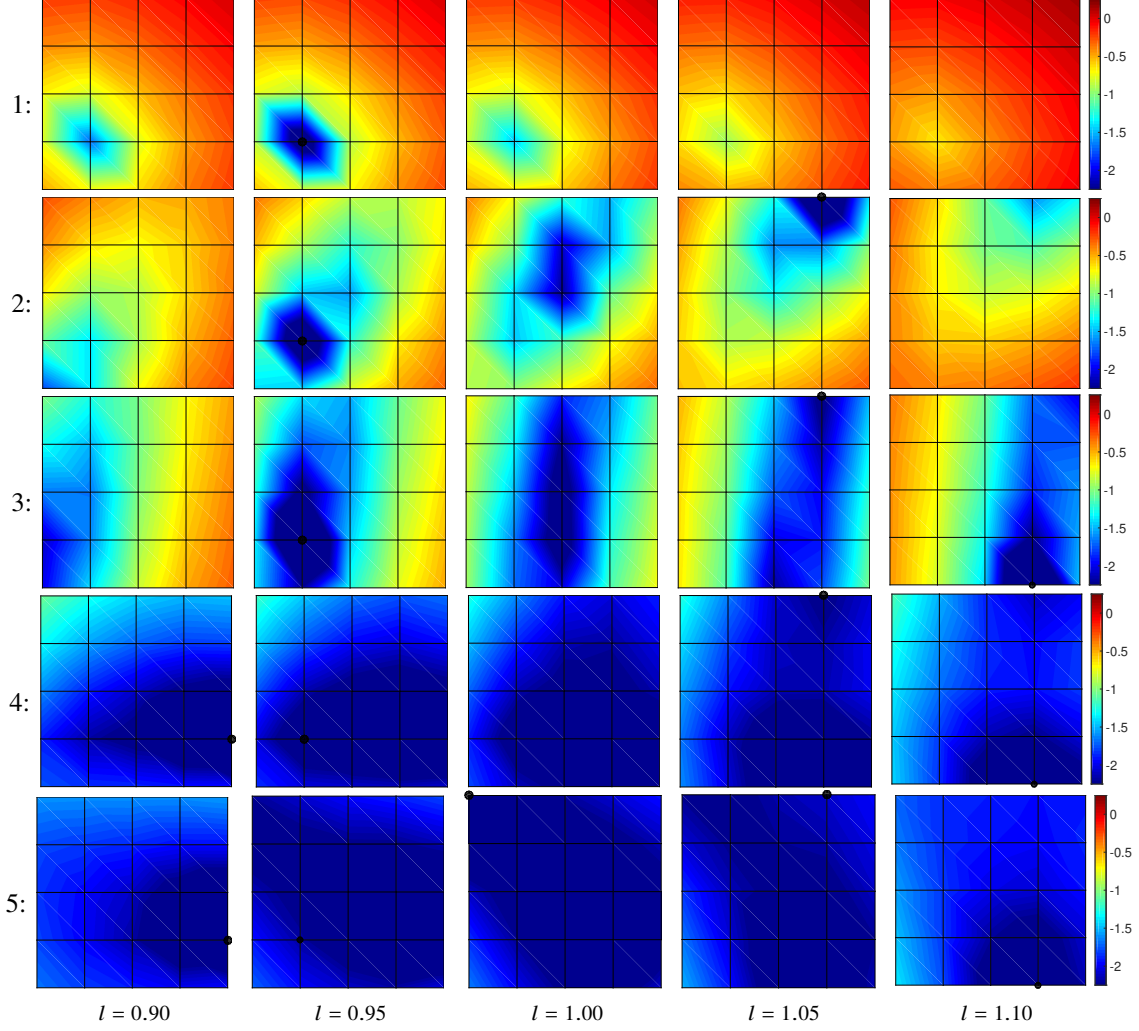


Fig. 14 Progression of the true average error ($\log_{10}$) on a $5 \times 5 \times 5$ (l, r_y, r_z) grid during the iterations performed by the greedy sampling algorithm.

using the MATLAB function `fmincon`, its active-set algorithm, and the following constraints:

$$\begin{aligned}
 10^{-5} &\leq s \leq 10^{-2} \\
 0.001 &\leq \beta \leq 0.1 \\
 0 &\leq \tau_{jj} \leq 20, \quad 1 \leq j \leq n \quad \text{and} \quad -10 \leq \tau_{ij} \leq 10, \quad 1 \leq i < j \leq n \\
 0.9999 &\leq s_M \leq 2.0
 \end{aligned}$$

The chosen optimization algorithm is initialized with $\alpha_0 = (s_0, \beta_0, s_{M_0}, \tau_0)$, where $s_0 = 4 \times 10^{-5}$, $\beta_0 = 4 \times 10^{-2}$, $s_{M_0} = 1$, and $\tau_0 = I_n$. The probability distributions of the random variables $\mathbf{s}$, β , and $\mathbf{s}_M$ are chosen as uniform distributions on the supports defined by their constraints. The probability model of the random matrix τ is constructed as follows. For $1 \leq i < j \leq n$, $\tau_{ij} = \tau_{0ij}(1 + \mathcal{N}_{ij})$, and for $i = 1, \dots, n$, $\tau_{ii} = \tau_{0ii}|0.728 + \mathcal{N}_{ii}|$, where $\{\mathcal{N}_{ij}\}_{1 \leq i \leq j \leq n}$ is a family of $n(n+1)/2$ independent variables. Each of these variables is a Gaussian, centered variable with a unit variance. Note that the absolute value is introduced in the expression of τ_{ii} in order to guarantee the required positiveness of the diagonal terms of the matrix τ .

The Monte Carlo method is used as a stochastic solver for estimating the cost function $J(\alpha)$ for every value of the hyperparameter vector α proposed by the optimizer. Specifically, 40 independent realizations of the QoIs are performed for each estimation of $J(\alpha)$. After 552 iterations, the value of this cost function is $0.99 \cdot 10^{-4}$: it is reached

for $\alpha = (s, s_M, \beta, \sigma)$ where $s = 4.92 \times 10^{-4}$, $s_M = 1.2437$, $\beta = 2.04 \cdot 10^{-2}$, and the 666 entries of τ are not reported for the sake of brevity.

5. Sample wall-clock performance

As stated earlier, the main purpose of the SHPROM is to ensure the computational feasibility of the nonparametric probabilistic method described in this paper for performing model-form UQ. Its construction, which is performed here in MATLAB, entails hundreds of thousands of evaluations of the cost function $J(\alpha)$. Even though many of these evaluations can be executed in parallel, a single such evaluation on a single processor is indicative of the consequential wall-clock time. Here, the reduction in wall-clock time that an SHPROM can offer for this function evaluation is demonstrated for the application considered herein by reporting the performance improvement this SHPROM delivers over its SPROM counterpart. To this end, the wall-clock time is reported in Table 2 for a single cost function evaluation on a single processor using each of these two stochastic models. While the overall computational approach is easily parallelizable (for example, by running many of the embarrassingly parallel Monte Carlo realizations concurrently), the choice of reporting the performance results for a serial implementation is made to avoid the otherwise distracting discussion of implementation-dependent parallel issues. Furthermore, it is noted that each wall-clock time reported in Table 2 accounts for performing 40 realizations: these include building the realizations of the SPROM/SHPROM matrices, and solving the stochastic reduced/hyperreduced generalized EVP at each of the nine parameter points where surrogate experimental (reference) data is assumed to be available (this entails the solution of 360 reduced-order generalized EVPs). The reader can observe that for the application considered herein, where the HDM has the dimension $N = 247,320$ and the SHPROM has the much smaller dimension $n = 36$ but its construction entails computations on a reduced mesh $\tilde{\mathcal{E}}$ with $N_{\tilde{\mathcal{E}}} = 11,724$ dofs, hyperreduction using the ECSW method reduces the total cost of the aforementioned function evaluation by a factor equal 31. This superlinear (with respect to the ratio $N/N_{\tilde{\mathcal{E}}} = 247,320/11,724 \approx 21$) speedup factor is mostly due to the fact that the most time-critical part of the function evaluation is the construction of the stochastic reduced stiffness and mass matrices for a random realization of the SROB, labeled “Construction generalized EVP” in Table 2. This part of the computation is accelerated by a factor larger than 46.

Table 2 Performance comparisons of the serial wall-clock timings (seconds) for a single cost function evaluation using the SPROM and SHPROM.

Computational model	Construction generalized EVP	Solution generalized EVP	Cost function (total)
SPROM	170.56 s	0.30 s	256.19 s
SHPROM	3.67 s	0.27 s	8.36 s

6. Modeling and quantification of model-form uncertainties

The SHPROM is applied here to perform the model-form UQ analysis of the HDM tainted by model-form uncertainty, while accounting for all modeling errors introduced by model reduction. The results of this analysis are reported in Fig. 15 for two queried but *unsampled* parameter points of the shape-parameter domain $\Omega_{\mathcal{Q}}$ located at two corners of this domain: the parameter point (0.9, 0.9, 0.9) and the parameter point (1.1, 1.1, 1.1). The mean values and confidence intervals corresponding to the quantiles 98% and 2% are constructed using the constructed SHPROM and 100 realizations of the QoIs.

The reader can observe that for this problem, the global HPRM problem reproduces accurately the data generated using the HDM tainted by model-form uncertainty. This is due to the greedy sampling procedure. The reader can also observe that:

- The mean values computed using the SHPROM capture reasonably well the surrogate experimental data, and in any case, much better than the deterministic HDM tainted by model-form uncertainty.
- For each QoI, the realizations generated using this SHPROM and the computed confidence intervals corresponding to the quantiles 98% and 2% encompass the deterministic value of its surrogate experimental value.

Both of these observations highlight the potential of the proposed nonparametric probabilistic method for model-form UQ in generalized eigencomputations in general, and vibration analysis in particular.

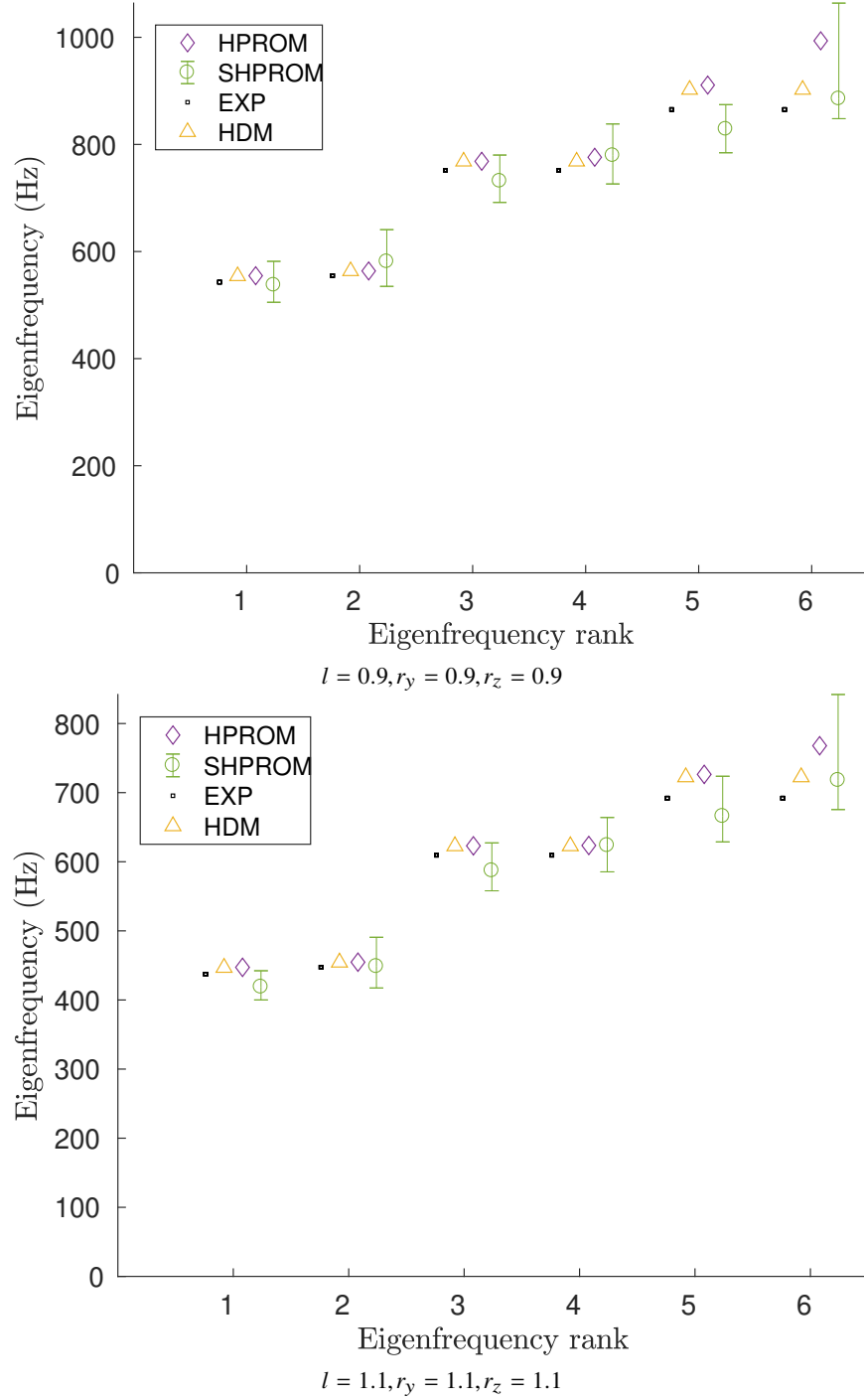


Fig. 15 Vibration analysis at two queried but *unsampled* parameter points of the shape-parameter domain: (confidence intervals correspond to the quantiles 98% and 2% constructed using the SHPROM)

VII. Technical thrust area: Multifidelity modeling

In the area of multifidelity methods, we have focused on approaches for forward UQ, including sampling-based methods and surrogate-based approaches. These methods reduce reliance on expensive high-fidelity fine-discretization models, and optimally exploit less expensive modeling alternatives. In Fig. 16 we depict relationships among the different activities, organized into recursive approaches that assume a hierarchical structure (left column) and generalized

	Recursive	Combined
Sampling	Multilevel Monte Carlo Multifidelity Monte Carlo	Approximate control Variates
Surrogates	Multi-index collocation Adaptive multi-index collocation Greedy adaptive MLMC	Bayesian latent variable networks

Fig. 16 Organization of multifidelity algorithms explored in this project. Bold-face indicates new SEQUOIA contributions.

approaches that can accommodate (and discover) non-hierarchical modeling relationships (right column).

A. Sampling-based approaches: highlights

The Monte Carlo algorithm is popular due to its simplicity, flexibility, and provably convergent behavior. Let $(\Omega, \mathcal{F}, P)$ a probability space and $Z \subset \mathbb{R}^d$ a vector valued random variable defined on this space. We consider a generic QoI Q as the mapping from Z to a scalar valued output. We desire to compute some statistics of Q , for instance its expected value $\mathbb{E}[Q]$. The MC estimator for the expected value of Q based on N realizations is

$$\mathbb{E}[Q] \simeq \hat{Q} = \frac{1}{N} \sum_{i=1}^N Q^{(i)}. \quad (2)$$

A classical results for MC is derived from the Central Limit Theorem and states that the mean square error for this estimator tends to be normally distributed with zero mean and variance $\text{Var}[Q]/N$, for $N \rightarrow \infty$. We are interested in computing statistics for computationally intensive applications for which the final number of evaluations N is limited. In recent years many strategies have been designed to accelerate the convergence of standard MC simulation by assimilating evaluations from cheaper to evaluate models [21–23].

In the following section, we describe our early experiences with multilevel and multilevel-multifidelity Monte Carlo approaches applied to forward UQ for the nozzle problem. We then describe a generalized approach [24] that enables us to interpret several multifidelity sampling approaches in the context of a generalized control variate for which the lower fidelity expected values are unknown and need to be evaluated at the same time of the high-fidelity expected value.

1. Initial nozzle model exploration

In an initial pilot study, we noticed a degradation of the performance of the multilevel-multifidelity (MLMF) approach due to uneven distribution of the correlation between the low- and medium-fidelity models for the three QoIs of interest, namely thrust and maximum thermal and mechanical stresses. In particular, the low-fidelity model demonstrated high correlation values only for the thrust, whereas the correlations for the stresses were poor. These correlations between the low- and medium-fidelity models are a key factor in the efficient redistribution of the computational burden across discretization levels and models, and therefore in the overall efficiency of the MLMF algorithm. This finding motivated an interaction between the UQ and the application team which resulted in improved low-fidelity structural and thermal models. The low-fidelity model is still based on the non-ideal nozzle solver for the fluid flow computation, but incorporates the finite element solver for the thermal and mechanical stresses computation. For an estimated mean squared error (MSE) of $\varepsilon^2/\varepsilon_0^2 = 0.01$ (normalized with respect to the pilot sample), the correlation properties of the two low-fidelity models are reported in Table 3. The updated model retains the high correlation values for the thrust, while significantly improving correlation for the stresses.

We previously deployed an MLMF estimator using three discretization levels for the medium-fidelity model and a control variate at its coarsest level. The control variate employs evaluations of the low-fidelity model. In Table 4 we compare the optimal sample allocations for different estimated MSE obtained by using the two low-fidelity models. An

	1D flow + hoop stress	1D flow + FEM
Thrust	9.98e-01	9.98e-01
Mechanical Stress	2.23e-02	9.66e-01
Thermal Stress	3.96e-01	9.93e-01

Table 3 Correlation for two LF models for the nozzle problem. The original LF model couples a 1D non-ideal fluid solver with a hoop stress model. The updated LF model replaces the stress components with FEM analyses. This correlation values correspond to the accuracy $\varepsilon^2/\varepsilon_0^2 = 0.01$.

overall trend is evident: the number of samples required for the coarse low-fidelity and coarse medium-fidelity have been reduced by approximately an order of magnitude to achieve a comparable MSE accuracy.

Accuracy ($\varepsilon^2/\varepsilon_0^2$)	LF	Medium Fidelity			LF (updated)	Medium Fidelity		
	Coarse	Coarse	Medium	Fine	Coarse	Coarse	Medium	Fine
0.1	N/A	N/A	N/A	N/A	404	20	20	20
0.01	21,143	1,757	20	20	3,091	177	31	20
0.003	69,580	5,775	36	20	N/A	N/A	N/A	N/A
0.001	212,828	17,715	109	34	32,433	1,773	314	20

Table 4 MLMF sample profiles, before and after LF model refinement, as a function of the normalized accuracy $\varepsilon^2/\varepsilon_0^2$.

The use of a well correlated low-fidelity model enables the MLMF algorithm to shift the computational effort from the medium- to the low-fidelity model, which in turn reduces the overall computational cost. The total computational cost for different values of $\varepsilon/\varepsilon_0$ is plotted in Figure 17, comparing MLMF (multilevel with coarse grid control variate), MLMC (multilevel only), and MC (no multilevel or control variate). For the simplified LF model, MLMC outperforms MLMF due to the uneven correlation across the QoIs. With the updated LF model, the MLMF estimator begins to outperform its MLMC counterpart by $\varepsilon^2/\varepsilon_0^2 = 0.1$.

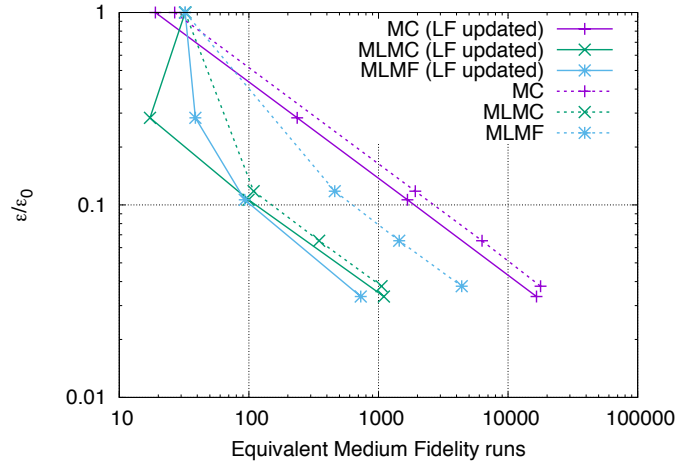


Fig. 17 Estimator standard deviation (ε) normalized w.r.t. its pilot sample value (ε_0), as a function of total computational cost (equivalent medium fidelity fine discretization evaluations).

2. Generalized Framework for Approximate Control Variates

Optimal Control Variate In the traditional control variate approach (denoted in the following as OCV) based on M low-fidelity models [25], the MC estimator Eq. (2) is modified by adding an unbiased term given by the difference between $\hat{Q}_i$, indicating the i th low-fidelity model estimator, and its expected value μ_i (which is assumed here to be known)

$$\hat{Q}^{\text{CV}} = \hat{Q} + \sum_{i=1}^M \alpha_i (\hat{Q}_i - \mu_i). \quad (3)$$

Moreover the unbiased terms are weighted by a sets of weights $\underline{\alpha} = [\alpha_1, \dots, \alpha_M]^T$. The optimal set of weights can be obtained by maximizing the variance reduction of the estimator [24]. If we denote $\mathbf{C} \in \mathbb{R}^{M \times M}$ as the covariance matrix among Q_i and $\mathbf{c} \in \mathbb{R}^M$ as the vector of covariances between Q and each Q_i , the highest variance reduction is obtained by choosing $\underline{\alpha}^* = -\mathbf{C}^{-1}\mathbf{c}$. If we further define $\bar{\mathbf{c}} = \mathbf{c}/\text{Var}[Q] = [\rho_1 \text{Var}[Q_1], \dots, \rho_M \text{Var}[Q_M]]^T$, where ρ_i is the Pearson's correlation coefficient between Q and Q_i , then the variance reduction corresponding to optimal weights becomes

$$\gamma^{\text{CV}}(\underline{\alpha}^*) = \frac{\text{Var}[\hat{Q}^{\text{CV}}]}{\text{Var}[\hat{Q}]} = 1 - R_{\text{OCV}}^2 = 1 - \bar{\mathbf{c}}^T \mathbf{C}^{-1} \bar{\mathbf{c}}, \quad 0 \leq R_{\text{OCV}}^2 \leq 1. \quad (4)$$

For one single low-fidelity model we obtain the classical result $R_{\text{OCV}-1}^2 = \rho_1^2$ and therefore the variance reduction is equal to the correlation of the low-fidelity model $\gamma^{\text{CV}}(\underline{\alpha}^*) = 1 - \rho_1^2$ where $\alpha_1 = \rho_1 \text{Var}[Q] / \text{Var}[Q_1]$.

Approximate Control Variate Traditional control variate assumes that the expected value of each low-fidelity model μ_i is known *a priori*. In the presence of complex engineering system simulations we instead have cheaper to evaluate low-fidelity models whose means we also need to estimate. To do so we consider $\mathbf{z}$ a set of samples used to evaluate the high-fidelity model Q and $\mathbf{z}_i$ an ordered set of N_i evaluations for Q_i . Furthermore, we partition $\mathbf{z}_i$ in two ordered subsets according to $\mathbf{z}_i^1 \cup \mathbf{z}_i^2 = \mathbf{z}_i$ (note that in general $\mathbf{z}_i^1 \cap \mathbf{z}_i^2 \neq \emptyset$).

Combining these subsets, we obtain the following form for the *approximate control variate*

$$\tilde{Q}(\underline{\alpha}, \mathbf{z}) = \hat{Q}(\mathbf{z}) + \sum_{i=1}^M \alpha_i (\hat{Q}_i(\mathbf{z}_i^1) - \hat{\mu}_i(\mathbf{z}_i^2)) = \hat{Q}(\mathbf{z}) + \sum_{i=1}^M \alpha_i \Delta_i(\mathbf{z}_i) = \hat{Q} + \underline{\alpha}^T \underline{\Delta}, \quad (5)$$

where $\underline{\Delta} = [\Delta_1(\mathbf{z}_1), \dots, \Delta_M(\mathbf{z}_M)]^T$, $\Delta_i(\mathbf{z}_i) = \hat{Q}_i(\mathbf{z}_i^1) - \hat{\mu}_i(\mathbf{z}_i^2)$ and the input values for the high- and low-fidelity models is denoted as $\mathbf{z} = (\mathbf{z}, \mathbf{z}_1, \dots, \mathbf{z}_M)$. Note that all the estimators in literature (for instance MLMC [21, 26] and MFMC [22]) can be obtained from the approximate control variate estimator (ACV) $\tilde{Q}$ by using different partitioning schemes for $\mathbf{z}_i^1, \mathbf{z}_i^2$ of each low-fidelity model. To facilitate easier analysis, we assume that the number of evaluations available for each low-fidelity model N_i is proportional to the number of evaluations of the high-fidelity model N , as $N_i = \lfloor r_i N \rfloor$ where $r_i \in \mathbb{R}_+$. In [24] we derived the expressions for both the optimal weights $\underline{\alpha}^{\text{ACV}}$ and the variance of $\tilde{Q}$. These results are reported here for convenience but the reader must refer to [24] for their derivation

$$\begin{aligned} \underline{\alpha}^{\text{ACV}} &= -\text{Cov}[\underline{\Delta}, \hat{Q}]^{-1} \text{Cov}[\underline{\Delta}, \hat{Q}] \\ \text{Var}[\tilde{Q}(\underline{\alpha}^{\text{ACV}})] &= \text{Var}[\hat{Q}] \left(1 - \text{Cov}[\underline{\Delta}, \hat{Q}]^T \frac{\text{Cov}[\underline{\Delta}, \underline{\Delta}]^{-1} \text{Cov}[\underline{\Delta}, \hat{Q}]}{\text{Var}[\hat{Q}]} \right) \\ &= \text{Var}[\hat{Q}] (1 - R_{\text{ACV}}^2). \end{aligned} \quad (6)$$

For reference, if only a single low-fidelity model is used, the classical results from [27, 28] is recovered

$$R_{\text{ACV}-1}^2 = \frac{r_1 - 1}{r_1} \rho_1^2. \quad (7)$$

To interpret this result, consider the term $\frac{r_1 - 1}{r_1} < 1$ as a penalization on the efficiency with respect to the traditional CV with a single low-fidelity model, *i.e.* OCV-1 Eq. (3), where a full reduction of ρ_1^2 is achieved.

Accelerating the Approximate Control Variate We demonstrated in [24] that a recursive sampling strategy can be exploited to accelerate the estimator to a target control variate level in case there is not enough data from the higher-fidelity models. The reason for that is that a recursive scheme can achieve the faster convergence to a given target control variate (for instance the OCV-1 or OCV-2). One possible approach is to partition the control variates in two groups: the first K variables form a K -level approximate control variate, and the last $M - K$ variables are used to reduce the variance of estimating μ_L for some $L \leq K$. The resulting estimator accelerates quickly to OCV- K , and L provides a degree of freedom for targeting a control variate level that contributes the greatest to the estimator variance. We report in the following only the final results for one estimator we derived, for the entire derivation the interested reader should refer to [24]. Let $K, L \leq M$ and $K \in \mathbb{Z}_+$ with $0 \leq L \leq K$. The ACV-KL estimator is given by the following expression and sampling scheme

$$\hat{Q}^{\text{ACV-KL}}(\underline{\alpha}, \mathbf{z}) = \hat{Q}(\mathbf{z}) + \sum_{i=1}^K \alpha_i \left(\hat{Q}_i(\mathbf{z}) - \hat{\mu}_i(\mathbf{z}_i) \right) + \sum_{i=K+1}^M \alpha_i \left(\hat{Q}_i(\mathbf{z}_L) - \hat{\mu}_i(\mathbf{z}_i) \right), \quad (8)$$

where we used that $\mathbf{z}_i^1 = \mathbf{z}$ for $i \leq K$ and $\mathbf{z}_i^1 = \mathbf{z}_L$ for $i > K$. Furthermore $\mathbf{z}_i^2 = \mathbf{z}_i$ for all i . The sets $\mathbf{z}_i \setminus \mathbf{z}_i^1$ can be chosen in several ways. Here we choose the same sampling strategy as ACV-MF: $\mathbf{z}_i^1 = \mathbf{z}$, $\mathbf{z}_i^2 = \mathbf{z}_i$ and $\mathbf{z}_j^{(k)} = \mathbf{z}_i^{(k)}$ for $j > i$ and $k \leq \min(r_i, r_j)N$.

Assume $r_i > r_L$ for $i > L$, then the optimal weights for the ACV-KL control variate are

$$\underline{\alpha}^{\text{ACV-KL}}(K, L) = - \left[\mathbf{C} \circ \mathbf{F}^{(K, L)} \right]^{-1} \left[\text{diag} \left(\mathbf{F}^{(K, L)} \right) \circ \mathbf{c} \right], \quad (9)$$

and the estimator variance is

$$\mathbb{V}ar \left[\hat{Q}^{\text{ACV-KL}}(\underline{\alpha}^{\text{ACV-KL}}(K, L)) \right] = \mathbb{V}ar \left[\hat{Q} \right] \left(1 - R_{\text{ACV-KL}}^2(K, L) \right), \quad (10)$$

where

$$R_{\text{ACV-KL}}^2(K, L) = \left[\text{diag} \left(\mathbf{F}^{(K, L)} \right) \circ \bar{\mathbf{c}} \right]^T \left[\mathbf{C} \circ \mathbf{F}^{(K, L)} \right]^{-1} \left[\text{diag} \left(\mathbf{F}^{(K, L)} \right) \circ \bar{\mathbf{c}} \right], \quad (11)$$

and $\mathbf{F}^{(K, L)} \in \mathbb{R}^{M \times M}$ such that

$$\mathbf{F}_{ij}^{(K, L)} = \begin{cases} \frac{\min(r_i, r_j) - 1}{\min(r_i, r_j)} & \text{if } i, j \leq K \\ \frac{(r_i - r_L)(r_j - r_L) + r_L(\min(r_i, r_j) - r_L)}{r_i r_j r_L} & \text{if } i, j > K \\ \begin{bmatrix} \frac{r_i - r_L}{r_i r_L} \\ \frac{r_j - r_L}{r_j r_L} \end{bmatrix} & \text{if } L < i \leq K, j > K \\ \begin{bmatrix} \frac{r_i - r_L}{r_i r_L} \\ \frac{r_j - r_L}{r_j r_L} \end{bmatrix} & \text{if } L < j \leq K, i > K \\ 0 & \text{otherwise} \end{cases}, \quad \text{for } i \neq j. \quad (12)$$

The diagonal elements are $\mathbf{F}_{ii}^{(K, L)} = \frac{r_i - 1}{r_i}$ if $i \leq K$ and $\mathbf{F}_{ii}^{(K, L)} = \frac{r_i - r_L}{r_i r_L}$ otherwise.

Numerical optimization procedure For the algorithms presented so far, the total computational cost can be expressed as a function of the cost of obtaining an evaluation of Q and Q_i , w and w_i respectively. An optimal sample allocation strategy can be obtained for MLMC and MFMC in closed form. This is not the case for the ACV estimators. In this latter case we resort to a numerical optimization that seeks to maximize the variance reduction subject to a constraint on the total cost $N \left(w + \sum_{i=1}^M w_i r_i \right)$. For ACV-KL we use a mixed-integer formulation where (K, L) are treated as categorical integer variable (addressed through enumeration). We mention here that the best choice of K and L is problem dependent; however, they can be estimated at negligible cost. Additional details on the optimization are reported in [24].

Parametric model problem In order to illustrate the features of the ACV estimators compared to the other multifidelity estimators in literature we present a parametric model problem that enables us to explore different scenarios in term of

	Q	Q_1	Q_2
Q	1	$AA_1/9 (\sin \theta \sin \theta_1 + \cos \theta \cos \theta_1)$	$AA_2/7 (\sin \theta \sin \theta_2 + \cos \theta \cos \theta_2)$
Q_1	<i>sym</i>	1	$A_1A_2/5 (\sin \theta_1 \sin \theta_2 + \cos \theta_1 \cos \theta_2)$
Q_2	<i>sym</i>	<i>sym</i>	1

Table 5 Correlation/covariance matrix for problem given by Eq. (13) with $A = \sqrt{11}$, $A_1 = \sqrt{7}$, and $A_2 = \sqrt{3}$.

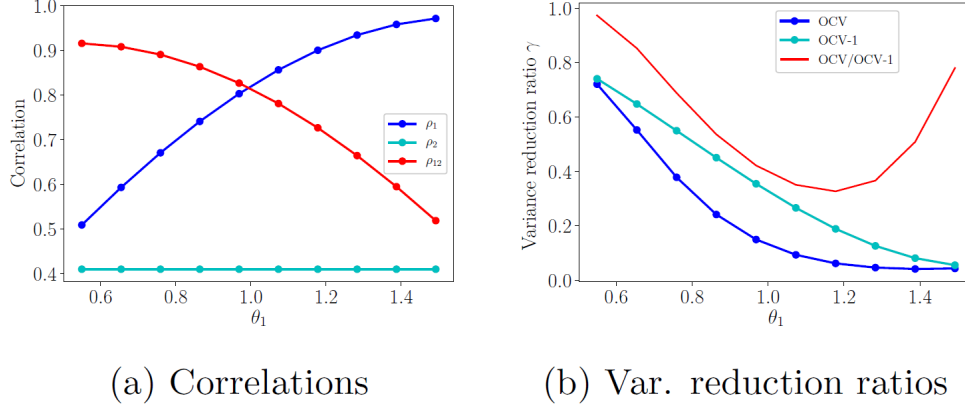


Fig. 18 Correlation (left) and variance reduction gap (right) for the parametric problem of Eq. (13) with $\theta = \pi/2$, $\theta_2 = \pi/6$ and $\theta_2 < \theta_1 < \theta$. Results from [24].

correlation among models and numerical cost. We consider a two-dimensional problem (two random variables) and two low-fidelity models that serve as control variate

$$\begin{aligned}
Q &= A \left(\cos \theta x^5 + \sin \theta y^5 \right), \\
Q_1 &= A_1 \left(\cos \theta_1 x^3 + \sin \theta_1 y^3 \right), \\
Q_2 &= A_2 (\cos \theta_2 x + \sin \theta_2 y),
\end{aligned} \tag{13}$$

where $x, y \sim \mathcal{U}(-1, 1)$ and all A_i and θ_i coefficients are real. For this particular example we choose to set $A = \sqrt{11}$, $A_1 = \sqrt{7}$ and $A_2 = \sqrt{3}$ to obtain unitary variance for each model. This choice further reduce the degree-of-freedom in the problem parameterization since the correlation and covariance matrices are identical. Specifically, the analytic correlation/covariance matrix is given in Table 5. To create enough variability in the scenarios we want to analyze, we fix the parameter for the high-fidelity and lowest fidelity model to $\theta = \pi/2$ and $\theta_2 = \pi/6$, respectively, and we let the middle fidelity model to vary, *i.e.* θ_1 varies uniformly in the bounds $\theta_2 < \theta_1 < \theta$. Each particular value of θ_1 induces a different correlation ρ_1 between Q_1 and Q and a different correlation ρ_{12} between Q_1 and Q_2 , whereas the correlation ρ_2 between Q_2 and Q remains fixed. These correlations are reported for different settings of θ_1 in a previous publication.

First we demonstrate that the variance reduction ratio of the control variate that considers both models, *i.e.* OCV, is larger than the one obtained using only a single control variate, *i.e.* OCV-1. This ratio of the estimator variance of OCV-1 and OCV is reported in Fig. 18. In Fig. 18 we also report the variance reduction obtained by OCV and OCV-1 with respect to MC. The greatest gap between OCV and OCV-1 occurs when θ_1 is approximately $\pi/3$ (the middle of its range).

The presence of a gap between OCV and OCV-1 indicates that the ACV estimators might outperform the recursive-based approaches as MLMC and MFMC because these can attain at most the variance reduction of OCV-1. Efficiently exploiting this gap ultimately depends on the possibility to obtain enough samples for the low-fidelity models and therefore, in turn it depends on the cost ratios between the different models. We explore several cost scenarios by assigning a relative cost of 1 for Q , $1/w$ for Q_1 and $1/w^2$ for Q_2 . For an equivalent total cost of 100 runs of Q , we consider the effects of w on the performance of each estimator. These results are illustrated in Fig. 19.

Some of the important features emerging from the analysis of Fig. 19 are reported in the following. First, the ACV-KL estimator generally outperforms all other estimators across the range of cost ratios enumerated in plots (a)

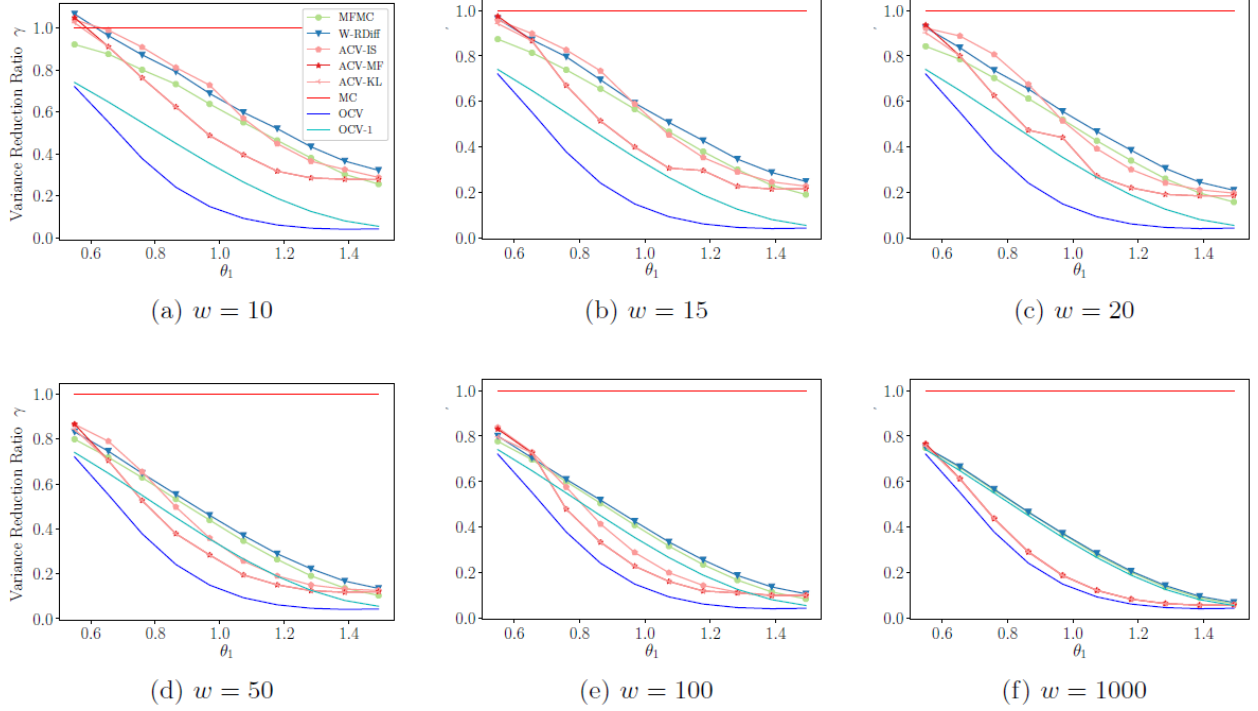


Fig. 19 Variance reduction for cost ratios of $[1, 1/w, 1/w^2]$ for Q , Q_1 , and Q_2 , results from [24].

through (e), and across the range of θ_1 from $\pi/6$ to $\pi/2$. However, in this particular example ACV-KL is nearly identical to ACV-MF since large deviations in (K, L) cannot occur for only two low-fidelity models. In most of the scenarios, these two estimators also outperform OCV-1 and the relative benefits with respect to OCV-1 increases whenever the low-fidelity models become less expensive. For very inexpensive low-fidelity models ($w = 1000$) all the ACV estimators almost converge to the OCV estimator, whereas the recursive estimators only converge to OCV-1 as suggested by the theory.

UCAV nozzle problem We demonstrated above that all the recursive sampling based schemes cannot achieve a variance reduction higher than the one corresponding to OCV-1. We proposed the ACV framework as a possible approach to overcome this issue and obtain converging estimators, *i.e.* multifidelity sampling estimators that converge to OCV in the limit of an infinite number of low-fidelity evaluations. As we demonstrated by means of the parametric model problem in the previous section, the efficiency of the ACV estimators ultimately depends on the presence of a separation between the performance, *i.e.* variance reduction, of OCV-1 and OCV. If this gap exists and the cost ratio is adequate we demonstrated that this framework outperforms the recursive schemes. In this section we analyze the nozzle problem to understand if a similar gap between OCV and OCV-1 exists for several QoIs, namely the thrust of the nozzle, a temperature failure criterion in the inner load layer, and a strain failure criterion in the thermal layer. We focus on an array of seven models ranging from the 1D ideal fluid solver up to a RANS 3D model (each with a different spatial resolution for the FEM analyses). The full range of models and their costs are reported in Table 6. We note that the meshes for both the CFD solver and the FEM solvers are denoted as COARSE, MEDIUM and FINE, however, they vary for each fidelity, for instance meshes COARSE for Euler 2D, Euler 3D and RANS 3D are all distinct.

A general observation is that the correlations among all the models are very high for all the three QoIs we considered. We report in Table 7 the variance reduction that could be potentially achieved by the OCV estimator (or potentially a convergent ACV estimator). The variance reduction is expressed as the term $1 - R^2$, *i.e.* for a generic estimator the absolute variance is given by the Monte Carlo estimator variance with the same number of high-fidelity evaluations times the variance reduction term $1 - R^2$. We also compare the ratio of the variance reduction between OCV-1 and OCV to make explicit the performance increase that could be attained by using the ACV framework (in the limit of infinite low-fidelity data).

We note that for all the QoIs considered a separation exists between OCV-1 and OCV. The variance reduction of

CFD	FEM (Thermal/Structural)	Cost
1D	COARSE	2.63e-04
Euler 2D COARSE	COARSE (axisymmetric)	9.69e-04
Euler 2D MEDIUM	MEDIUM (axisymmetric)	3.18e-03
Euler 2D FINE	FINE (axisymmetric)	9.05e-03
Euler 3D COARSE	COARSE	1.16e-02
Euler 3D MEDIUM	MEDIUM	3.58e-02
RANS 3D COARSE	COARSE	1.00

Table 6 Relative computational cost for several model fidelities for the nozzle problem. All the cost are normalized with respect to the 3D RANS solver.

QoI	Variance reduction		
	OCV	OCV-1	Ratio OCV/OCV-1
Thrust	0.020595	0.050432	0.41
Temperature failure criterion	0.0043612	0.0075662	0.58
Strain failure criterion	6.2981e-04	0.011720	0.05

Table 7 Comparison of OCV and OCV-1 performance for three different QoIs in the nozzle problem.

OCV-1 with respect to the plain MC estimator is greater than one order of magnitude for all QoIs and an additional performance increase of approximately 40% could be obtained for Thrust and the temperature failure criterion by using the OCV estimator. Moreover, for the strain failure criterion the variance of OCV is only 5% that of OCV-1. We note that these values correspond to the maximum attainable variance reduction under the assumption that the statistics of the low-fidelity models are known (or equivalently the cost of the low-fidelity model is zero).

3. Leveraging Active Subspaces for Multifidelity Sampling

In the previous section we demonstrated that the performance of all the multifidelity sampling based algorithms are ultimately related to the correlation matrix among the models. Moreover, in many circumstances, in order to increase the efficiency of the lower fidelity models, the physics and the numerical approximations need to be much simplified. This might potentially lead to a lack of consistency with respect to the original model parametrization. We recently proposed [29] to employ a dimension reduction strategy before deploying a multifidelity sampling estimator to seek a greater correlation and overcome the inconsistent models' parametrization. As an example of this strategy we proposed to adopt an Active Subspace [30] dimension reduction step to accelerate a control variate with a single low-fidelity model. The extension to several control variates and ACV estimators is straightforward therefore here we focus on a single control variate for illustration purpose. The idea is very simple and in order to present it we consider two models Q and Q_1 , the high- and low-fidelity model respectively. We assume z and z_1 the input vector of random parameters for Q and Q_1 respectively. Moreover we consider that the model can have a different dimensionality, *i.e.* $z \in \mathbb{R}^d$ and $z_1 \in \mathbb{R}^{d^1}$. For simplicity of exposure we consider the input parameters to be a vector of i.i.d. standard normal Gaussian variables, but we will relax this assumption later on. For each model, we perform an Active Subspace analysis and we identify the active directions independently. This step is accomplished by using only the pilot samples (which are usually generated in any case for a control variate estimator) and does not strictly require the knowledge of the gradient of a function because a low-order regression might be adopted (we tested several problems with only a linear regression and obtained satisfactory results). This step produces a vector of active variables z_A and $z_{1,A}$ for both models

$$\begin{aligned} z_A &= (\mathbf{W}_A)^T z \\ z_{1,A} &= (\mathbf{W}_{1,A})^T z_1, \end{aligned} \tag{14}$$

where $\mathbf{W}_A$ and $\mathbf{W}_{1,A}$ indicates the orthogonal matrices whose columns are the normalized eigenvectors for the high- and low-fidelity model respectively. These matrices are obtained through the standard Active Subspaces analysis, *i.e.* the solution of the eigenvalue problem for the expected value of the matrix generated by the outer product for the gradient

(see [30] for more details). Given the linearity of the transformation, the orthonormality of the eigenvectors and the normal distribution of both z and z^1 , both variables z_A and $z_{1,A}$ are standard normal variables. Therefore it is possible to identify a shared space between the models and a vector of standard normal variables $\mathbb{R}^{d_A^*} \ni z^* \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{d_A^*})$ where $d_A^* = \max(d_A, d_A^1)$.

In a simple version of the algorithm the approximate control variate estimator can be built by evaluating a set of N_{pilot} samples for $\{z^*\}_{i=1}^{N_{pilot}}$ for which the models are evaluated after the rotation to the original coordinates by using Eq. (14). At this point the approximate control variate approach (ACV-1) can be deployed as usual just sampling the variable z^* to create a new relationship between the models through the matrices $\mathbf{W}_A$ and $\mathbf{W}_{1,A}$. As we demonstrated in [29] if the original distribution of the input variables of the low- and high-fidelity model is not a standard Gaussian, a non-linear transformation can be adopted and the product between this function and the original model becomes the subject of both the Active Subspace analysis and the control variate.

UCAV nozzle problem In this section we present some preliminary results regarding the use of the Active Subspaces to accelerate the convergence of an ACV-1 estimator for the nozzle problem. We consider only the ACV-1 estimator for simplicity, but the extension to multiple model is straightforward. The high-fidelity model in this case uses 3D Euler for the CFD solver and a COARSE mesh for the FEM thermal and structural solver. Two low-fidelity models are separately considered; both models are based on the same numerical approximation, namely 2D Euler solutions for the CFD and axi-symmetric FEM mesh. However, albeit the nominal parametrization of the two low-fidelity model is the same, one model uses a consistent parametrization with the 3D geometry, *i.e.* the preservation of the nozzle internal area, section-by-section, is enforced whereas for the other one the axi-symmetric mesh is defined starting from the mid-plane section of the 3D geometry. Henceforth, we refer to these models as consistent and inconsistent parametrization, respectively. The consistency of the parameterization is enforced underneath by the aero-thermo-structural solver, therefore we consider it as a distinct model with a slightly higher computational cost due to this pre-processing stage. The computational cost for the three models used in this section is reported in Table 8.

CFD	FEM (Thermal/Structural)	Parameterization	Cost
3D Euler COARSE	COARSE		1.00
2D Euler COARSE	COARSE (axisymmetric)	Consistent	0.201
2D Euler COARSE	COARSE (axisymmetric)	Inconsistent	0.135

Table 8 Relative computational cost for the models used for the Active Subspace tests for the nozzle problem. All the costs are normalized with respect to the 3D Euler COARSE solver.

The numerical test has been configured to include 40 uncertain parameters that describe operative conditions and material properties as described in [10] and 96 design parameters that describe control points coordinates for the nozzle geometry. Moreover, the set of random parameters is constituted by a combination of uniform and log-normal distributed variables.

Multiple scenarios are described in [10]. Here, we highlight the third scenario, for which high- and low-fidelity model with inconsistent parametrizations are evaluated for the same set of nominal samples. In this case, we also consider the design parameters of the high-fidelity model to be uniform variable with a range of $\pm 2.5\%$ with respect to their nominal values. The inconsistency in this case stems from the different dimensionality of the two input spaces (136 versus 40 for the high- and low-fidelity, respectively).

We omit for brevity the representation of the active subspaces for the four QoIs, but we report their scatter plots in Fig. 20. We can note that the additional 96 uncertainty parameters in the design variable add only a limited amount of variability to the scatter plots especially for the thermal and mechanical stresses. However, their impact is more evident for the mass and thrust when considering the original variables. By using the PCE regression along the active variable we are able to compute, *a posteriori* and without additional realizations of the low-fidelity model, low-fidelity responses for a set of active variables consistent with the high-fidelity samples. Therefore, after this projection, the scatter plots reveal more clearly the correlation between the models.

In Table 9 we report the variation of the correlation squared between the samples taken in the physical variables and its estimated value from the PCE regression along the active variable. This value enable us to estimate the reduction in the estimator standard deviation w.r.t. to plain MC approach for both the standard OCV-1 with the sampling performed in the original/physical variables and OCV-1 (AS) with the sampling carried out in the active variables.

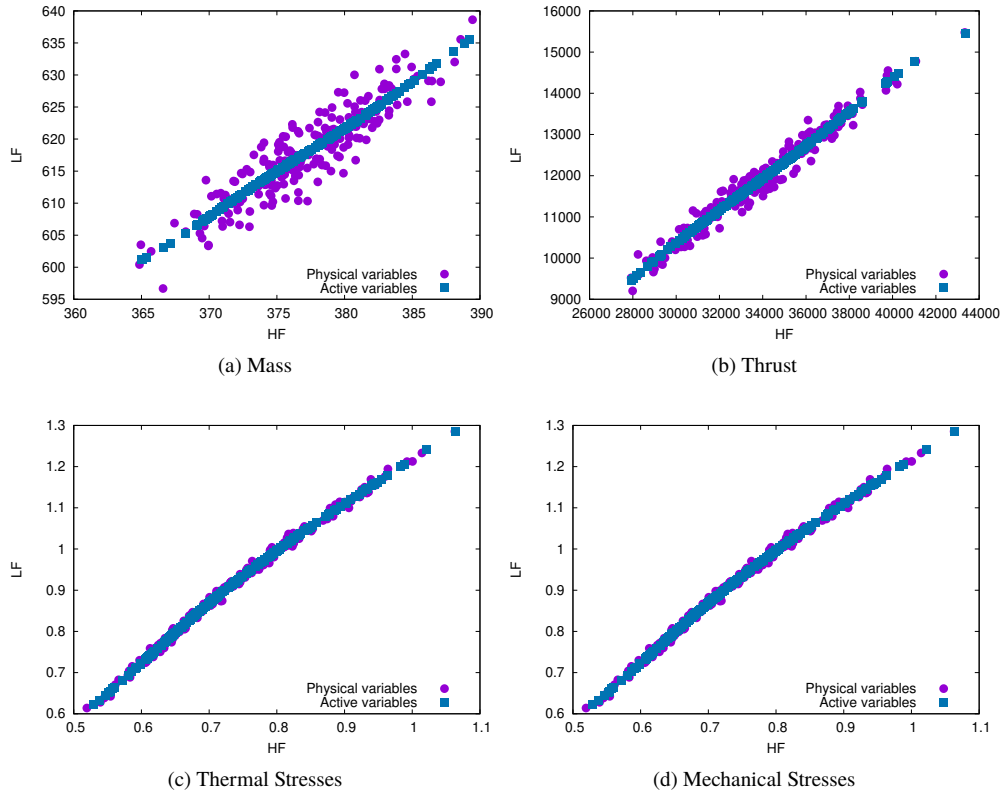


Fig. 20 Scatter plot for the QoIs w.r.t. the physical and active variables for the nozzle problem in the case of inconsistent parameterization and uncertain design input in the high-fidelity model (Scenario 3).

QoIs	ρ^2	ρ^2_{AS}	Estimator St.Dev		
			MC	ACV-1	ACV-1 (AS)
Mass	0.822	0.999	1	0.178	0.001
Thrust	0.956	0.998	1	0.044	0.002
Thermal Stress	0.982	0.998	1	0.018	0.002
Mechanical Stress	0.985	0.986	1	0.015	0.014

Table 9 (Estimated) Standard Deviation for OCV-1 and OCV-1 (AS) (normalized w.r.t. MC) for the Sequoia application problem in the case of inconsistent parameterization and uncertain design input in HF (Scenario 3).

The results reported in Table 9 are promising and predict an increase of performance of at least one order of magnitude for all the QoIs but the mechanical stresses. However, we remark that these results are obtained by resorting to a PCE regression along the active direction which is used to evaluate the correlation between the models. Albeit we can visually verify that the regression accurately capture the model trend, more rigorous tests need to be performed to confirm these results for this specific application.

B. Surrogate-based approaches: highlights

In this section, we describe two approaches for greedy adaptive refinement of stochastic expansions utilizing a multilevel, multifidelity, or multi-index construction. In the first approach, we describe a general construction that supports polynomial chaos expansion (PCE) methods, based on either projection or regression approaches, and stochastic collocation (SC) methods, built using either nodal or hierarchical interpolants. In the second section, we describe a fully integrated approach that employs adaptive sparse grid refinement across a union of random variable and model resolution dimensions.

1. Initial nozzle model exploration

The family of multilevel Monte Carlo approaches has primary strengths in insensitivity to nonsmoothness and to dimension; thus, they are robust methods that can be very effective for large-scale problems. In an initial nozzle deployment with moderate dimensionality (15 random variables), we can also evaluate methods that exploit smoothness in the QoIs by using our multilevel polynomial chaos expansion (ML PCE) approach. To support random sample sets on unstructured grids (as well as legacy data, fault tolerance, etc.), a compressed sensing strategy can be used to efficiently recover a sparse set of polynomial coefficients. From the hybridization of multilevel sampling and polynomial chaos with compressed sensing, we can replace the Monte Carlo estimator with known estimator variance with an improved estimator with parameterized estimator variance [31]. In Figure 21, the performance of this ML PCE is compared with the previous MC, MLMC and MLMF results. The ML PCE employs only realizations from the MF model on the same coarse, medium and fine discretizations of the previous MLMF results; no LF evaluations are used. For the ML PCE estimator, the actual number of realizations per-level is reported in Table 10 for different target tolerances.

Target accuracy	MF		
1E-1	11	10	10
1E-3	118	10	10
1E-4	374	10	10
1E-5	1182	35	11
1E-6	4048	132	70

Table 10 Samples profile for ML PCE as function of the target normalized MSE.

It is evident that a significant improvement in convergence rate is obtained, reducing evaluation requirements by orders of magnitude for a given accuracy target. Overall, ML PCE and MLMF MC approaches provide a useful tandem for exploiting smoothness in low-to-moderate dimension and robustly addressing nonsmooth responses in higher dimension, respectively.

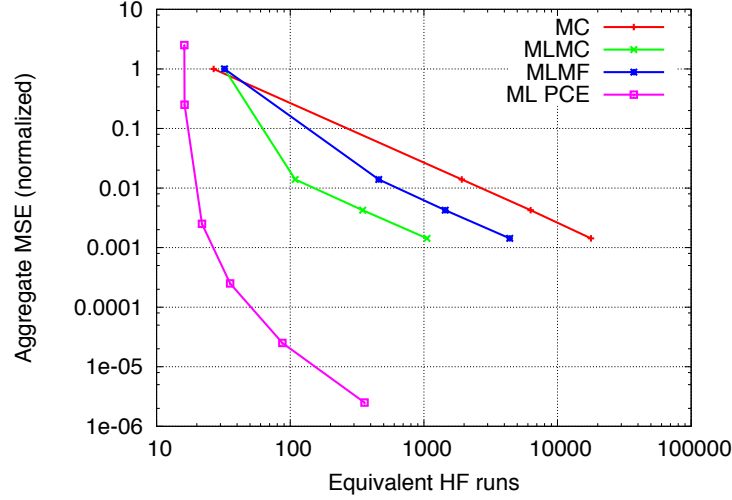


Fig. 21 Convergence comparison for sparse regression multilevel polynomial chaos expansion (ML PCE) against MC, MLMF and MLMC methods. MSE values are normalized with respect to the accuracy obtained by a set of 20 pilot samples per level.

2. Greedy multilevel/multifidelity adaptation

In the area of multilevel PCE approaches, we have previously explored bi-fidelity approaches [32, 33], followed by approaches that shaped a sample profile based on convergence rate estimation for statistical estimator variance, on satisfying the restricted isometry property for compressed sensing, and on greedy multilevel refinement [34]. Of these, the greedy multilevel approach has been the most successful, and we describe the latest advancements here.

In a greedy approach, one competes a set of refinement candidates in order to select one or more *best* candidates, as indicated by maximal change in statistical quantities of interest (QoI) per unit of computational cost. This supports goal orientation in that the statistical QoI can involve expected values, variances, probabilities of failure, or other statistical goals. Each selected refinement is then used to generate additional candidates that continue the evolution of the approximation in the most fruitful directions.

Multilevel-multifidelity schemes add an additional dimension to greedy refinement, in that now we compete one or more candidates per level/fidelity and normalize their impact based both on the number of evaluations and on the relative cost of the evaluations for a particular level or fidelity. Level candidate generators of interest include:

- *Uniform refinement*: coarse-grained refinement with one expansion order / grid level candidate per model level
 - Tensor / sparse grids: projection PCE and nodal/hierarchical SC
 - Regression PCE: least squares / compressed sensing using a fixed sample ratio
- *Anisotropic refinement*: coarse-grained refinement with one expansion order / grid level candidate per model level
 - Tensor / sparse grids: projection PCE and nodal/hierarchical SC
- *Index-set-based refinement*: fine-grained refinement with multiple index set candidates per model level; exponential growth in size of candidate set with dimension.
 - Generalized sparse grids: projection PCE and nodal/hierarchical SC
- *Basis selection* [35]: coarse-grained refinement with a few expansion order frontier advancements per model level
 - Regression PCE

Model problem We present preliminary results for the steady state diffusion problem defined as follows:

$$-\frac{d}{dx} \left[a(x, \xi) \frac{du}{dx}(x, \xi) \right] = 10, \quad (x, \xi) \in (0, 1) \times I_{\xi}, \quad (15)$$

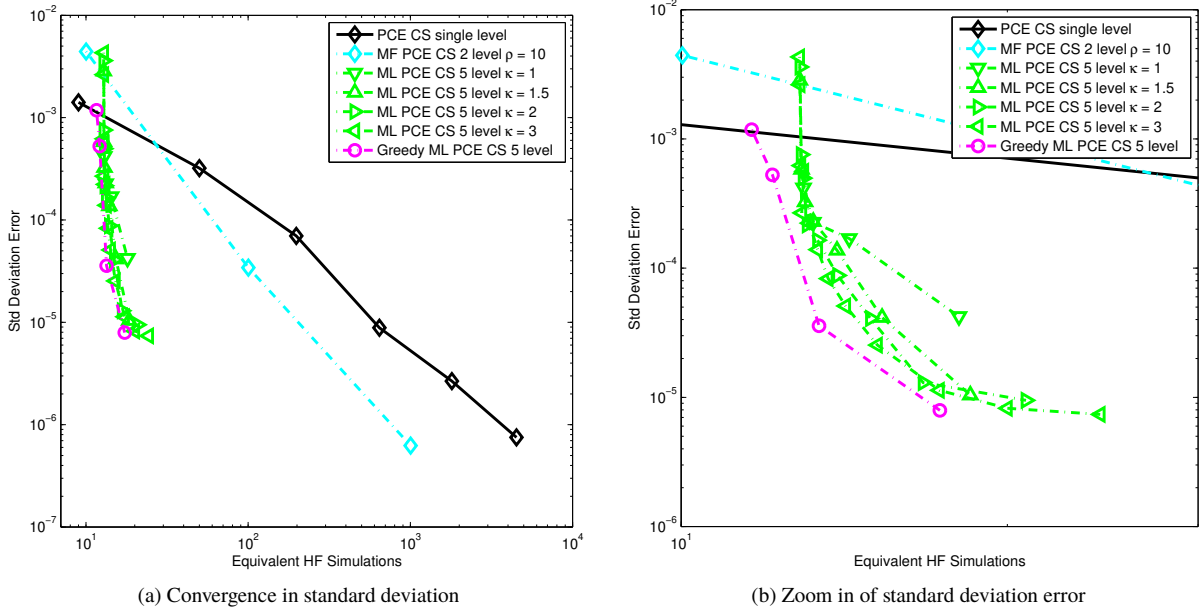


Fig. 22 Convergence for greedy multilevel PCE based on compressed sensing. Test problem is steady state diffusion with nine random variables and one, two, or five discretization levels.

where x is the spatial coordinate, ξ a vector of independent random input parameters and $a(x, \xi)$ denotes the (random) diffusivity field. The following Dirichlet boundary conditions are also assumed

$$u(0, \xi) = 0, \quad u(1, \xi) = 0. \quad (16)$$

We are interested in quantifying the uncertainty in the solution u at specified spatial locations: $\bar{x} = 0.05, 0.5, 0.95$. We represent the random diffusivity field a using the following expansion

$$a(x, \xi) = 1 + \sigma \sum_{k=1}^d \frac{1}{k^2 \pi^2} \cos(2\pi k x) \xi_k \quad (17)$$

where $\xi_k \in [-1, 1]$, $k = 1, \dots, d$ are bounded random variables. The form of Eq. (17) is similar to that obtained from a Karhunen-Loève expansion. If the variables are i.i.d. uniform in $[-1, 1]$ the the diffusivity satisfies satisfies the auxiliary properties

$$\mathbb{E}[a(x, \xi)] = 1 \quad \text{and} \quad 1 - \frac{\sigma}{6} < a(x, \xi) < 1 + \frac{\sigma}{6}. \quad (18)$$

For the particular test case in this section $d = 9$, therefore $I_\xi = [-1, 1]^9$ and five uniform spatial discretizations of 4, 8, 16, 32 and 64 elements are selected.

Greedy multilevel PCE with compressed sensing In Figure 22, we present results for greedy multilevel PCE with uniform refinement of candidate expansion orders using compressed sensing. The efficacy of the greedy approach is demonstrated by comparing its convergence envelope to several distinct sample allocations for the same number of levels (previous rate estimation results using compressed sensing). This envelope encompasses all of the other sample allocations.

In Table 11, we present the final sample profiles for this greedy multilevel compressed sensing approach. Here, the minimum evaluation requirement per level is 9 samples, resulting from initialization using a first-order basis with 10 terms and an enforced collocation ratio of 0.9 (note: a second-order candidate in nine dimensions contains 55 terms such that the collocation ratio requires 50 terms).

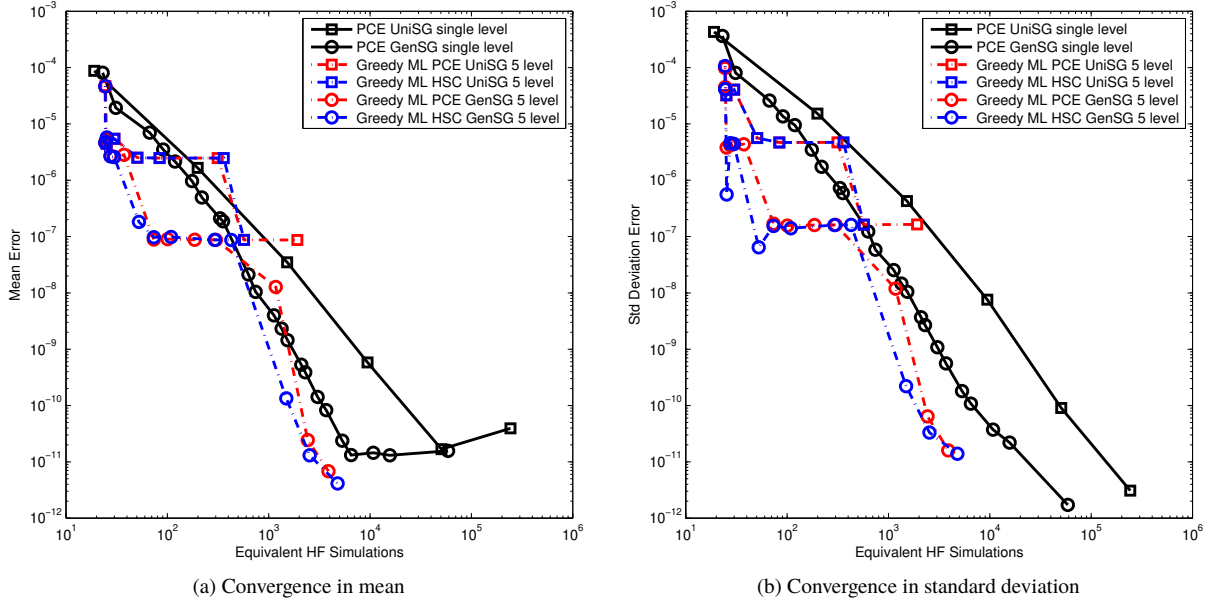


Fig. 23 Convergence for greedy multilevel PCE based on (generalized) sparse grids. Test problem is steady state diffusion with nine random variables and one or five discretization levels (solid and dashed lines, respectively).

Greedy multilevel PCE with generalized sparse grids The greedy approach can be also based on sparse grid refinement built using uniform refinement candidates (level-based) or generalized sparse grid candidates (index set-based). For these two cases, we report in Figure 23 the convergence for mean and standard deviation, where it is evident that the greedy multilevel polynomial chaos (ML PCE) and multilevel hierarchical stochastic collocation (ML HSC) variants outperform their corresponding single-level benchmarks. Moreover, the ML HSC case shows initial promise relative to ML PCE since it can directly calculate hierarchical increments in statistics and avoid loss of precision due to subtractive cancellation. Finally, the greedy uniform refinement cases appear promising early on, but then stagnate for the highest grid levels that are currently affordable.

In Table 12, we present the final sample profiles for greedy multilevel PCE using generalized sparse grid refinement candidates. In nine dimensions for generalized sparse grids initiated from level-zero, all levels incur a minimum $2n + 1 = 19$ evaluation cost due to the initial set of level-one candidate index sets. The table shows how this minimum evaluation requirement is eventually replaced with an optimized multilevel sample profile, which also sheds light on the relative efficiency for low sample levels.

3. Adaptive Multi-index collocation

The multi-level surrogate based approaches we described in the previous section are useful when the level of fidelity is controlled by a single hyper-parameter specifying the level of discretization. In many applications, however, multiple hyper-parameters may control the model discretization, such as the mesh and time step sizes. In these situations

Conv Tol	N_1	N_2	N_3	N_4	N_5
1.e-1	198	9	9	9	9
1.e-2	644	198	9	9	9
1.e-3	1802	644	9	9	9
1.e-4	4505	1802	50	9	9

Table 11 Final sample profiles for greedy multilevel compressed sensing applied to steady state diffusion (9 random variables, 5 discretization levels).

Conv Tol	N_1	N_2	N_3	N_4	N_5
1.e-2	43	23	19	19	19
1.e-4	211	83	19	19	19
1.e-6	391	271	156	19	19
1.e-8	1359	743	327	59	19
1.e-10	3535	2311	1039	391	19
1.e-12	10319	5783	2783	1343	43
1.e-14	26655	14991	8063	3703	1535

Table 12 Final sample profiles for greedy multilevel refinement applied to steady state diffusion (9 random variables, 5 discretization levels).

multi-index stochastic collocation (MISC), can be used to deal with multivariate hierarchies with multiple refinement hyper-parameters [36].

In our recent work [37] we developed an adaptive extension of MISC and use this algorithm to reduce the cost of quantifying uncertainty in the performance of a supersonic jet engine nozzle. Our work utilizes a sparse grid approximation that spans both random dimensions and model resolution dimensions and uses generalized sparse grid refinement to select model evaluations from multiple candidates based on maximizing benefit per unit cost. In the following we will describe our algorithm and summarize our existing numerical results.

Given a sequence of models $\{\alpha \mid \alpha \leq (l^{\alpha_1}, \dots, l^{\alpha_{n_\alpha}})\}$ of varying physical fidelity, which can be used to approximate the QoI f , MISC approximates the QoI as a linear combination of low-resolution tensor-product interpolants

$$\hat{f}_{\mathcal{J}}(z) = \sum_{[\alpha, \beta] \in \mathcal{J}} c_{\alpha, \beta} \hat{f}_{\alpha, \beta}(z). \quad (19)$$

The tensor-product interpolants are themselves a weighted linear combination of tensor-product of univariate Lagrange polynomials

$$\phi_{i,j}(z_i) = \prod_{k=1, k \neq j}^{m_{\beta_i}} \frac{z_i - z_i^{(k)}}{z_i^{(j)} - z_i^{(k)}}, \quad i \in [d], \quad (20)$$

defined on a set of univariate points $z_i^{(j)}, j \in [m_{\beta_i}]$ (we use Clenshaw-Curtis points). Specifically the multivariate interpolant is given by

$$\hat{f}_{\alpha, \beta}(z) = \sum_{j \leq \beta} \hat{f}_{\alpha}(z^{(j)}) \prod_{i \in [d]} \phi_{i, j_i}(z_i). \quad (21)$$

The partial ordering $j \leq \beta$ is true if all the component wise conditions are true.

Model problem The accuracy and efficiency of the MISC approximation is dependent on the index set $\mathcal{J}$. In Figure 24 we plot the tensor-product grids and corresponding interpolants that make up an isotropic MISC approximation of the function

$$f(z) = \cos\left(\frac{\pi}{4}\left(z_1 + \frac{1}{2}\right)\right), \quad (22)$$

with $\mathcal{J}(l) = \{[\alpha, \beta] \mid (\max(0, l - 1) \leq \|\alpha + \beta\|_1 \leq l + n_\alpha + n_\beta - 2)\}$. The MISC approximation uses a sequence of models $\hat{f}_{\alpha_1} = \cos(\frac{\pi}{4}(z_1 + \frac{1}{2} + \epsilon_{\alpha_1}))$ where $\epsilon_i > \epsilon_{i+1} \geq 0$ to approximate f . For each of these models MISC builds an interpolant $\hat{f}_{\alpha_1, \beta}$ that converges to $\hat{f}_{\alpha_1}$ as β increases. Refinement of both α_1 and β is needed for the interpolants to converge to f . MISC achieves its efficiency by allocating resources to balance physical prediction bias and error of the stochastic interpolant. For the isotropic MISC, the number of evaluations allocated to an approximate model $\hat{f}_{\alpha_1}$ decreases as the level of fidelity α_1 increases. For example, 5 samples are allocated to evaluating $\hat{f}_0$, 3 to $\hat{f}_1$ and 1 to $\hat{f}_2$. Despite the fact only one sample is allocated to evaluating $\hat{f}_2$, the MISC approximation, shown in the right plot of Figure 25, is more accurate than any of the tensor-product interpolants that make up the MISC approximation (gray).

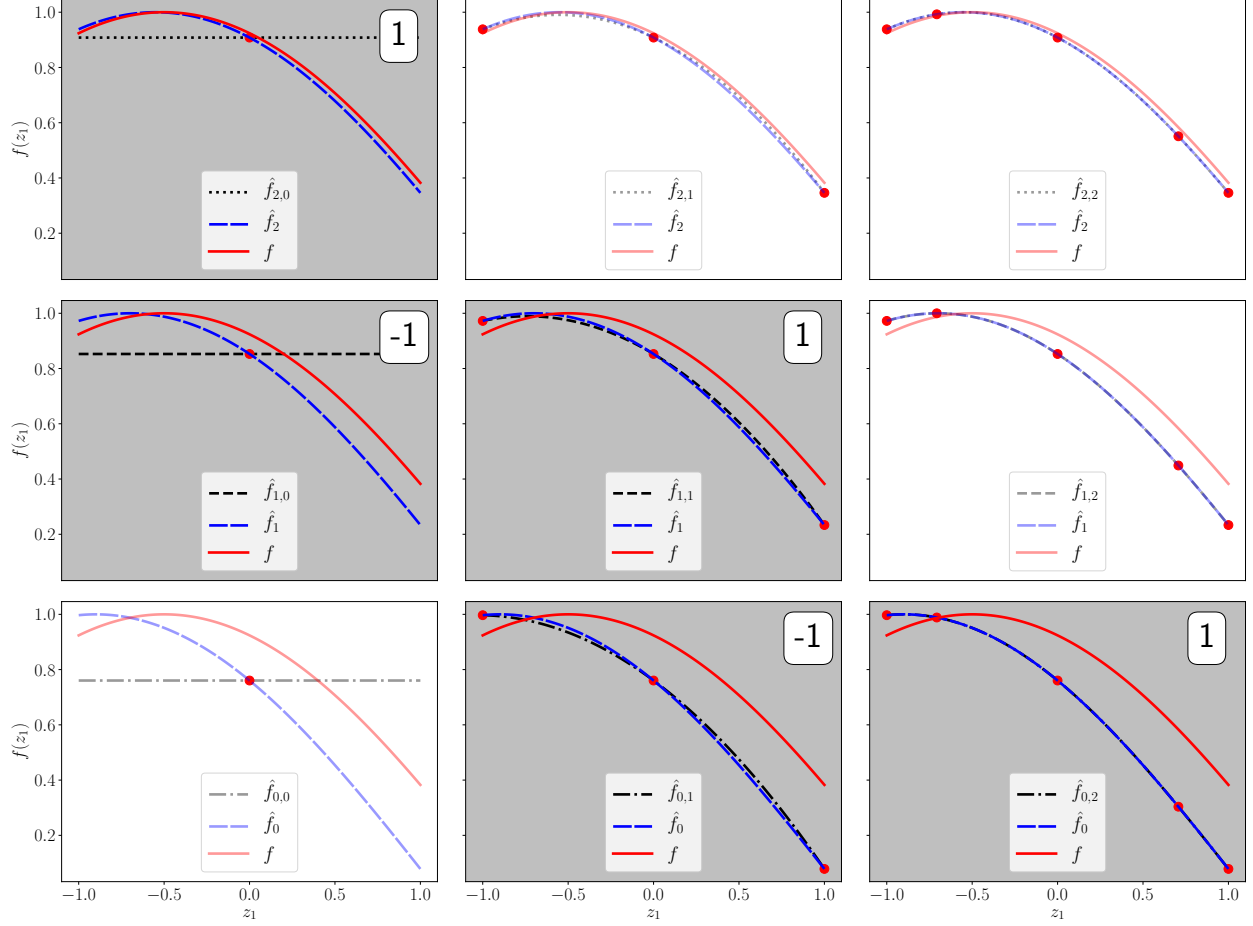


Fig. 24 The tensor-product grids and interpolants of the level $l = 2$ isotropic MISC approximation $\hat{f}_{\mathcal{J}(2)}(z)$. The grids plotted with gray background are included in the MISC approximation, whereas the remaining grids are ignored. The numbers in the top-right of each plot are the coefficients $c_{\alpha,\beta} = (-1)^{l-\|\alpha+\beta\|_1} \binom{d-1}{l-\|\alpha+\beta\|_1}$ in (19). The dash-dotted black lines represent the tensor-product interpolants $\hat{f}_{\alpha,\beta}$. These interpolants converge to the lower fidelity models $\hat{f}_\alpha$ (dashed blue lines) as β increases, and to the high-fidelity model f (solid red line) as both α and β are increased.

Understanding of this observation can be gained by viewing the discrepancies between models $\hat{f}_{i+1} - \hat{f}_i$ depicted in the middle plot of Figure 25. The magnitude of these discrepancies decreases substantially as α_1 increases. Thus a constant approximation $\hat{f}_{2,0}$ of $\hat{f}_2$ is sufficient to balance the deterministic and stochastic errors. Specifically the constant approximation is useful because it reduces the bias in the MISC interpolant by decreasing the physical discretization error. In contrast the approximation $\hat{f}_{0,2}$ is useful as it captures the variation of the function f with respect to the variable z_1 thereby reducing the stochastic error.

The isotropic MISC approximation balances computational expense with accuracy. The same number of model evaluations are used for two physical discretization parameters $\alpha \neq \alpha^*$ if $\|\alpha\|_1 = \|\alpha^*\|_1$. The isotropic formulation does not account for differences in the cost of evaluating the two models. In [37] we propose using a greedy algorithm to adaptively choose the most cost effective interpolants to add to the MISC approximation. Our adaptive algorithm is based on the algorithm proposed in [38]. The algorithm greedily chooses tensor product interpolants that maximize the change in the interpolant whilst minimizing cost. Specifically letting

$$\Delta E_{\alpha,\beta} = \|\hat{f}_{\mathcal{J} \cup \{\alpha,\beta\}} - \hat{f}_{\mathcal{J}}\| \quad \Delta W_{\alpha,\beta} = |\text{Work}[\hat{f}_{\mathcal{J} \cup \{\alpha,\beta\}}] - \text{Work}[\hat{f}_{\mathcal{J}}]| \quad (23)$$

respectively denote the difference between two successive MISC approximations and the work needed to update the

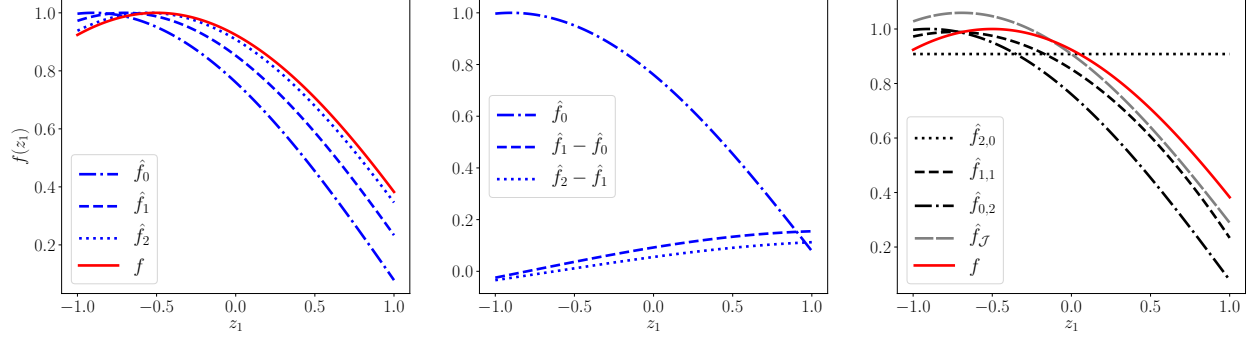


Fig. 25 (Left) The function f (22) and the approximations $\hat{f}_{\alpha_1}$. (Middle) The discrepancies between successive approximations $\hat{f}_{\alpha_1}$ and $\hat{f}_{\alpha_1+1}$. (Right) The most accurate (largest β_1) single fidelity interpolants $\hat{f}_{\alpha_1, \beta_1}$, which use only evaluations of one approximate model $\hat{f}_{\alpha_1, \alpha_1} = 0, 1, 2$, and the MISC approximation $\hat{f}_{\mathcal{J}}$ which combines each of these interpolants (and some less accurate interpolants).

MISC approximation, we choose grids using the following indicator to guide refinement

$$\gamma_{\alpha, \beta} = \frac{1}{\Delta W_{\alpha, \beta}} \left(\kappa \Delta E_{\alpha, \beta}^{\mu} + (1 - \kappa) \Delta E_{\alpha, \beta}^{\sigma^2} \right), \quad \kappa \in [0, 1], \quad (24)$$

where

$$\Delta E_{\alpha, \beta}^{\mu} = \frac{1}{|\hat{f}_{0,0}|} \left| \mathbb{E}[\hat{f}_{\mathcal{J} \cup \{\alpha, \beta\}}(z)] - \mathbb{E}[\hat{f}_{\mathcal{J}}(z)] \right| \quad \Delta E_{\alpha, \beta}^{\sigma^2} = \frac{1}{|\hat{f}_{0,0}|^2} \left| \text{Var}[\hat{f}_{\mathcal{J} \cup \{\alpha, \beta\}}(z)] - \text{Var}[\hat{f}_{\mathcal{J}}(z)] \right| \quad (25)$$

The index $[\alpha, \beta]$ with the largest refinement indicator is refined by adding all indices $[\alpha, \beta] + \mathbf{e}_k$, $k \in [n_{\alpha} + n_{\beta}]$ that satisfy the following (downward-closed) admissibility criterion

$$[\alpha, \beta] + \mathbf{e}_k - \mathbf{e}_j \in \mathcal{J} \text{ for } j \in [n_{\alpha} + n_{\beta}], l_k > 1. \quad (26)$$

UCAV nozzle problem We report here some results regarding the performance of the adaptive multi-index Stochastic Collocation (AMISC) algorithm for four QoIs, namely the mass, thrust, the thermal and mechanical stress failure criteria.

For this problem, we are interested in approximating four QoIs, so we must adjust slightly the refinement indicator introduced above. More precisely, we use a worst case error indicator that refines the multi-index set in the AMISC approximation according to the largest error across the set of QoIs. Letting $\gamma_{\alpha, \beta}^k$ denote the error indicator Eqn. (24), we use

$$\gamma_{\alpha, \beta} = \max_{k \in [4]} \gamma_{\alpha, \beta}^k \quad (27)$$

across the set of k QoI. Moreover we select $\kappa = 1/2$ to obtain a balanced evaluation of both mean and variance.

We first compare the performance of the AMISC algorithm with a single-level adaptive sparse grid in Figure 26 (only thrust and temperature layer failure are reported for brevity). Work is measured relative to the cost of the highest fidelity. We terminate refinement of each approximation after the cost of 3400 high-fidelity samples is reached. We observe that AMISC can approximate each QoI with the same accuracy as a single fidelity approximation at only a fraction of the cost. The AMISC approximation at some points on the convergence curves is an order of magnitude less expensive. This approaches the limit of potential acceleration, as there is only a factor of 20 difference in cost between the cheapest and most expensive models.

Note that the convergence curves of the approximation of thrust (top right) is noisy in the right tail. This is related to the tolerance we use for the non-linear CFD solver; in particular, the error is saturating at the error dictated by this tolerance.

For the thermal-layer failure criteria we observe that AMISC is cheaper to construct for accuracies up to 10^{-4} , however the single-fidelity approximation is cheaper to construct for higher accuracy. This situation is caused by

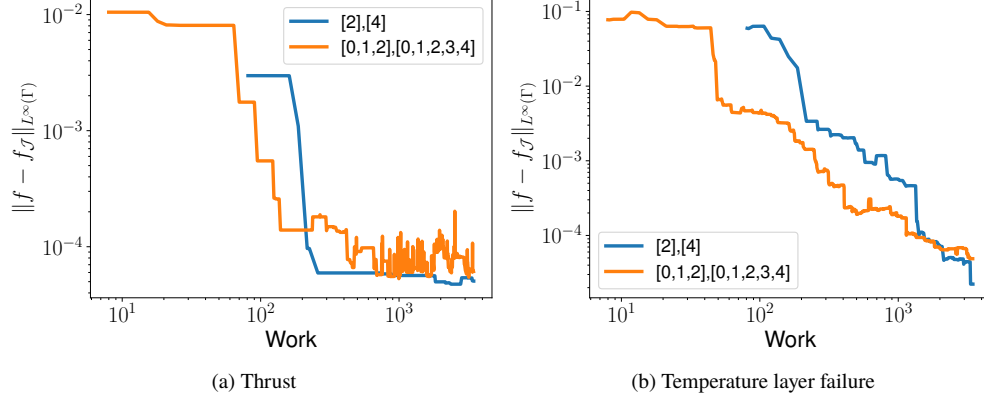


Fig. 26 Comparison of the convergence of single-fidelity and multi-fidelity approximations of thrust and temperature layer failure. The $([0, 1, 2], [0, 1, 2, 3, 4])$ specify the physical discretization indices used for the AMISC approximation and $([2], [4])$ the physical discretization indices used for the single high-fidelity sparse grid approximation.

the error of the two coarsest structural meshes to be smaller than the deterministic prediction error of those models. Therefore, only the high-fidelity structural model can reduce the error. The downward closed assumption for the sparse grid index set implies that the coarsest structural meshes are always evaluated before the highest fidelity structural mesh even if they are not effective. Therefore an overhead cost is associated to these low-fidelity evaluations and the single-level sparse grid approximation of the load-layer failure criteria will eventually become more efficient than the AMISC approximation.

In Figure 27, we plot the number of model evaluations allocated to each model discretization. The AMISC algorithm determined that the CFD and structural model are both large sources of deterministic prediction error. This is reflected by the large number of samples allocated to the most refined CFD and structural models. However, it is not necessary to refine both the CFD and structural meshes simultaneously. This is indicated by the fact only one sample is assigned to any model indexed by (α_1, α_2) with both α_1 and α_2 non zero. The error in the CFD model output, which is used as input to the structural model, does not significantly affect the prediction of the structural quantities of interest – the load layer and temperature layer failure criteria.

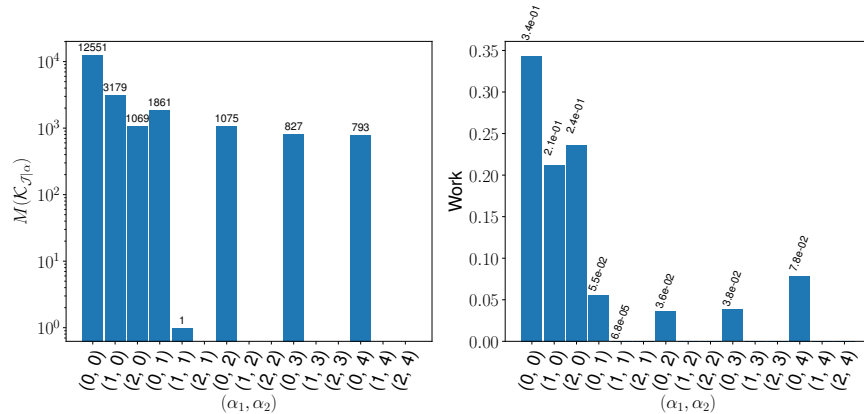


Fig. 27 The number of simulation and fraction of total work assigned to each nozzle model.

The advantage of surrogate-based approach with respect to sampling-based methods is that once the surrogate has been constructed it can be used to compute additional metrics of interest. For instance, for this problem we used AMISC to

compute the sensitivity indices, namely the Sobol indices, that evaluate the contribution of each parameter or group of parameters to the overall variance. In Figure 28, we plot the dominant sources of uncertainty in the nozzle predictions. We only report the parameter combinations that contribute to a portion of the first 99.9% of variance of at least one nozzle QoI. It is interesting to note that, for each QoI, one parameter acting independently of all other parameters contributes over 90% of the total variance. Moreover almost all contributions to 99.9% of the variance of each QoI are caused by individual parameters. Only two second-order interactions are significant. However, the convergence curves suggest that resolving multivariate interactions becomes important as we drive the surrogate error to tight tolerances.

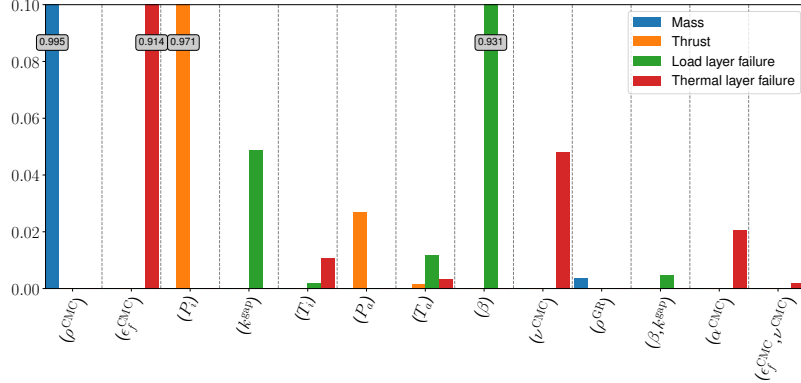


Fig. 28 Sobol indices that contribute a portion of the first 99.9% of the variance of at least one nozzle QoI. The labels of the x-axis represent the parameter of the Sobol index. The maximum value of any Sobol index is one. We zoom in on the y-scale to better highlight the smaller Sobol index values, and report values greater than the 0.1 plot scale in the gray boxes.

The most sensitive parameter identified for each QoI matches physical intuition. The density of the CMC heat layer ρ^{CMC} is the largest contribution to the uncertainty in the mass of the nozzle, consistent with the observations that the minimum density of this layer is an order of magnitude larger than the maximum density of the other layers and the volume of this layer is greater than any other layer. The inlet stagnation pressure P_i has the greatest influence on thrust. The heat transfer coefficient β which transfers heat from the nozzle to the environment has the greatest impact on the Load layer temperature failure ratio, which is the ratio of the temperature in the load layer to the maximum allowable temperature. When β is decreased, less heat is able to escape the nozzle and thus the temperature in the nozzle increases. Finally, the failure strain in the CMC heat layer ϵ_f^{CMC} has the greatest impact on the failure of the thermal failure.

C. Multifidelity modeling with Bayesian networks

The surrogate methodologies described above require explicit ordering and hierarchical assumptions. However, the approximate control variate methodology has shown that enforcing such assumptions, which lead to MLMC and MFMC in the sampling case, creates inefficiency in the multifidelity models. In contrast to the pure sampling based approaches – the surrogate methodology can be much more brittle to inaccuracies in the modeling. For instance, it might require significantly more resources to accurately approximate discrepancies between two models which were ordered hierarchically, but in practice are actually not so. For example, it is common that a computational simulation of a high-fidelity model may not be fully converged over the entire input space and therefore exhibit some noisy behavior in a quantity of interest or its derivative.

We propose a new methodology for leveraging multiple information sources that first specifies hidden variables for each model and then *learns the structure between the latent variable models*. These latent variables serve to *explain* any perceived or measured relationships between the information sources or models. Once such a network structure is learned, any additional analysis including uncertainty quantification, experimental design, and optimization can be made more efficient by exploiting the structure.

Specifically, our new model for the relationships between physical models consists of a network of M latent variables $\theta = \{\theta_1, \dots, \theta_M : \theta_i \in \Theta_i \subseteq \mathbb{R}^{p_i}\}$, each of which can be multivariate with p_i univariate parameters. These latent variables correspond to each of the models we seek to manage. For example, suppose we assume a

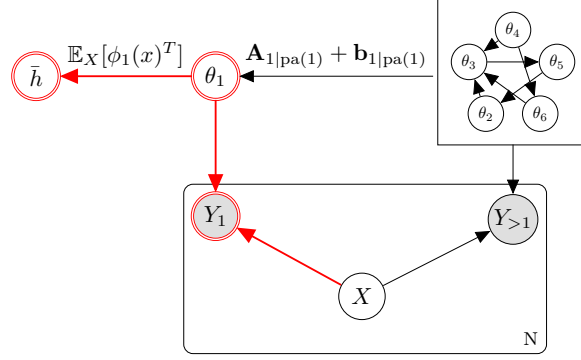


Fig. 29 Multifidelity UQ for a network-informed model. Data Y generated from uncertain parameters X are first used to construct the network of low-fidelity models. Then this network feeds into the high fidelity model. Statistics can then be extracted from this high-fidelity model. For instance here, we have $\bar{h} = \mathbb{E}[Q]$. Furthermore, note that the predictions are now probabilistic over the latent variable space.

linear relationship between models and latent parameters, then we define a set of basis functions for each model $\phi = \{\phi_1, \dots, \phi_M : \phi_i : X_i \rightarrow \mathbb{R}^{p_i \times d_{y_i}}\}$. Each model output becomes a random variable

$$Q_i = \widehat{Q}_i(x_i, \theta_i) = \phi_i^T(x_i)\theta_i \equiv \sum_{k=1}^{p_i} \phi_{ik}(x_i)\theta_{ik}, \quad x_i \in X_i. \quad (28)$$

Essentially, we introduce a new source of uncertainty due to incomplete knowledge of the latent variables. While we cannot reduce the original source of uncertainty (the input variables), we can reduce the uncertainty due to the latent variables as more data is obtained. The primary purpose of introducing the latent variables in the UQ context is to provide *explanatory variables* for observed correlations in model outputs. Furthermore, they allow us to consider relationships between models with different input/output spaces by developing relationships between their respective random variables. This notion of explanatory behavior is especially important in managing physical models where engineering and science intuition can often qualitatively specify the relationships between models. As an example, these latent variables could be the rotated variables obtained from active subspace approach described above and the basis functions could be the rotation matrices.

Next introduce a network relationship between each of the parameters θ_i to obtain a *network-informed high-fidelity model*. A graph $\mathcal{G} = (\mathbf{V}, \mathbf{E})$ is a tuple of vertices (or nodes) and edges. A Bayesian network (BN) structure is a directed acyclic graph whose nodes are random variables and edges are conditional probability distributions. The network encodes the set of conditional independence assumptions, called local Markov assumptions. The parents of random variable X_i in a graph are denoted as $X_{\text{pa}(i)}$. The set of conditional probability distributions (CPDs), given the graph, is denoted as $\mathbf{p} = \{P(X_i | X_{\text{pa}(i)}) : X_i \in \mathbf{V}\}$.

For simplicity, we consider linear-Gaussian CPDs so that

$$\theta_i | \theta_{\text{pa}(i)} \sim \mathcal{N}(\mathbf{A}_{i|\text{pa}(i)}\theta_{\text{pa}(i)} + \mathbf{b}_{i|\text{pa}(i)}, \mathbf{\Gamma}_{i|\text{pa}(i)}),$$

where $\mathbf{c} \sim \mathcal{N}(\mathbf{b}, \mathbf{\Gamma})$ denotes that $\mathbf{c}$ is normally distributed with mean $\mathbf{b}$ and covariance $\mathbf{\Gamma}$. Then, the expectation of the high-fidelity model over the input uncertainty is a *random variable* depending on the latent parameters

$$\mathbb{E}_X[\widehat{Q}(x_1, \dots, x_M, \theta_1, \dots, \theta_M)] = \mathbb{E}_X[\widehat{Q}(x_1, \theta_1)] = \mathbb{E}_X[\phi_1(x_1)^T]\theta_1, \quad (29)$$

where the first equality uses the Markov property (implicitly indicated by Equation (28)) that indicates that given θ_1 and x_1 , f is independent of the other latent parameters. Together this procedure is shown in Figure 29

Learning the *hyperparameters* of a fixed graph is an important aspect before using it for prediction. Let $\tau = \{\tau_1, \dots, \tau_M\}$ denote the set of hyperparameters defining the distribution of $\theta_i | \theta_{\text{pa}(i)}$. For example for a network of

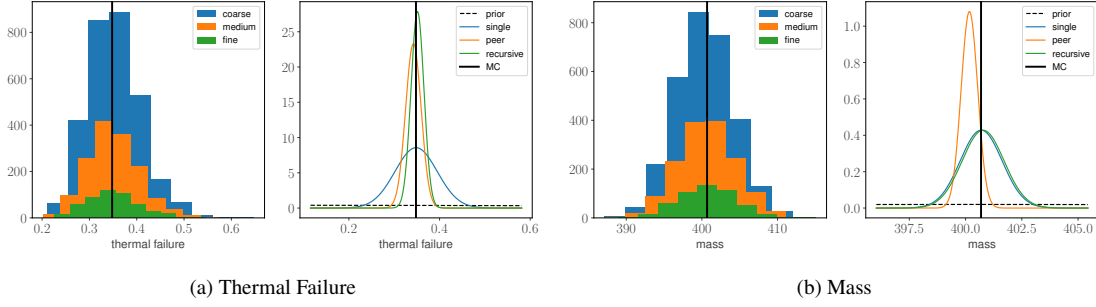


Fig. 30 Posteriors over expectation of the thermal failure probability and the mass as a function of 40 uncertain inputs. The left figure in each subplot shows the raw data, with the the black line indicating the mean of the finest mesh. The right plot in each subplot shows the posterior and the prior of the estimated mean of the high-fidelity model for both a peer and recursive model graph.

$M - 1$ low-fidelity peer models we factor the joint distribution, given the hyperparameters, according to

$$\begin{aligned} \theta_j | \tau_j &\sim \mathcal{N}(\mathbf{b}_j, \mathbf{\Gamma}_j) \quad \text{for } j = 2, \dots, M \\ \theta_1 | \theta_2, \dots, \theta_M, \tau &\sim \mathcal{N}\left(\mathbf{b}_{1|pa(1)} + \sum_{j=2}^M \mathbf{A}_{1|j} \theta_j, \mathbf{\Gamma}_{1|pa(1)}\right) \end{aligned}$$

where $\mathbf{b}_j \in \mathbb{R}^{p_j}$, $\mathbf{A}_{1|j} \in \mathbb{R}^{p_1 \times p_j}$, $\mathbf{\Gamma}_j \in \mathbb{R}^{p_j \times p_j}$, and $\mathbf{b}_{1|pa(1)} \in \mathbb{R}^{p_1}$, $\mathbf{\Gamma}_{1|pa(1)} \in \mathbb{R}^{p_1 \times p_1}$. In this case the hyperparameters are

$$\begin{aligned} \tau_j &= \{\mathbf{b}_j, \mathbf{\Gamma}_j\} \quad \text{for } j = 2, \dots, M \\ \tau_1 &= \{\mathbf{b}_{1|pa(1)}, \mathbf{A}_{1|2}, \dots, \mathbf{A}_{1|M}, \mathbf{\Gamma}_{1|pa(1)}\}, \end{aligned}$$

UCAV nozzle problem Next we apply the Bayesian network approach to a SEQUOIA model consisting with 40 uncertain inputs and three mesh resolutions. Our goal here is to demonstrate the difference in posteriors due to differing graphs. The first graph is the peer graph where we have $\theta_1 = a_1 \theta_2 + a_2 \theta_3 + \xi$. The second graph is a recursive graph $\theta_1 = a_1 \theta_2 + \xi_1$ and $\theta_2 = a_2 \theta_3 + \xi_2$. The peer model indicates that the high fidelity model is a combination of the two low fidelity models. The recursive model is a standard recursive co-kriging (see e.g., [39]) approach in parametric form with a single, constant basis function. For each graph we choose the statics of ξ , ξ_1 , and ξ_2 to ensure that the marginal priors of all of the parameters $\theta_1, \theta_2, \theta_3$ are identical. Only the joint marginals differ amongst the graph. The posteriors over the mean of two quantities of interest, recall this is a Bayesian procedure and we have more than a point estimate, are shown in Figure 30.

VIII. Technical thrust area: Optimization under Uncertainty

A. DUU with Ridge Functions

We have continued to develop approaches to DUU based on identifying and exploiting low-dimensional ridge structure in objective and constraint functions. We have developed these approaches to apply to general complex, physics-based systems with multiple constraints, several design variables, and several random variables representing uncertainty. We have demonstrated the approaches on the high-fidelity MULTI-F nozzle model that has 138 design variables and 40 uncertain variables. The associated DUU problem involves one objective function that is a statistical mean over random variables and 15 constraint functions that are probabilities of certain failure modes in thrust or stresses.

The first step to solve the DUU with this approach is to search for one-dimensional ridge structure in the objective and constraints as functions of the combined design / uncertain space—a 178-dimensional space for the nozzle model. With knowledge that we will use the resulting ridge approximations for DUU, we have developed a novel *stretched sampling* approach for choosing the points in the design/uncertain space to evaluate the quantities of interest used to

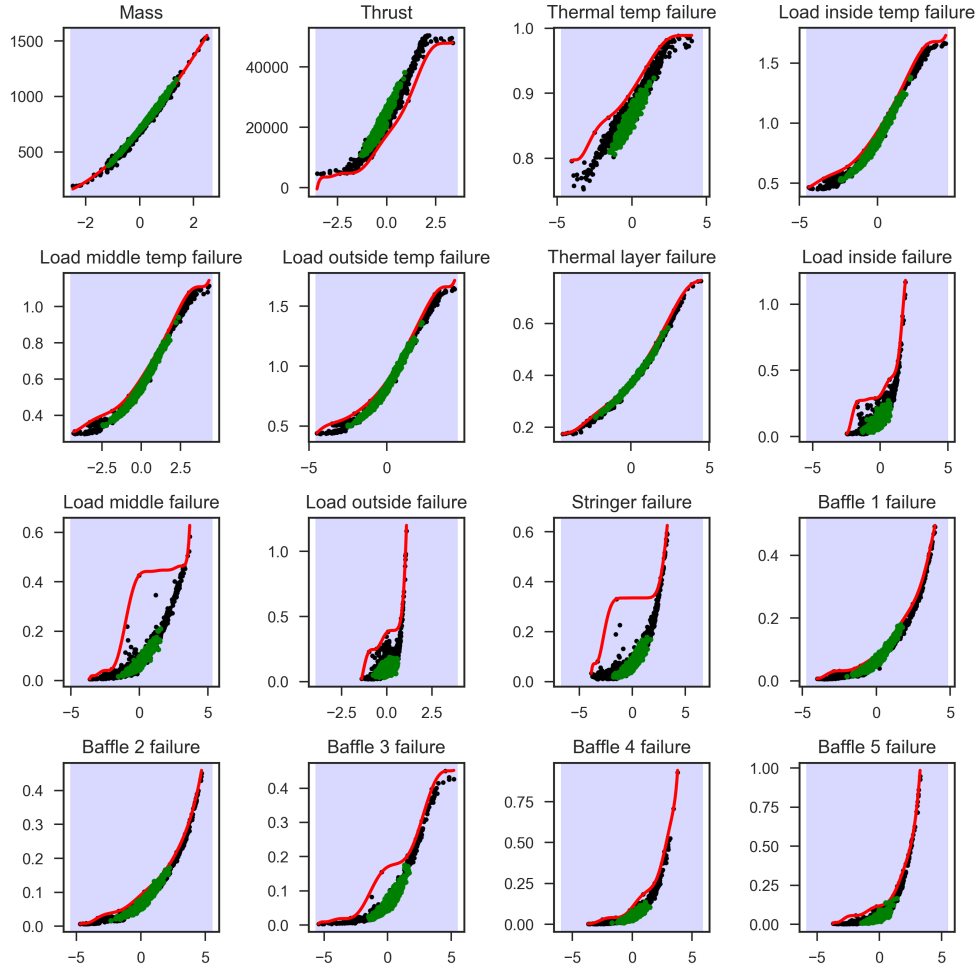


Fig. 31 Comparing stretched samples (black dots) to random samples (green). The green dots verify the ridge structure uncovered by the stretched sampling.

build the ridge approximations. The problem with standard random sampling approaches is that the samples tend to cluster toward the origin of the projected space thus not resolving potentially large variation in quantities of interest that are important for solving the DUU. Figure ?? shows the 1d shadow plots (plotting the output of interest on the vertical axis versus one linear combination of the 178 inputs on the horizontal axis) using ridge approximations (red surfaces) constructed using the stretched sampling heuristic. The black dots show the stretched training data, and the green dots are a testing set of 1000 points sampled randomly according to the input domain and its measure. Notice how the green dots cluster toward the center of the projected domains across all quantities of interest.

The red surfaces in Figure ?? are constructed to bound the training data in such a way that is conservative for the DUU. These bounding functions are used, along with the failure probability tolerances, to transform the DUU to a linear program that can be solved simply and reliably with widely available software (e.g., Matlab optimization toolbox or SciPy) in seconds on a laptop. This is possible when the bounding functions are 1-dimensional ridge and monotonic with respect to the ridge direction, which are both assumptions we can check with a simple shadow plot. We have extended these ideas somewhat. For 1d, the ridge approximations need not be monotonic with respect to the ridge direction as long as we can identify the multiple points where the approximation crosses the failure threshold; each crossing creates an additional linear constraint in the linear program. In 2d and higher, it is possible to exploit the low-dimensional structure with response surfaces and the additive structure in the projected design/uncertain space to develop model-based *safety margins*. In 2d and higher, the resulting transformed, low-dimensional optimization

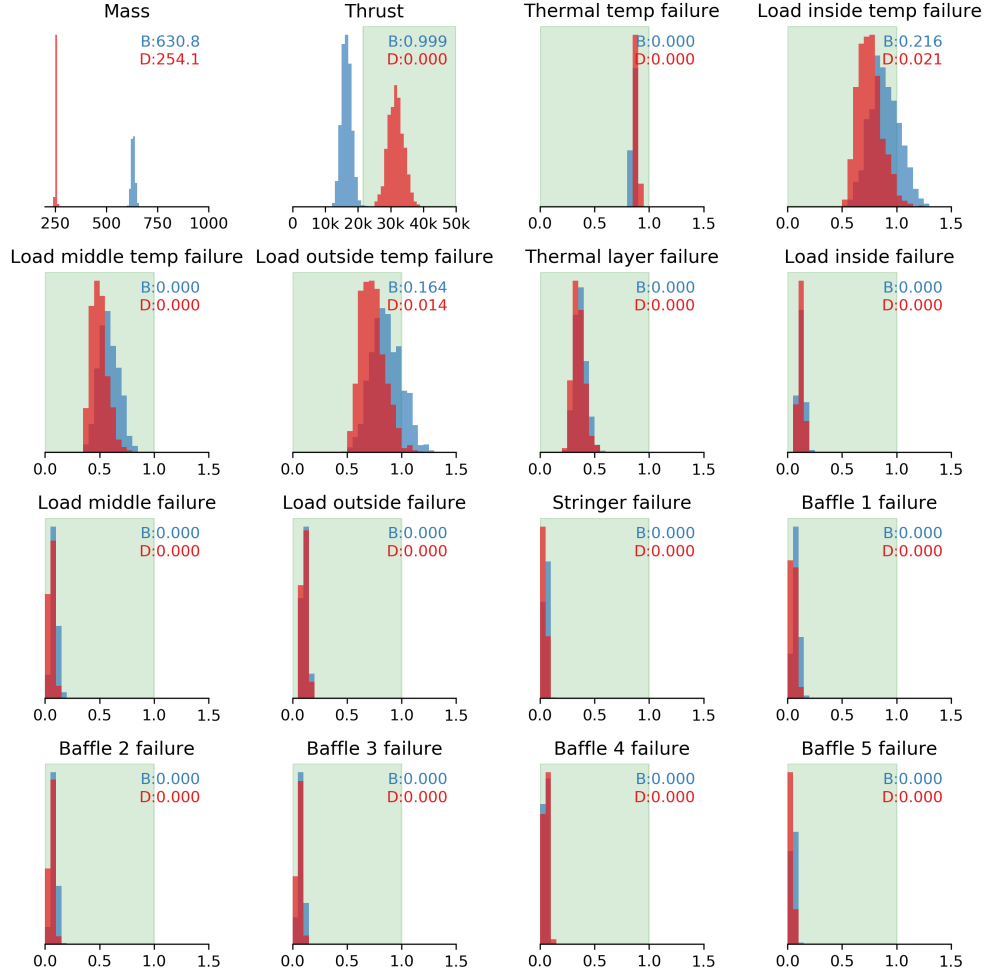


Fig. 32 Comparison between pdfs from DUU solution with ridge approximation with pdfs of baseline nozzle design.

problem is not necessarily a linear program without additional approximations.

Figure ?? compares the pdfs of the QOIs at the DUU-with-ridge-approximation solution to pdfs of the baseline nozzle design, where the failure tolerance is 10^{-3} . According to this failure criterion, the baseline design is infeasible, and the DUU-with-ridge-approximation is significantly improved—as seen in the constraint pdfs whose support is more in the feasible (green) region. However, this approximation can potentially be improved by *zooming in* to a smaller subset of the design/uncertain space and reconstructing the ridge approximations. This is something we will explore in future studies.

Figure ?? compares pdfs of the DUU-with-ridge-approximation solution to pdfs of a related deterministic optimization solution. Notably, the deterministic solution is infeasible with respect to the DUU problem—specifically in the *thrust* failure constraint. We have found a trade-off between the DUU failure tolerance and the optimized objective function, namely, the optimized objective function increases as the failure tolerance decreases. Such trade-offs do not exist in the deterministic optimization, since there is no notion of a failure *probability*, per se, in the optimization.

B. Reliability DUU

At Stanford we have focused our efforts on the reliable nozzle design problem and have made several developments over the past quarter to the problem’s methodology and solution including the reliability estimation, extension to the 3D nozzle parameterization, and comparison with an industry standard method for handling stochastic optimization

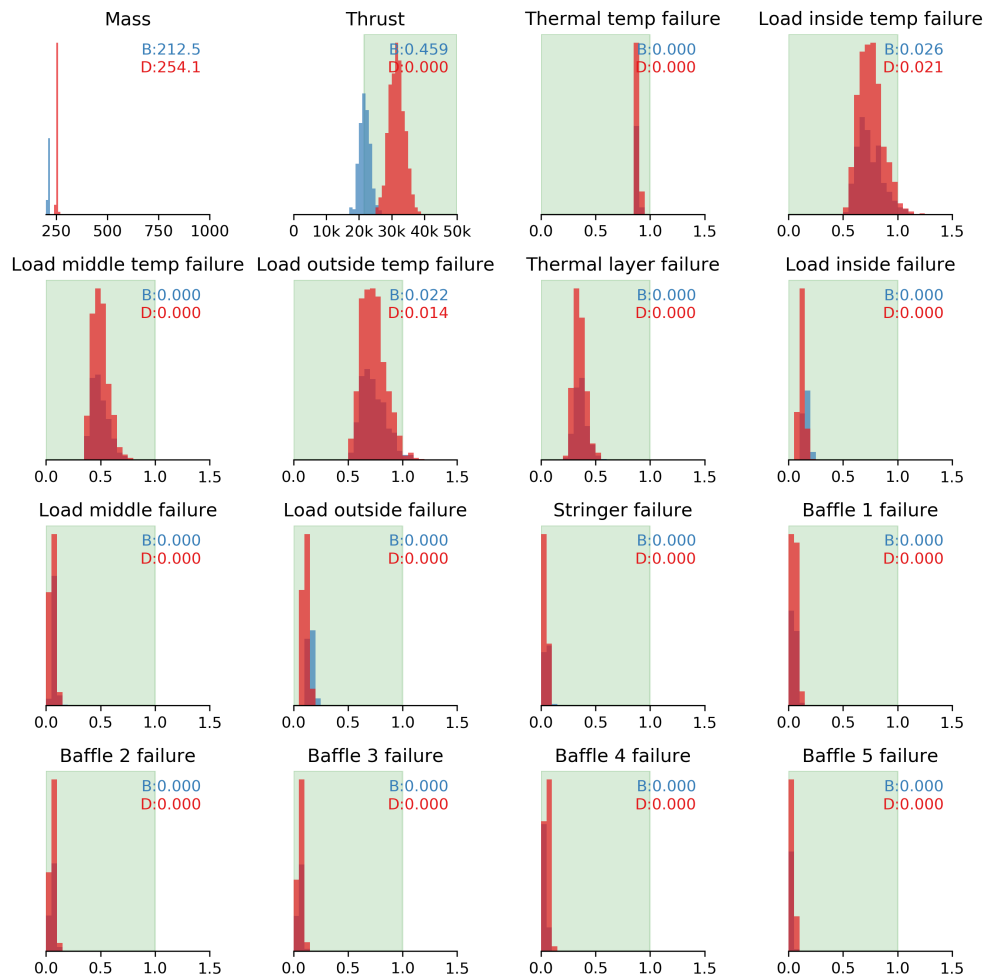


Fig. 33 Comparison between pdfs from DUU solution with ridge approximation with pdfs of a deterministic-optimized nozzle design.

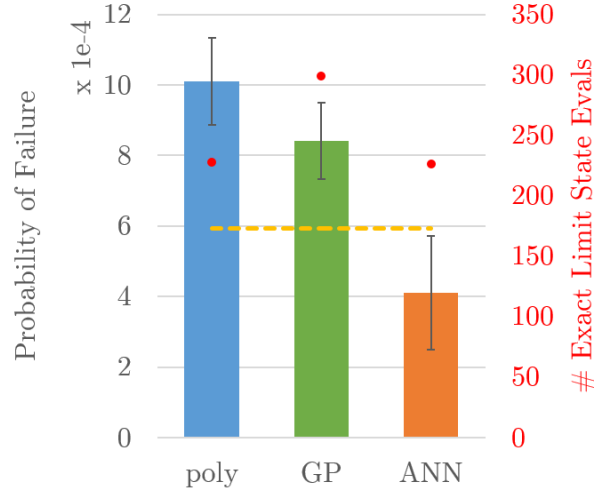


Fig. 34 Comparison of the estimated probability of failure of the nozzle structure via the adaptive directional importance sampling (ADIS) method for various response surface types. Quadratic polynomial (poly), Gaussian process (GP) and multilayer perceptron (ANN) response surfaces are compared. The exact probability of failure is shown as the dashed horizontal line and the number of model simulations required to estimate the probability of failure are shown as red dots. Error bars denote one standard deviation.

problems. Previous successes included the solution of a reduced and relaxed reliable nozzle design problem using the low- and medium-fidelity 2D-axisymmetric nozzle models. One important limitation was the incorporation of large failure probabilities in the reliability problem (10^{-2}) which arose due to the use of polynomial chaos surrogate models for the estimation of tail probabilities.

To address the above limitation we have been extending an adaptive directional importance sampling (ADIS) method as mentioned in the previous report. In this method, random directions are sampled in standard normal space and then searched along until the limit state surface is found (*i.e.* the point at which failure first occurs). A response surface is built when enough exact data has been obtained and is adaptively updated and used to provide an estimate of the failure probability. We have shown that this technique allows accurate estimation of probabilities as low as 5×10^{-5} with as little as 50 exact simulations in the case of the low-fidelity model quantities of interest, and requires several hundred simulations for similar probabilities in the case of the medium-fidelity nozzle quantities of interest (see figure 34). We have begun comparisons with other state-of-the-art reliability estimation techniques such as EGRA (Efficient Global Reliability Analysis), and are in the process of incorporating the method into the reliable optimization process.

Over the past quarter we have also started extending the solution of the reliable nozzle design problem to the 3D nozzle parameterization and model. Initial results show this problem to be more expensive in terms of number of iterations. So far, a reduced reliable optimization problem has been solved using the low-fidelity model: minimum weight subject to one each of thrust, stress, and temperature constraints. The stress and temperature constraints correspond to the most critical material layers in the nozzle wall. Figure 35 compares the optimal reliable non-axisymmetric nozzle on the right with its equivalent axisymmetric nozzle shape on the left. The optimization is performed using the low-fidelity axisymmetric nozzle model but uses the 3D nozzle parameterization that defines the non-axisymmetric nozzle model. The axisymmetric and non-axisymmetric nozzle models are linked by an equivalent cross-sectional area. Next quarter we will be exploring optimization results for higher-fidelity models. Due to the relatively high correlation between QoI for the axisymmetric and non-axisymmetric nozzle models, we anticipate using the axisymmetric nozzle in a multifidelity framework can substantially reduce computational cost.

Finally, we have been investigating an industry standard technique for DUU known as the basis value technique. In this technique, which is typically used to account for uncertainties in material properties, a one-sided tolerance interval known as a basis value is defined and used to replace random parameters in the optimization problem. A basis value is effectively a mean estimate of the random parameter knocked down by its estimated variance scaled by an appropriate factor. Optimizations performed using the basis value technique will provide a good comparison for the reliable and robust design techniques we have developed in the SEQUOIA project. It is also more sophisticated than the solution of

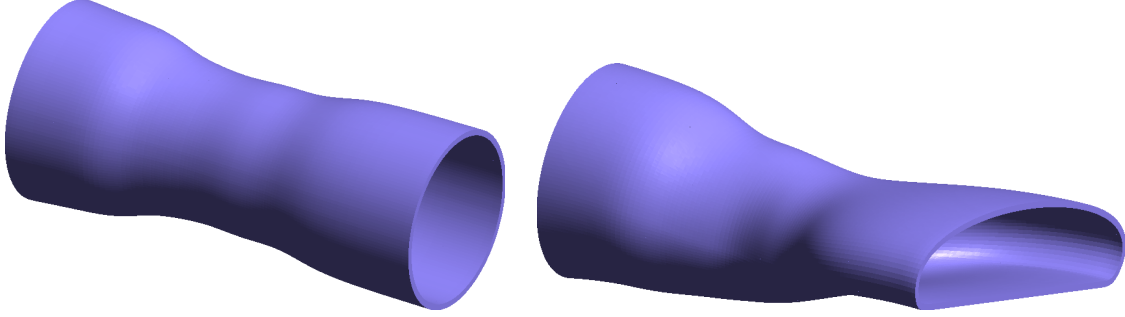


Fig. 35 Comparison of axisymmetric nozzle geometry (left) which is used during multiphysics analyses to determine an optimal reliable non-axisymmetric nozzle geometry (right).

an equivalent deterministic optimization problem.

C. Multifidelity OUU

At Sandia, the emphasis has been on multilevel-multifidelity optimization algorithms that move beyond common bi-fidelity approaches to exploit deep hierarchies of model form and spatial discretization. We have developed trust-region model management (TRMM) and multigrid optimization (MG/Opt) approaches that target both model form and discretization level hierarchies. We have deployed these methods extensively to model problems (solution matching for steady state diffusion and shape design for a transonic airfoil) and have performed robust DUU studies for the UCAV nozzle design. Figure 36 shows this robust nozzle result.

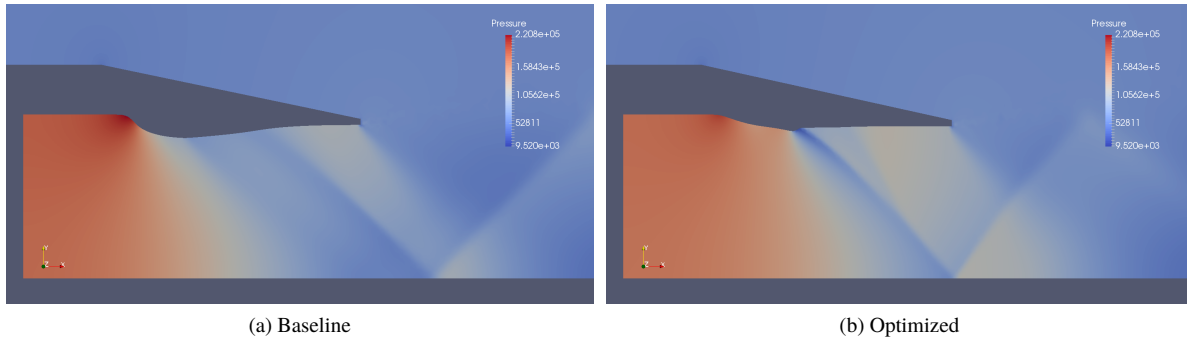


Fig. 36 Nozzle geometry and pressure contours resulting from the robust DUU problem, a) baseline configuration b) design minimizing variance in thrust subject to preserving reference weight and mechanical/thermal stresses.

In [40], we formulated several multilevel-multifidelity approaches and presented initial convergence results for trust-region model management (TRMM) applied recursively to a deep model hierarchy and for multigrid optimization (MG/Opt) applied recursively to model forms, discretization levels, or both. A particular focus has been on second-order recursive TRMM. In the case of full or finite-difference Newton, a full Newton solver can be used with the latest instantaneous Hessian data (low fidelity Hessian plus Hessian correction), and the TRMM process then inherits the quadratic convergence properties of full-Newton optimization solvers. Quasi-Newton approaches, on the other hand, provide extra challenges for multi-fidelity TRMM. Since quasi-Newton Hessian approximations must accumulate low rank updates over a series of design cycles, a number of iterations (increasing with design dimension) is required before one can obtain an accurate full-rank representation of the Hessian matrix. Multifidelity TRMM performs sub-problem optimization solves within a sequence of trust regions, and can incur repeated spin-up costs if their Hessian accumulations are restarted from scratch each time. Thus, in order for quasi-2nd-order TRMM to best compete with its single-fidelity peers, warm starting of accumulated Hessian data is required across the trust region cycles. This requires either some level of optimizer intrusion to restore internal Hessian states or access to a generic full Newton

solver for which Hessian accumulations can be managed externally from the solver. Leveraging a full Newton solver for external Hessian preservation, Table 13 shows convergence results for the transonic airfoil model problem, including quasi-second-order TRMM using BFGS updates. Figure 37 shows the baseline and optimal airfoil shapes.

	Coarse eval	Medium eval	Fine eval	C_D	C_L
Reference NACA 0012	–	–	–	0.10345	0.80118
Single-fidelity SLP	–	–	1190	0.065024	0.80199
Single-fidelity SQP	–	–	806	0.064904	0.80118
Single-fidelity NIP	–	–	2294	0.066894	0.80119
Three-level 1 st -order TRMM	43630	4882	187	0.064968	0.80153
Three-level quasi-2 nd -order TRMM	17020	248	93	0.068695	0.81966

Table 13 Single- and multi-fidelity optimization results for transonic airfoil design, minimizing drag subject to preserving reference lift using SU2 Euler CFD. SQP, NIP, and quasi-2nd-order TRMM all accumulate Hessian approximations based on BFGS updates.

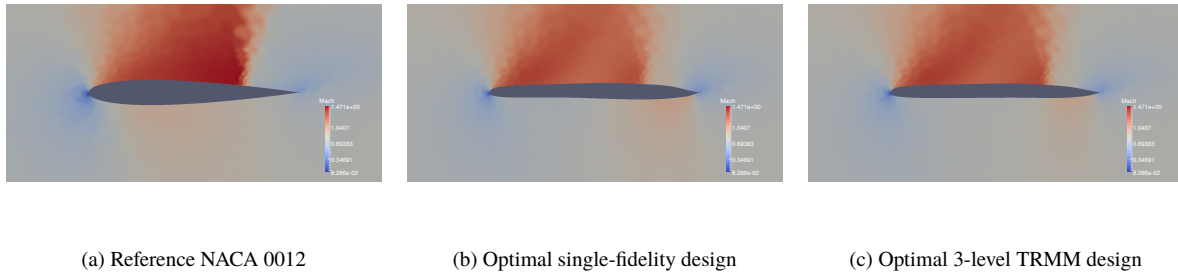


Fig. 37 Optimized shapes for transonic airfoil design using Euler CFD.

It is evident that the single-fidelity approaches require somewhere between 800 and 2300 fine-grid evaluations, whereas the multifidelity TRMM approaches reduce this by approximately an order of magnitude, with greater reliance on less-expensive coarse- and medium-grid evaluations. These results point to the utility of restricting optimizer interfacing to only the lowest-fidelity, coarsest-grid models, and relying on trust region recursion for higher fidelity verification and correction.

1. Sequential approaches: highlights

In this section, we describe a sequential approach for reliability-based design optimization (RBDO) that generalizes existing sequential methods [8]. It uses a new kriging-based reliability analysis algorithm called AKIRA (Adaptive Kriging Inverse Reliability Analysis) to estimate chance constraints and their gradients efficiently using high-fidelity information. Anchored decomposition is used to separate a RBDO problem into a series of reliability analyses and deterministic optimizations, thus naturally leading to a sequential method for solving the RBDO problem. Low-fidelity models are used in the limit state approximations and are managed with a trust region model management (TRMM) framework. The developed sequential RBDO method has some notable differences from many other methods:

Multifidelity optimization. A lower-fidelity model is used during the solution of the subproblem in the TRMM framework which potentially decreases the overall cost of the RBDO problem.

Trust Region Model Management. A trust region method is used to ensure convergence to an optimum of the high-fidelity model.

Anchored decomposition. The use of anchored decomposition motivates quantile constraint approximations for points other than the MPP. For example taking the mean as an anchor may have benefits for problems with coupled constraints and is helpful for system RBDO problems.

System inverse reliability analysis. A system RBDO problem is solved using the more robust inverse reliability analysis.

Adaptive Kriging Inverse Reliability Analysis The Adaptive Kriging Inverse Reliability Analysis (AKIRA) builds upon the work of existing kriging-based methods which solve the forward reliability analysis problem [41–46]. In a kriging-based method for reliability analysis kriging response surfaces are constructed for each limit state and a probability of failure is estimated from the kriging models, typically via a sampling method, which bypasses the need for many costly evaluations of the original limit state functions. Kriging response surfaces are typically chosen since they provide an empirical error estimate, the kriging variance, which can be used as a measure of uncertainty for the response surface prediction of the underlying limit state. As a result, the use of adaptive refinement methods are well-suited for kriging response surfaces where the approximate kriging model is refined only in regions impacting the probability of failure. An additional benefit of the kriging response surface is its exact interpolation property for sampled data points unlike other regression approaches.

The AKIRA algorithm solves the system inverse reliability analysis problem which has been shown to be more stable than the forward reliability analysis problem for RBDO [47]. In doing so, a response level $z = g(x, y)$ is calculated for a prescribed probability of failure $\bar{p}_f$ and limit state $g(x, y)$ with deterministic parameters x and random parameters y . In AKIRA, the system inverse reliability analysis problem is solved as a constrained least squares problem. The associated quantile constraint used in RBDO is then

$$z(x; y) > \bar{z}, \quad (30)$$

where $\bar{z}$ is a lower limit on the limit state level $z(x; y)$.

Figure 38 shows the convergence of AKIRA on the supersonic nozzle problem for three limit states, thrust, inner load layer temperature ratio, and the thermal layer structural failure criterion [8]. 14 random parameters are considered with a prescribed probability of failure $\bar{p}_f = 10^{-6}$. Only 95 total function evaluations are required to accurately model the limit states in the regions contributing the most to the probability of failure and thereby converge the limit state levels to sufficient accuracy for RBDO in the supersonic nozzle problem.

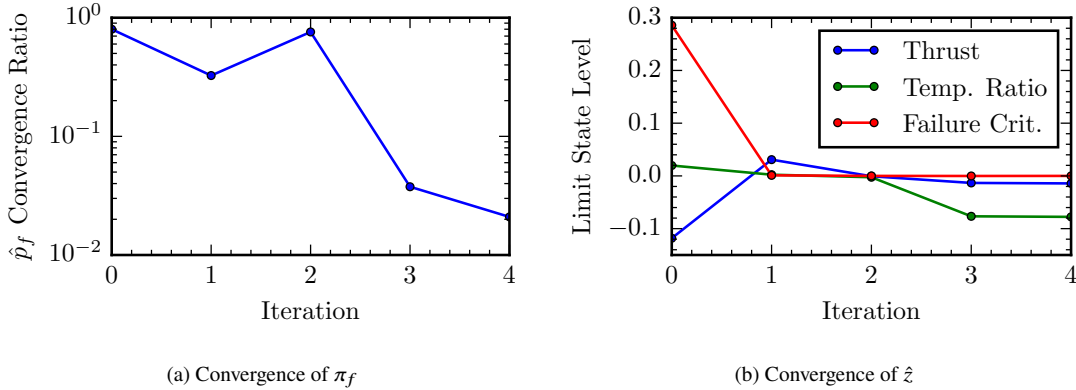


Fig. 38 Convergence of AKIRA for the supersonic nozzle problem with three normalized limit states: thrust F , inner load layer temperature ratio T , and thermal layer structural failure criterion S [8]. Convergence occurs after 4 iterations and 95 total function evaluations for each limit state. The convergence ratio π_f for $\hat{p}_f$ illustrates the convergence of AKIRA’s kriging model in the regions important to the probability of failure calculation (values < 0.1 correspond to high accuracy). Convergence of the estimated limit state levels $\hat{z}$ for each constraint is shown on the right.

Anchored Decomposition Approximation Any function approximation separable in the deterministic design variables and random parameters may be used to achieve a sequential approach for RBDO. The ANOVA decomposition [48] is one such approximation and has been extensively studied, however obtaining any terms greater than first-order can be

expensive for high-dimensional problems. Here we choose the anchored decomposition [49] which is expressed as

$$f(x) = f(x_1, \dots, x_n) = f_0 + \sum_{i=1}^n f_i(x_i) + \sum_{1 \leq i \leq j}^n f_{ij}(x_i, x_j) + \dots + f_{12\dots n}(x_1, \dots, x_n). \quad (31)$$

where $f^{i_1 i_2 \dots i_s}$ is the function $f(x)$ with all input variables x_i whose $i \notin \{i_1 i_2 \dots i_s\}$ are fixed at the anchor a_i , $1 \leq s \leq n$.

The truncated anchored decomposition is used as an approximation of $f(x)$ and is naturally suited for most optimization algorithms where $f(x)$ is evaluated sequentially at a series of points. In such cases, the anchor is selected to be the previous point at which information about $f(x)$ has been obtained and only a locally accurate approximation of $f(x)$ is required. Incorporation of other previously evaluated points can also be used to build in curvature information. Application to sequential RBDO yields the approximation

$$f(x, y) \approx f(x_c, y) + f(x, y_c) - f(x_c, y_c), \quad (32)$$

where the anchor $a = (x_c, y_c)$ for deterministic x and random y .

Multifidelity Adaptation Adaptation of the anchored decomposition approximation in (32) simply follows by replacing the terms in the approximation with lower-fidelity models $f_{lo}(x, y)$ and then relying on a trust region model management scheme to ensure convergence to an optimum of the high-fidelity model $f_{hi}(x, y)$. The anchor for the deterministic design variables x is taken to be the trust region center x_c . Choices for the anchor of the random parameters y such as $E[y]$ or the most probable point of failure were both investigated; for problems with multiple constraints it is advantageous to use the mean of the random parameters as the anchor as this typically results in fewer function evaluations [8].

The quantile constraint from (30) obtained from an inverse reliability analysis can be reformulated using the anchored decomposition from (32) and the multifidelity adaptation described above [8]. Additionally incorporating an additive first order correction to ensure provable convergence in the trust region scheme leads to the formulation of an approximate multifidelity quantile constraint

$$g_{lo}(x, y_c) - s - \nabla_x s^\top (x - x_c) \geq \bar{z}, \quad (33)$$

where s plays the role of an additive safety margin and is given by

$$s \equiv g_{lo}(x_c, y_c) - z_{hi}(x_c), \quad (34a)$$

$$\nabla_x s \equiv \nabla_x g_{lo}(x_c, y_c) - \nabla_x z_{hi}(x_c). \quad (34b)$$

Equation (33) is an approximate deterministic constraint whose parameters are updated each trust region using a trust region model management procedure. When the anchor is taken as the most probable point of failure (MPP), the sequential RBDO method is essentially the Sequential Optimization and Reliability Assessment (SORA) proposed by Du and Chen [50]. Thus it may be called a SORA method. When the design variables are deterministic and do not parameterize the random variables, the anchor is taken as the MPP, and an additive correction is used for the limit state approximations then the sequential RBDO method is also similar to Wu et al's proposed method with a shift in the limit state [51]. However the described method generalizes both approaches in that it accommodates a variety of anchor points other than the MPP.

2. Numerical Results

We first present a comparison of the described sequential RBDO algorithm with other similar algorithms for a standard short column design problem, followed by results for the supersonic nozzle application. Additional examples can be found in [8].

Short Column Problem The short column problem [42, 52–56] concerns a short column with rectangular cross-section subject to two bending moments M_1 and M_2 and an axial load P . The deterministic design variables are the means of the breadth b and height h of the column cross-section; thus the design variables $x = [\mu_b, \mu_h]^\top$. The loading and the column's yield strength are considered to be independent normal random variables. The objective is to minimize the column mass such that failure by yielding does not occur with probability greater than 1.35×10^{-3} .

Table 14 compares the RBDO results for the short column problem. The proposed sequential method performs comparably to SORA and is more efficient than double-loop RBDO with FORM even when the mean is used as the anchor point. When an improved low-fidelity correction (*e.g.* two-point 2nd order correction) is used in the trust region method, using the mean as an anchor can become competitive with SORA and using the MPP as an anchor can exceed the performance of SORA in terms of number of function evaluations.

Table 14 Comparison of RBDO results for the short column problem. The problem is first solved without geometric uncertainties, that is with deterministic $\mu_b = b$ and $\mu_h = h$, and then with geometric uncertainties. Several items of interest are compared in this table: 1) the use of the mean or the MPP as the anchor point for the separable approximation of the limit state, 2) the use of additive 1st order or two-point 2nd order corrections for the limit state approximations, and 3) the results from the proposed method with results from a similar sequential RBDO method SORA and inverse FORM with HNV (a popular double-loop RBDO method) [55]. $f(x^*)$ is the optimal value of the objective function, μ_b^* and μ_h^* are the optimal mean breadth and height of the column's cross section, and N is the number of function evaluations required during the optimization.

Method	$f(x^*)$	μ_b^*	μ_h^*	N
Without geometric uncertainties				
MPP anchor; additive 1st order correction	0.1918	0.3097	0.6194	115
Mean anchor; additive 1st order correction	0.1918	0.3114	0.6160	270
MPP anchor; two-point 2nd order correction	0.1918	0.3097	0.6194	115
Mean anchor; two-point 2nd order correction	0.1918	0.3097	0.6194	248
SORA [55]	0.1915	–	–	119
Inverse FORM w/ HNV [55]	0.1913	–	–	450
With geometric uncertainties				
MPP anchor; two-point 2nd order correction	0.2363	0.3606	0.6554	220
Mean anchor; two-point 2nd order correction	0.2357	0.3459	0.6813	459
SORA [55]	0.2372	–	–	252
Inverse FORM w/ HNV [55]	0.2372	–	–	570

Supersonic Nozzle Problem The full supersonic nozzle RBDO problem contains 40 random parameters related to material properties, inlet and atmospheric conditions, and heat transfer properties as well as 96 deterministic design variables parameterizing the wall geometry and thicknesses. Furthermore 12 chance constraints ensure failure by excessive temperatures or stress or low thrust production do not occur. A global sensitivity study however reveals that only 14 of the 40 random parameters have a non-negligible contribution to the critical quantities of interest, namely the mass, thrust, temperature in the inner load layer, and strain in the thermal layer. The total order Sobol' indices for the 3D Euler nozzle model and FEM thermal and structural models are shown in Figure 39 for the four critical quantities of interest and the random parameters inducing the greatest variability.

The results of the global sensitivity study motivated the study of a reduced optimization problem with three constraints, 14 random parameters, and 54 deterministic design variables. The reduced quantile constrained design problem is

$$\begin{aligned}
& \underset{x \in S^d}{\text{minimize}} && E[M(x, y)] \\
& \text{s.t.} && \frac{z_F(x; \bar{p}_f)}{21,500} \geq 1 \\
& && z_T(x; \bar{p}_f) \leq 1 \\
& && z_S(x; \bar{p}_f) \leq 1 \\
& && Ax \leq b,
\end{aligned} \tag{35}$$

where M is the mass of the nozzle, z_F is the limit state level for the thrust F , z_T is the limit state level for the temperature

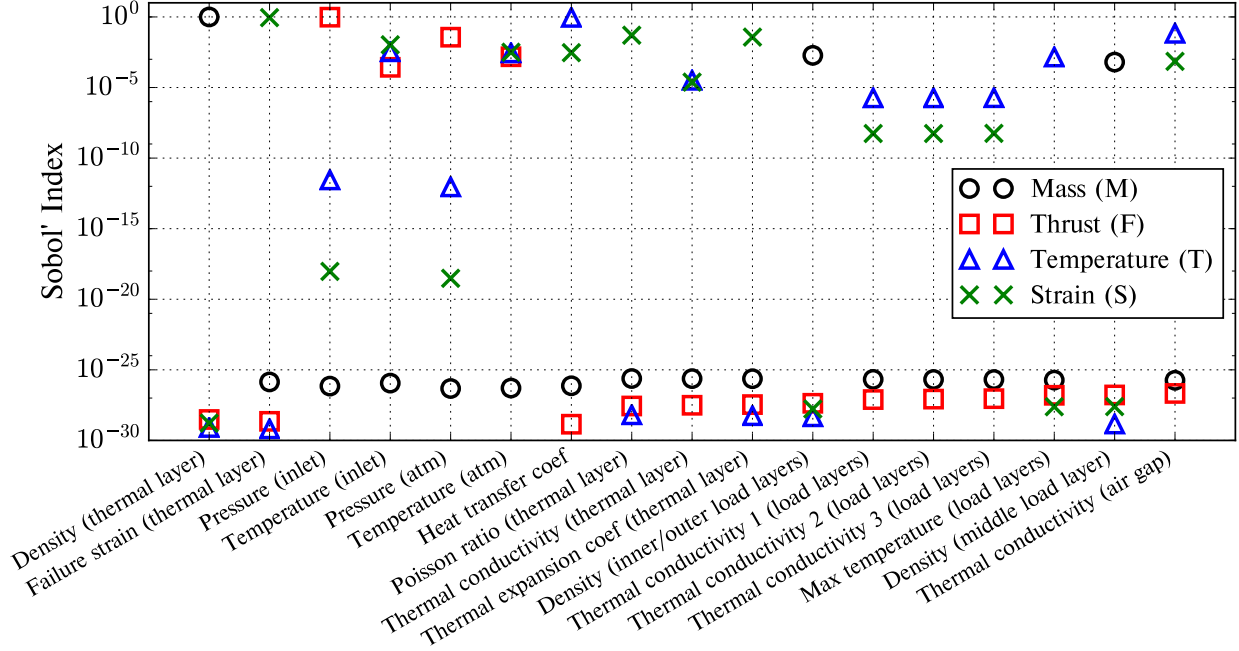


Fig. 39 Total order Sobol' indices for the 17 random variables which contribute most to the variance of the four quantities of interest considered in the nozzle problem. Eliminating variables with indices less than 10^{-5} reduces the problem's random dimension from 40 to 14. First order Sobol' indices (not shown) are nearly the same implying interaction effects can be neglected. The Sobol' indices are calculated using a sparse grid level 1 polynomial chaos expansion with three anisotropic adaptive refinement iterations, resulting in a total of 3,405 function evaluations for the 3D Euler nozzle model and FEM thermal and structural models.

ratio T in the inner load layer, and z_S is the limit state level for the strain-based failure criterion S in the thermal layer. $Ax \leq b$ is a set of linear constraints to ensure reasonable geometries and $\bar{p}_f$ is the prescribed system probability of failure, $\bar{p}_f = 10^{-6}$.

Results for the problem in (35) are compared with the solution of a similar standard practice deterministic design problem formulated as

$$\begin{aligned}
 & \underset{x \in \mathbb{R}^{54}}{\text{minimize}} && M(x, \tilde{y}) \\
 & \text{s.t.} && \frac{F(x, \tilde{y})}{21,500\gamma} \geq 1 \\
 & && \gamma T(x, \tilde{y}) \leq 1 \\
 & && \gamma S(x, \tilde{y}) \leq 1 \\
 & && Ax \leq b,
 \end{aligned} \tag{36}$$

where $\tilde{y}$ are the fixed B basis values (a set of conservative values) for the random parameters y and γ is a safety factor varied between 1.2 and 1.5.

Table 15 summarizes and compares the solutions of the standard practice deterministic design problem from (36) and the reduced reliability-based design optimization problem from (35) for various model fidelities. Several trends are immediately clear:

- 1) The standard practice deterministic design using safety factors and basis values does not achieve the desired reliability. In some cases, *e.g.* the low-fidelity model, the reliability is lower for the deterministic design, however in other cases, *e.g.* the high-fidelity model, the reliability is much greater than required. Basis values and safety factors are unable to directly control system reliability.
- 2) The computational cost of the RBDO problem is an order of magnitude greater than the deterministic problem. The additional cost is due to the need to perform multiple deterministic optimizations interspersed with reliability

Table 15 Comparison of standard practice deterministic and reliability-based design optimization (RBDO) results for the nozzle problem. The deterministic optimization results use conservative material properties (B basis values) and safety factors to account for uncertainty. The RBDO results directly account for uncertainty and use the multifidelity sequential method presented in this section. In the table, “Fidelity” refers to the highest fidelity model used in the optimization; for example “Low” refers to a single-fidelity optimization using the low-fidelity nozzle models, whereas “High” refers to multifidelity optimizations using the high-fidelity model. The total number of function evaluations N_{eval} and gradient evaluations N_{grad} are shown. In multifidelity optimizations the number corresponding to the low-fidelity model is shown on the left, and the number corresponding to the high-fidelity model is shown on the right separated by “/”. The optimization objective is expected mass, and $\hat{p}_f$ is the nozzle system failure probability estimated using AKIRA or a kriging-based forward reliability analysis. Finally, the active constraints at the optimum are listed: F refers to thrust, T refers to the inner load layer temperature ratio, and S refers to the thermal layer structural failure criterion.

Optimization	Fidelity	N_{eval}	N_{grad}	Objective	$\hat{p}_f$	Critical Constraints
Det.	Low	51	49	329.16	2.58×10^{-5}	F, T
RBDO	Low	386	26	341.73	10^{-5}	F, T
RBDO*	Low	916	70	1,033.11	1.19×10^{-6}	F, T
Det. ($\gamma = 1.5$)	High	8	7	384.17	$< 2.31 \times 10^{-8}$	F
Det. ($\gamma = 1.2$)	High	5	4	380.18	$< 2.31 \times 10^{-8}$	F
RBDO	High	8 / 317	6 / 7	297.51	10^{-6}	F, T

slightly infeasible for $\bar{p}_f = 10^{-6}$

- analyses to solve the RBDO problem.
- 3) The number of high-fidelity model evaluations exceeds the number of low-fidelity model evaluations. In contrast to most multifidelity methods, the low-fidelity model does not appear to substantially decrease the number of high-fidelity model evaluations. The reason is that reliability analyses must be performed using the high-fidelity model to ensure convergence to an optimum for the high-fidelity model, and such reliability analyses constitute the majority of the cost of solving the RBDO problem. In addition the subproblem optimizations for the nozzle problem using the low-fidelity model are usually easily solved within the trust region bounds.
 - 4) The critical constraints in the nozzle problem are the thrust and temperature in the inner load layer. The structural failure criteria in the thermal layer is not an active constraint due to the high performance of the ceramic matrix composite.

The results for the high-fidelity model show the potential for excessive conservatism with well-intentioned deterministic design procedures. Table 15 shows the deterministic design using a safety factor $\gamma = 1.5$ and B basis values has an expected mass of 384 kg with a system failure probability at least two orders of magnitude smaller than the desired reliability. On the other hand, the reliable design features 22.6% reduction in mass while still meeting the requested reliability $\bar{p}_f = 10^{-6}$. Using a smaller safety factor, $\gamma = 1.2$, which may be more appropriate for military aircraft, does little to ameliorate the overdesign of the deterministic optimization. Figure 41 qualitatively compares the optimal standard practice deterministic design ($\gamma = 1.2$) with the optimal reliable design nozzle geometry and Figure 40 compares the optimal shapes of the inner nozzle walls. The obvious larger diameter of the standard practice design is reflected by its increased mass compared to the reliable design. Finally, Figure 41 gives a qualitative comparison of the optimal nozzle geometries and flow at the most probable point of failure for the standard practice and reliable designs. Additional analysis of results can be found in [8].

D. Model Form OUU

The objective of this task is to demonstrate the use of embedded model form UQ techniques in robust design. Specifically, aerodynamic design is pursued with a consideration of uncertainties due to turbulence modeling. OUU is performed based on RANS simulations aircraft engine nozzle, optimizing for robustness in the presence of model inadequacy. Model inadequacy is introduced into the simulations via the data-driven embedded UQ procedure detailed in TA2.

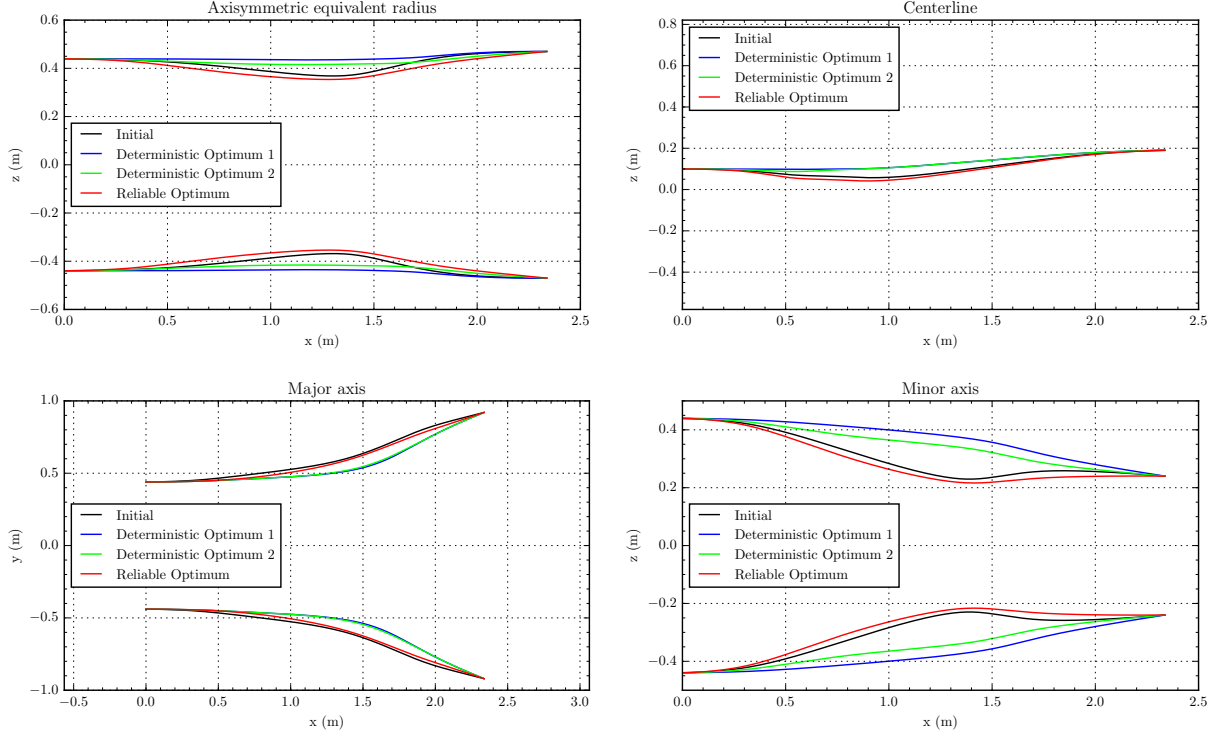


Fig. 40 Comparison of optimal inner wall geometry for the high-fidelity model.

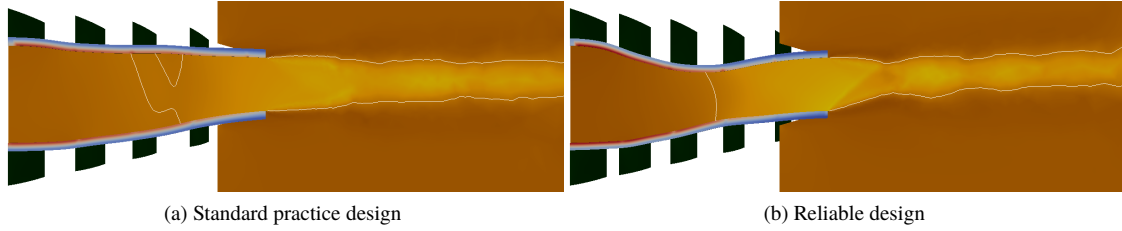


Fig. 41 Qualitative comparison of optimal nozzle geometry for the high-fidelity model and its effect on temperature and Mach number. The coupled aerothermostructural analysis is performed for the optimal nozzle geometry at the most probable point (MPP) of failure for the 14 random variables, *i.e.* for the combination of random variables that is most likely to yield system failure. The nozzles are shown split down their centerline with flow from left to right. Temperatures in the nozzle wall are shown (blue = cooler, red = hotter), as well as the Mach number (darker = smaller, lighter = larger). The color scales for both nozzles are the same, which enables a qualitative comparison between the optimal geometries. The sonic lines which mark where the Mach number is 1 in the fluid flow are shown as white contours. Since the inlet flow is subsonic, the flow after the sonic lines is supersonic. Note that a shock begins to form at the center of the nozzle behind the throat for the standard practice design in the off-design conditions at the MPP.

1. Problem Formulation

The goal is to optimize the parametrized shape of the nozzle for the non-afterburning F100-PW-220 turbofan engine which is described in prior sections. The known quantities are the inlet diameter of the nozzle along with Mach number of the aircraft, stagnation pressure and stagnation temperature (at the nozzle inlet) for various flight regimes.

Shape parametrization: The aerodynamic model uses the compressible Reynolds Averaged Navier–Stokes (RANS) as the governing equations with the Spalart–Allmaras turbulence model for closure.

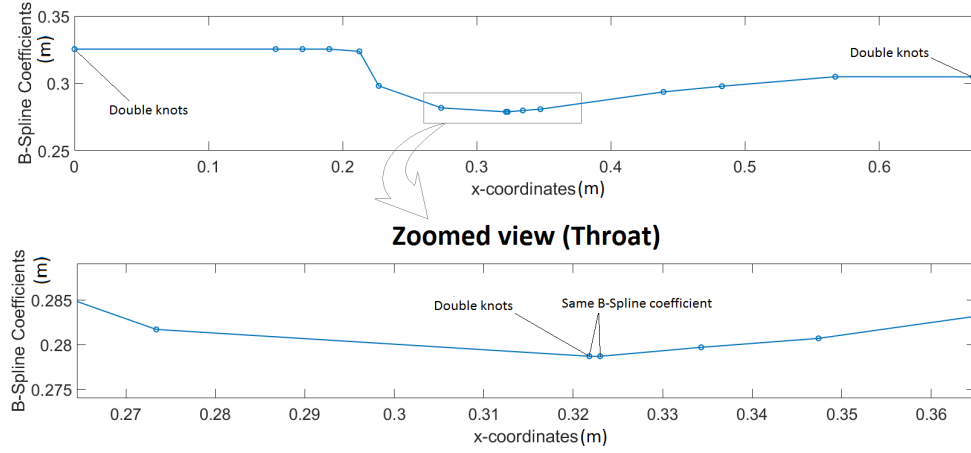


Fig. 42 B-spline control polygon

The inner nozzle surface is parametrized using 15 control points, which are used to generate the 2-D axisymmetric geometry as a cubic B-Spline curve. There are repeated knots at the control points defining the inlet, the exit and the throat to have desired geometrical characteristics at these locations. The first four control points are kept fixed to ensure a smooth geometry near the inlet and the remaining eleven control points are free to be varied to optimize the nozzle geometry. Since each control point consists of an x-coordinate and a B-spline coefficient (y-coordinate of the control point) defined along with it, a total of 21 design variables have been used (as the B-spline coefficients for the two narrowly-spaced control points defining the throat are equal, see figure 42).

Figure 42 shows the manner in which control points and knots have been setup for the control polygon of a sample nozzle geometry. Appropriate constraints have been used to constrain the control polygon to ensure that

- the variation in the coordinates of the control points is confined to $\pm 40\%$ of the initial value;
- the nozzle has proper convergence and divergence in the respective parts;
- the control point corresponding to the throat always has the minimum cross-section; and
- the gradients in the geometry are not exceedingly steep, especially at the throat.

Optimization problem The robust design optimization problem can be summarized as follows:

$$\underset{\mathbf{x}}{\text{minimize}} \quad \max_{\mathcal{M}^i} (\text{Thrust}(\mathbf{x}, \mathcal{M}^i)) - \min_{\mathcal{M}^i} (\text{Thrust}(\mathbf{x}, \mathcal{M}^i)),$$

subject to the non-linear constraints

$$\langle \text{Thrust}(\mathbf{x}, \mathcal{M}^i) \rangle > 21 \text{ kN},$$

$$\text{Mass}(\mathbf{x}) < 75 \text{ kg}.$$

For comparison, the deterministic optimization can be formulated as follows

$$\underset{\mathbf{x}}{\text{maximize}} \quad \text{Thrust}_{SA}(\mathbf{x})$$

subject to the non-linear constraint

$$\text{Mass}(\mathbf{x}) < 75 \text{ kg}$$

where $\mathbf{x}$ is the vector of design variables and $\mathcal{M}^i$ denotes the i^{th} model being used for predictive simulations (can be baseline SA or one of the Machine Learning augmented models). Linear geometric constraints [57] are used to restrict unrealistic shapes.

The thrust values required to estimate the discrepancy interval and the mean thrust were calculated using the baseline SA model and models augmented using FIML. A brief description of the model augmentation process is given as follows:

- 1) The SA model was augmented with a machine learned function $\beta(\eta(x))$, which is introduced as a multiplicative correction to the production term.
- 2) The Wilcox stress-omega model was used as a proxy to obtain higher fidelity data. The MAP value of beta at every node in the mesh is obtained through Bayesian inversion operating on skin friction data.
- 3) The locally non-dimensional flow features η , were calculated at each node. In this work, three features were taken which were functions of shear strain magnitude, wall shear stress at the nearest wall, and the distance to the nearest wall, the expressions for which are given in the previous section.
- 4) Using a suitable Machine Learning technique, viz. Adaboost, a model that maps the aforementioned features to β was created, which was then augmented to the base SA model.

The *uncertainty interval* for the problem is determined using the baseline SA model and six ML-augmented SA models (epistemic variables in this setting) which are used to run RANS simulations on the same geometry/grid resulting in seven different realizations of thrust values.

The optimization can be carried out using any feasible gradient-based or gradient-free algorithm. For this work, the gradient-based optimization code NPSOL[58] was used.

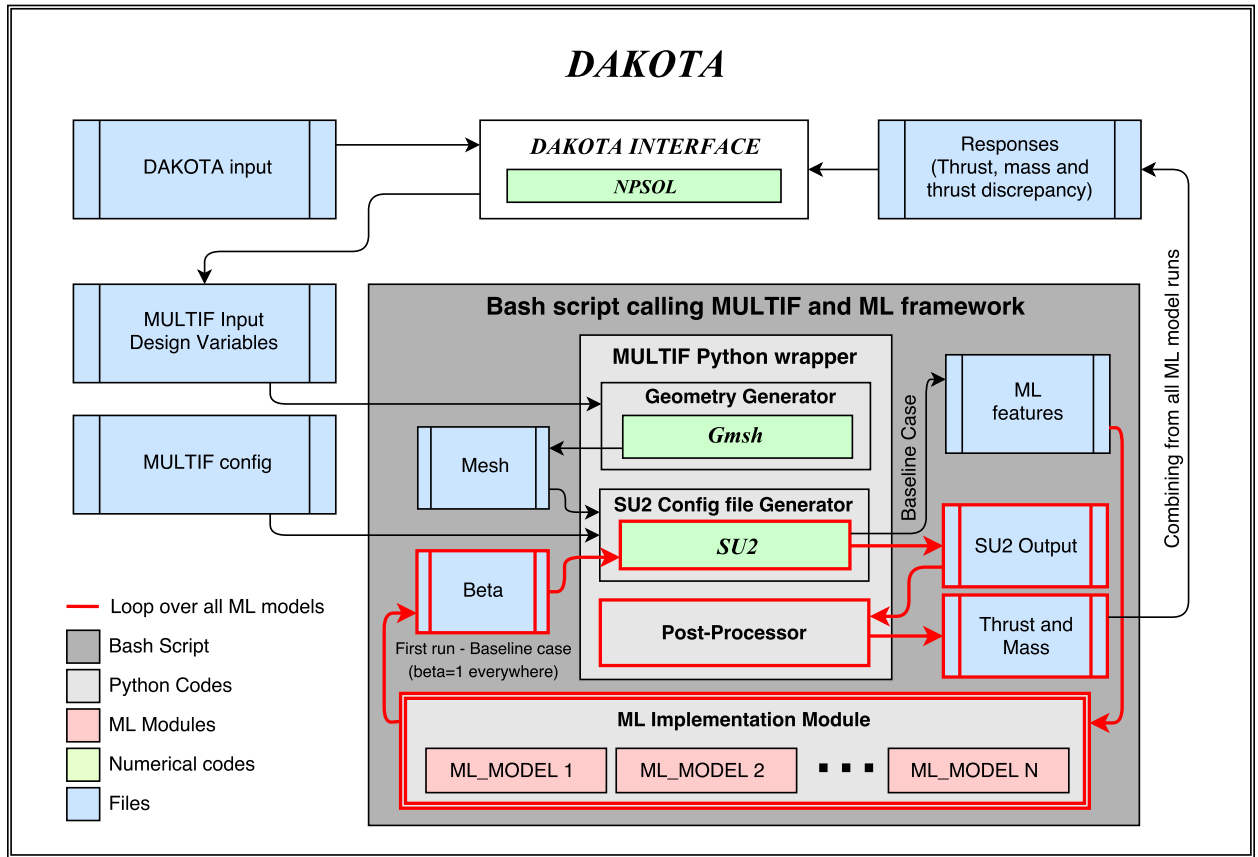


Fig. 43 Simplified flowchart for the DUU framework with Machine learning augmented models

The flowchart for the optimization process is shown in figure 43.

2. Results

Using NPSOL within DAKOTA, a deterministic optimization to maximize thrust, and an interval-based DUU optimization to minimize model uncertainty were performed. Model-form uncertainty in the thrust, as mentioned earlier, is defined as an interval that consists of the base SA model and six other Machine-Learning augmented SA models. Mass is calculated based on the volume obtained from the parametrized geometry and material densities for the structural layers as defined for the nozzle. The considerable difference between the deterministic optimization and DUU runs is evident in the nozzle shapes produced by the two geometries as shown in figure 44. The optimum thrust

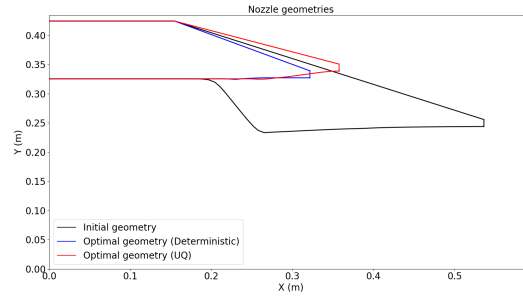


Fig. 44 Comparison between baseline and optimum geometries

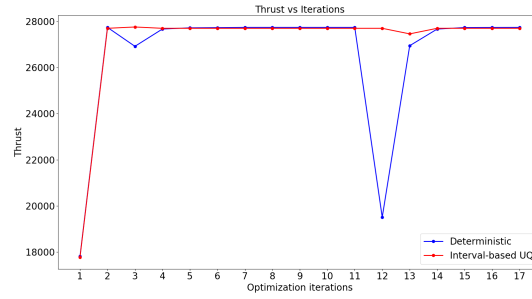


Fig. 45 Convergence of objective

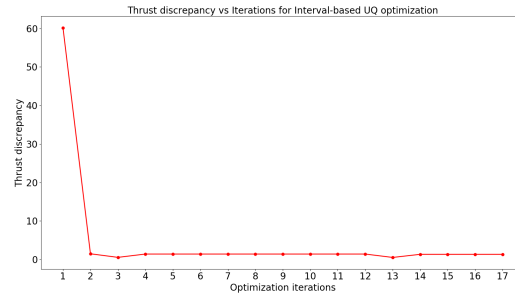


Fig. 46 Thrust discrepancy

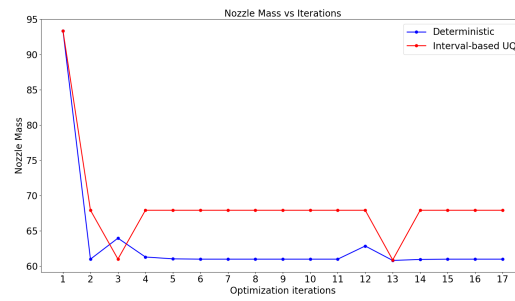


Fig. 47 Mass vs iterations

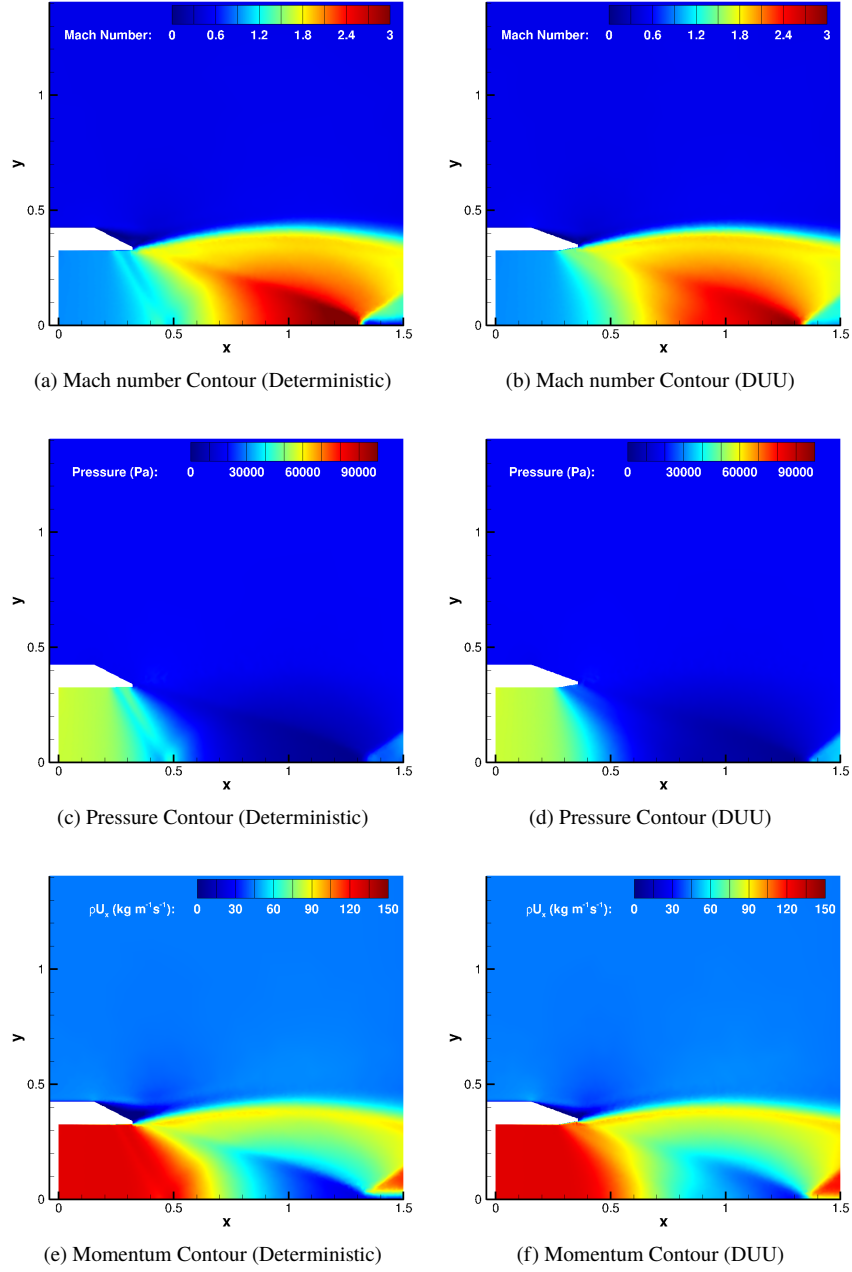


Fig. 48 Optimum Flow Contours (deterministic and DUU)

values for the deterministic (27738.2 N) and DUU (27699.2 N) runs are similar as seen in figure 45, and the two optimal geometries (figure 44) exhibit noticeable difference in the flow path towards the exit plane of the nozzle. The thrust discrepancy decays from a considerable value of around 60 N to a value of about 1 N for the optimal case, as seen in figure 46.

It should be noted that the discrepancy being calculated here corresponds to the ML models trained using 6 representative configurations. It is well-recognized that more comprehensive datasets could lead to different and more trustworthy DUU solutions. The optimal mass, as shown in figure 47, for the deterministic optimization case was 61 kg while it was 67.93 kg for the DUU case.

The flow contours corresponding to the optimal deterministic and DUU nozzle geometries are shown in figure 48.

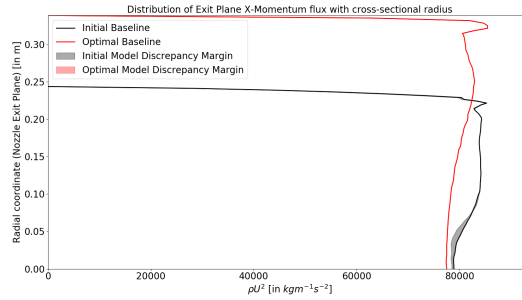


Fig. 49 Exit plane Momentum flux profile for initial and optimum cases

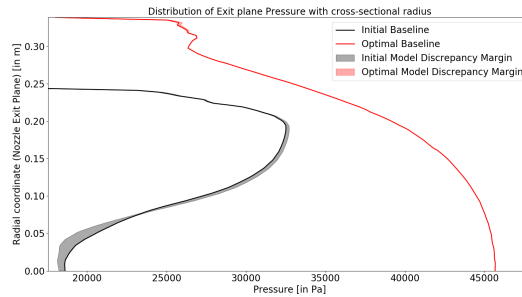


Fig. 50 Exit plane Pressure profile for initial and optimum cases

We see that the pressure increase at the exit plane is significantly higher than the previous case (figure 50), although the streamwise momentum flux is reduced as seen in figure 49, but the overall effect still remains a significant increase in thrust.

Acknowledgments

Support for this research was provided by the DARPA program EQUiPS. Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. The views expressed in the article do not necessarily represent the views of the U.S. Department Of Energy or the United States Government.

References

- [1] Alonso, J., Eldred, M., Constantine, P., Duraisamy, K., Farhat, C., Iaccarino, G., and Jakeman, J., "Scalable Environment for Quantification of Uncertainty and Optimization in Industrial Applications (SEQUOIA)," *19th AIAA Non-Deterministic Approaches Conference*, Grapevine, Texas, 2017.
- [2] Fenrich, R., "F100-PW-220 Engine Model," private communication, 2015.
- [3] Fenrich, R., "F100-PW-220 Non-Ideal Nozzle Model," private communication, 2016.
- [4] Fenrich, R., "Axisymmetric Nozzle Optimization Formulationsl," private communication, 2016.
- [5] Economou, T. D., Palacios, F., Copeland, S., Lukaczyk, T., and Alonso, J. J., "SU2: An Open-Source Suite for Multiphysics Simulation and Design," *AIAA Journal*, Vol. 54, No. 3, 2016, pp. 828–846.
- [6] Farhat, C., "A general-purpose finite element structural analyzer," <https://bitbucket.org/frg/aero-s>, 2018.
- [7] Fenrich, R. W., Menier, V., Avery, P., and Alonso, J. J., "Reliability-based design optimization of a supersonic nozzle," *Proceedings of the 18th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Denver, Colorado, 2017.

- [8] Fenrich, R. W., “Efficient Sequential Reliability-Based Design Optimization with Adaptive Kriging Inverse Reliability Analysis,” Ph.D. thesis, Stanford University, Stanford, California, 2018.
- [9] Geraci, G., Eldred, M. S., and Iaccarino, G., “A multifidelity multilevel Monte Carlo method for uncertainty propagation in aerospace applications,” *19th AIAA Non-Deterministic Approaches Conference*, Grapevine, TX, 2017.
- [10] Fenrich, R. W., and Alonso, J. J., “Reliable multidisciplinary design of a supersonic nozzle using multifidelity surrogates,” *Proceedings of the 18th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Denver, Colorado, 2017.
- [11] Iaccarino, G., Mishra, A. A., and Ghili, S., “Eigenspace perturbations for uncertainty estimation of single-point turbulence closures,” *Physical Review Fluids*, Vol. 2, No. 2, 2017, p. 024605.
- [12] Mishra, A. A., and Iaccarino, G., “Uncertainty estimation for reynolds-averaged navier–stokes predictions of high-speed aircraft nozzle jets,” *AIAA Journal*, 2017, pp. 3999–4004.
- [13] Mishra, A. A., Mukhopadhyaya, J., Iaccarino, G., and Alonso, J., “Uncertainty Estimation Module for Turbulence Model Predictions in SU2,” *AIAA Journal*, Vol. 57, No. 3, 2018, pp. 1066–1077.
- [14] Kennedy, M. C., and O’Hagan, A., “Bayesian calibration of computer models,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 63, No. 3, 2001, pp. 425–464.
- [15] Sargsyan, K., Najm, H. N., and Ghanem, R., “On the statistical calibration of physical models,” *International Journal of Chemical Kinetics*, Vol. 47, No. 4, 2015, pp. 246–276.
- [16] Spalart, P., and Allmaras, S., “A one-equation turbulence model for aerodynamic flows,” *30th Aerospace Sciences Meeting and Exhibit Reno, NV, U.S.A.*, 1992.
- [17] Soize, C., and Farhat, C., “A nonparametric probabilistic approach for quantifying uncertainties in low-dimensional and high-dimensional nonlinear models,” *International Journal for Numerical Methods in Engineering*, Vol. 109, No. 6, 2017, pp. 837–888.
- [18] Farhat, C., Bos, A., Avery, P., and Soize, C., “Modeling and quantification of model-form uncertainties in eigenvalue computations using a stochastic reduced model,” *AIAA Journal*, Vol. 56, No. 3, 2017, pp. 1198–1210.
- [19] Farhat, C., Bos, A., Tezaur, R., Chapman, T., Avery, P., and Soize, C., “A stochastic projection-based hyperreduced order model for model-form uncertainties in vibration analysis,” *2018 AIAA Non-Deterministic Approaches Conference*, 2018, p. 1410.
- [20] Soize, C., and Farhat, C., “Probabilistic learning for modeling and quantifying model-form uncertainties in nonlinear computational mechanics,” *International Journal for Numerical Methods in Engineering*, Vol. 117, No. 7, 2019, pp. 819–843.
- [21] Giles, M. B., “Multilevel Monte Carlo methods,” *Acta Numerica*, Vol. 24, 2015, pp. 259–328.
- [22] Peherstorfer, B., Willcox, K., and Gunzburger, M., “Optimal model management for multifidelity Monte Carlo estimation,” *SIAM Journal on Scientific Computing*, Vol. 38, No. 5, 2016, pp. A3163–A3194.
- [23] Geraci, G., Eldred, M., and Iaccarino, G., “A multifidelity control variate approach for the multilevel Monte Carlo technique,” *Center for Turbulence Research Annual Research Briefs*, Center for Turbulence Research, Stanford University, 2015, pp. 169–181.
- [24] Gorodetsky, A., Geraci, G., Eldred, M., and Jakeman, J., “A Generalized Approximate Control Variate Framework for Multifidelity Uncertainty Quantification,” *arXiv preprint arXiv:1811.04988v3 [stat.CO]*. Submitted, 2019.
- [25] Lavenberg, S. S., Moeller, T. L., and Welch, P. D., “Statistical results on control variables with application to queueing network simulation,” *Operations Research*, Vol. 30, No. 1, 1982, pp. 182–202.
- [26] Giles, M. B., “Multilevel Monte Carlo path simulation,” *Operations Research*, Vol. 56, No. 3, 2008, pp. 607–617.
- [27] Pasupathy, R., Taaffe, M., Schmeiser, B. W., and Wang, W., “Control-variate estimation using estimated control means,” *IIE Transactions*, Vol. 44, No. 5, 2014, pp. 381–385.
- [28] Ng, L. W.-T., and Willcox, K., “Multifidelity approaches for optimization under uncertainty,” *Int. J. Numer. Meth. Engng.*, Vol. 100, No. 10, 2014, pp. 746–772.
- [29] Geraci, G., Eldred, M. S., Gorodetsky, A. A., and Jakeman, J. D., “LEVERAGING ACTIVE DIRECTIONS FOR EFFICIENT MULTIFIDELITY UNCERTAINTY QUANTIFICATION,” *7th European Conference on Computational Fluid Dynamics (ECFD 7)*, Glasgow, UK, 2018.

- [30] Constantine, P., “Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies,” *SIAM Spotlights*, 2015.
- [31] Eldred, M. S., Geraci, G., and Jakeman, J., “Multilevel Monte-Carlo Hybrids Exploiting Multifidelity Modeling and Sparse Polynomial Chaos,” *SIAM Conference on Uncertainty Quantification*, Lausanne, Switzerland, 2016.
- [32] Ng, L. W.-T., and Eldred, M., “Multifidelity uncertainty quantification using non-intrusive polynomial chaos and stochastic collocation,” *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*, 2012, p. 1852.
- [33] Eldred, M. S., Ng, L. W. T., Barone, M. F., and Domino, S. P., “Multifidelity Uncertainty Quantification Using Spectral Stochastic Discrepancy Models,” *Handbook of Uncertainty Quantification*, Springer International Publishing, Cham, 2015, pp. 1–45. doi:10.1007/978-3-319-11259-6_25-1, URL http://link.springer.com/10.1007/978-3-319-11259-6_25-1.
- [34] Eldred, M. S., Geraci, G., Gorodetsky, A., and Jakeman, J., “Multilevel-Multifidelity Approaches for Forward UQ in the DARPA SEQUOIA project,” *2018 AIAA Non-Deterministic Approaches Conference, AIAA SciTech Forum*, Kissimmee, Florida, 2018.
- [35] Jakeman, J. D., Eldred, M. S., and Sargsyan, K., “Enhancing ℓ_1 -minimization estimates of polynomial chaos expansions using basis selection,” *Journal of Computational Physics*, Vol. 289(C), 2015, pp. 18–34.
- [36] A. Haji-Ali, L. T., F. Nobile, and Tempone, R., “Multi-index stochastic collocation for random pdes,” *Computer Methods in Applied Mechanics and Engineering*, Vol. 306, 2016, pp. 95 – 122.
- [37] Jakeman, J., Eldred, M., Geraci, G., and Gorodetsky, A., “Adaptive Multi-index Collocation for Uncertainty Quantification and Sensitivity Analysis,” *in review*, 2018.
- [38] Gerstner, T., and Griebel, M., “Multilevel Monte Carlo path simulation,” *Computing*, Vol. 71, No. 1, 2003, pp. 65–87.
- [39] Gratiet, L. L., and Garnier, J., “Recursive co-kriging model for design of experiments with multiple levels of fidelity,” *International Journal for Uncertainty Quantification*, Vol. 4, No. 5, 2014, pp. 365–386.
- [40] Monschke, J. A., and Eldred, M. S., “Multilevel-Multifidelity Acceleration of PDE-Constrained Optimization,” *58th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, AIAA SciTech Forum, Grapevine, TX*, 2017.
- [41] Bichon, B. J., “Efficient Surrogate Modeling for Reliability Analysis and Design,” Ph.D. thesis, Vanderbilt University, Nashville, TN, 2010.
- [42] Dubourg, V., Sudret, B., and Bourinet, J.-M., “Reliability-based design optimization using kriging surrogates and subset simulation,” *Structural and Multidisciplinary Optimization*, Vol. 44, No. 5, 2011, pp. 673–690.
- [43] Echard, B., Gayton, N., and Lemaire, M., “AK-MCS: An active learning reliability method combining kriging and Monte Carlo simulation,” *Structural Safety*, Vol. 33, No. 2, 2011, pp. 145–154.
- [44] Chevalier, C., Bect, J., Ginsbourger, D., Vazquez, E., Picheny, V., and Richet, Y., “Fast parallel kriging-based stepwise uncertainty reduction with application to the identification of an excursion set,” *Technometrics*, Vol. 56, No. 4, 2014, pp. 455–465.
- [45] Hu, Z., and Mahadevan, S., “Global sensitivity analysis-enhanced surrogate (GSAS) modeling for reliability analysis,” *Structural and Multidisciplinary Optimization*, Vol. 53, No. 3, 2016, pp. 501–521.
- [46] Hu, Z., Nannapaneni, S., and Mahadevan, S., “Efficient kriging surrogate modeling approach for system reliability analysis,” *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, Vol. 31, No. 02, 2017, pp. 143–160.
- [47] Tu, J., Choi, K. K., and Park, Y. H., “A new study on reliability-based design optimization,” *ASME Journal of Mechanical Design*, Vol. 121, No. 4, 1999, pp. 557–564.
- [48] Archer, G. E. B., Saltelli, A., and Sobol, I. M., “Sensitivity measures, ANOVA-like techniques and the use of bootstrap,” *Journal of Statistical Computation and Simulation*, Vol. 58, No. 2, 1997, pp. 99–120.
- [49] Rabitz, H., Aliş, Ömer. F., Shorter, J., and Shim, K., “Efficient input-output model representations,” *Computer Physics Communications*, Vol. 117, No. 1-2, 1999, pp. 11–20.
- [50] Du, X., and Chen, W., “Sequential optimization and reliability assessment method for efficient probabilistic design,” *Journal of Mechanical Design*, Vol. 126, No. 2, 2004, p. 225.

- [51] Wu, Y.-T., Shin, Y., Sues, R., and Cesare, M., "Safety-factor based approach for probability-based design optimization," *Proceedings of the 42nd AIAA/ASME/ASC/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Seattle, Washington, 2001.
- [52] Kuschel, N., and Rackwitz, R., "Two basic problems in reliability-based structural optimization," *Mathematical Methods of Operations Research*, Vol. 46, No. 3, 1997, pp. 309–333.
- [53] Royset, J., Kiureghian, A. D., and Polak, E., "Reliability-based optimal structural design by the decoupling approach," *Reliability Engineering and System Safety*, Vol. 73, 2001, pp. 213–221.
- [54] Agarwal, H., "Reliability Based Design Optimization: Formulations and Methodologies," Ph.D. thesis, University of Notre Dame, Notre Dame, Indiana, 2004.
- [55] Aoues, Y., and Chateauneuf, A., "Benchmark study of numerical methods for reliability-based design optimization," *Structural and Multidisciplinary Optimization*, Vol. 41, No. 2, 2010, pp. 277–294.
- [56] Lobato, F. S., Gonçalves, M. S., Jahn, B., Cavalini, A. A., and Steffen, V., "Reliability-based optimization using differential evolution and inverse reliability analysis for engineering system design," *Journal of Optimization Theory and Applications*, Vol. 174, No. 3, 2017, pp. 894–926.
- [57] Srivastava, V., and Duraisamy, K., "Aerodynamic design of aircraft engine nozzles with consideration of model form uncertainties," *2018 AIAA Non-Deterministic Approaches Conference*, 2018, p. 2175.
- [58] Gill, P. E., Murray, W., Saunders, M. A., and Wright, M. H., *NPSOL 5.0 User's Guide*, Stanford University, 1986.