Parametric Reduced-Order Models for Probabilistic Analysis of Unsteady Aerodynamic Applications

T. Bui-Thanh[*]   K. Willcox[†]   O. Ghattas[‡]

Methodology is presented to derive reduced-order models for large-scale parametric applications in unsteady aerodynamics. The specific case considered in this paper is a computational fluid dynamic (CFD) model with parametric dependence that arises from geometric shape variations. The first key contribution of the methodology is the derivation of a linearized model that permits the effects of geometry variations to be represented with an explicit affine function. The second key contribution is an adaptive sampling method that utilizes an optimization formulation to derive a reduced basis that spans the space of geometric input parameters. The method is applied to derive efficient reduced-order models for probabilistic analysis of the effects of blade geometry variation for a two-dimensional model problem governed by the Euler equations. Reduced-order models that achieve three orders of magnitude reduction in the number of states are shown to accurately reproduce CFD Monte Carlo simulation results at a fraction of the computational cost.

I.   Introduction

Model reduction is a powerful tool that permits the systematic generation of cost-efficient representations of large-scale systems that, for example, result from discretization of partial differential equations (PDEs). We present an approach for deriving reduced models for probabilistic analysis in large-scale unsteady aerodynamic applications. The key challenges that

---

[*]Research Assistant, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139 (`tanbui@mit.edu`)

[†]Associate Professor, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139 (`kwillcox@mit.edu`)

[‡]Professor, Geological Sciences, Mechanical Engineering, Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, TX, 78712 (omar@ices.texas.edu)

must be addressed in this setting are the formulation of a computationally efficient representation of the parametric dependence that describes the uncertainty, and the derivation of reduced-order models that capture variation over a parametric input space.

Recent years have seen the widespread use of computational fluid dynamics (CFD) to solve problems arising in applications of interest in unsteady aerodynamics. In many cases, CFD formulations lead to large-scale systems of equations that are computationally expensive to solve. In many unsteady aerodynamic applications, a small number of inputs and outputs of interest can be identified, and computationally efficient reduced-order models can be obtained that preserve the desired input-output mapping. For example, the proper orthogonal decomposition (POD) method of snapshots[1] has been used widely throughout CFD applications such as aeroelasticity[2–4] and flow control.[5,6]

Quantifying the impact of variations in input parameters on system outputs of interest is critical to a number of applications, such as shape design and probabilistic analyses. For example, mistuning, or blade-to-blade variation, is an important consideration for aeroelastic analysis of bladed disks, since even small variations among blades can have a large impact on the forced response and consequently the high-cycle fatigue properties of the engine. In such applications—where the physical system must be simulated repeatedly for different inputs—the availability of reduced models can greatly facilitate the design and/or analysis task. However, to be useful in such a setting, the reduced model must provide an accurate representation of the high-fidelity CFD model over a wide range of parameters.

Most reduction techniques for large-scale systems employ a projection framework that utilizes a reduced-space basis. Methods to compute the basis in the large-scale setting include approximate balanced truncation,[7–10] Krylov-subspace methods,[11–13] and POD.[1,14,15] In the latter two cases, the quality of the reduced-order model is critically dependent on the information, generated from sampled solutions of the large-scale system, that is used to create the reduced basis. In general, selecting an appropriate set of samples to generate this information has been via an ad-hoc process. Empirical knowledge of the problem at hand has been used to sample a parameter space to generate a POD or Krylov basis for cases where the number of input parameters is small, for example optimal control applications[5,6,16,17] and parametrized design of interconnect circuits,[18] and in the case of multiple parameters describing inhomogeneous boundary conditions for parabolic PDEs.[19]

For reduction of large-scale linear time-invariant systems using multipoint rational Krylov approximations, Ref. 20 proposes a systematic method for selecting interpolation points based on a rigorous optimality criterion. To address the more general challenge of sampling a high-dimensional parameter space to build a reduced basis, the greedy algorithm was introduced in Refs. 21–24. The key premise of the greedy algorithm is to adaptively choose samples by finding the location of maximum reduced model error estimate, over a pre-

determined discrete set of parameters. The greedy algorithm was applied to find reduced models for the parametrized steady incompressible Navier-Stokes equations.[22] It was also combined with *a posteriori* error estimators for parametrized parabolic PDEs, and applied to several optimal control and inverse problems.[23, 24]

Here, we formulate the task of determining appropriate sample locations as a greedy optimization problem, which is solved using an efficient adaptive algorithm. The optimization formulation treats the parameter space as continuous; that is, we do not require the *a priori* selection of a discrete parameter set. Further, our selection criteria uses the true computed error between full-order and reduced-order outputs; thus, our approach is applicable in cases for which error estimators are unavailable. Unlike other sampling methods, the optimization-based approach scales well to systems with a large number of parameters. To further address the challenge of achieving a computationally efficient representation of the dependence of the CFD model on geometric parameters, we propose a linearization strategy that yields an affine parametric dependence.

This article is organized as follows. Section II describes the discontinuous Galerkin (DG) CFD model used in this work and formulates a linearized model for capturing the effects of geometry variations on unsteady aerodynamic response. Section III presents an overview of projection-based model reduction and then describes the proposed optimization-based approach to determine the reduced basis. Section IV demonstrates the methodology through an example that considers the effects of variations in blade geometry on the forced response of a subsonic compressor blade row. Finally, Section V presents conclusions.

II.   Linearized CFD Model with Geometric Variability

In this section, we present an overview of the DG CFD model that is used in this work. We then derive a linearized unsteady model that incorporates geometric variability in a computationally efficient way.

A.   CFD model

Recent developments in the field of CFD have led to the use of higher-order finite element discretizations for flow modeling. These schemes have advantages over traditional finite-volume methods by introducing higher-order accuracy compactly within grid elements and thus providing a significant decrease in the computational cost to obtain reliably accurate solutions. A DG formulation is used in this work. A linearized unsteady flow solver was developed and implemented as part of a larger CFD development effort that includes an adaptive meshing utility, a multigrid solution algorithm, gradient-based optimization capability, and high-order visualization.[25] Here, we briefly review the DG discretization and

solution method for the two-dimensional Euler equations, which are described in more detail in Ref. 26 and 27.

1. Nonlinear CFD model

For two-dimensional flows, the Euler equations are given by

$$\frac{\partial \mathbf{w}}{\partial t} + \nabla \cdot \mathcal{F}(\mathbf{w}) = 0, \tag{1}$$

where $\mathbf{w} = [\rho,\ \rho u,\ \rho v,\ \rho E]^T$ is the conservative state vector, and $\mathcal{F} = (\mathbf{F}^x, \mathbf{F}^y)$ is the inviscid Euler flux, where $\mathbf{F}^x = [\rho u,\ \rho u^2 + P,\ \rho uv,\ \rho uH]^T$ and $\mathbf{F}^y = [\rho v,\ \rho uv,\ \rho v^2 + P,\ \rho vH]^T$. Here, ρ is the density, u and v are respectively the $x-$ and $y-$component of velocity, E is the energy, P is the pressure, and $H = E + P/\rho$ is the total enthalpy. The equation of state is

$$P = (\gamma - 1)\left[\rho E - \frac{1}{2}\rho\left(u^2 + v^2\right)\right], \tag{2}$$

where γ is the ratio of specific heats.

As in the continuous finite element method, the first step in the DG method is to discretize the domain under consideration, Ω, into elements Ω^e. Next, a space of polynomials of degree at most p, $\mathcal{U}_h^p(\Omega^e)$, is defined on each element, where h denotes a representative element size for the discretization (e.g. the size of the smallest element). On each element Ω^e, the approximate solution $\mathbf{w}_h$ can be found by enforcing the nonlinear conservation law (1) locally, for all test functions $\mathbf{v}_h \in \mathcal{U}_h^p(\Omega^e)$:

$$\int_{\Omega^e} \mathbf{v}_h^T \frac{\partial \mathbf{w}_h}{\partial t} d\Omega^e \ - \ \int_{\partial\Omega^e} \nabla \mathbf{v}_h^T \cdot \mathcal{F}(\mathbf{w}_h) d\Omega^e + \int_{\partial\Omega^e\backslash\partial\Omega} \left(\mathbf{v}_h^+\right)^T \mathcal{H}(\mathbf{w}_h^+, \mathbf{w}_h^-, \hat{\mathbf{n}}) ds$$

$$+ \int_{\partial\Omega^e\cap\partial\Omega} \left(\mathbf{v}_h^+\right)^T \mathcal{H}_{\mathrm{bd}}(\mathbf{w}_h^+, \mathbf{w}_h^-, \hat{\mathbf{n}}) ds \ = \ 0, \tag{3}$$

where $\partial\Omega$ and $\partial\Omega^e$ are the boundaries of the entire domain Ω and the element Ω^e, respectively, and $\hat{\mathbf{n}}$ denotes the outward-pointing normal on the boundaries of the element. The terms $\mathcal{H}(\mathbf{w}_h^+, \mathbf{w}_h^-, \hat{\mathbf{n}})$ and $\mathcal{H}_{\mathrm{bd}}(\mathbf{w}_h^+, \mathbf{w}_h^-, \hat{\mathbf{n}})$ are numerical flux functions for interior and boundary edges, respectively, where $()^+$ and $()^-$ denote values taken from the interior and exterior of the element. The interior flux function is computed using the Roe-averaged flux function[28] and contributes over element boundaries that do not belong to the domain boundary, denoted by $\partial\Omega^e\backslash\partial\Omega$. The fluxes on the common boundaries of $\partial\Omega^e$ and $\partial\Omega$, denoted by $\partial\Omega^e \cap \partial\Omega$, are computed using the inner state and boundary condition data.

The final form of the DG discretization is constructed by selecting a basis for $\mathcal{U}_h^p(\Omega^e)$.

The approximate solution $\mathbf{w}_h$ on each element is assumed to be a linear combination of the finite element basis functions ψ_j,

$$\mathbf{w}_h(t, x, y) = \sum_{j=1}^{n_b} \bar{\mathbf{w}}_j(t)\psi_j(x, y), \tag{4}$$

where $\bar{\mathbf{w}}_j(t)$ gives the modal content of ψ_j on element Ω^e, and n_b is the number of basis functions required to describe $\mathcal{U}_h^p(\Omega^e)$ (e.g. $n_b = 1$ for $p = 0$ and $n_b = 3$ for $p = 1$). The complete set of unknown quantities for the DG formulation thus comprises the values of $\bar{\mathbf{w}}_j(t)$, $j = 1, \ldots, n_b$ for every element in the spatial domain. These quantities are contained in the vector $\bar{\mathbf{w}}(t) \in \mathbb{R}^n$, where n is the total number of unknowns, which depends both on the number of elements in the discretization and on the polynomial order p.

This spatial discretization together with the application of appropriate boundary conditions leads us to the following set of nonlinear ordinary differential equations (see Ref. 25 and 27 for more detail),

$$\mathbf{E}\frac{d\bar{\mathbf{w}}}{dt} + \mathbf{R}(\bar{\mathbf{w}}, \mathbf{u}) = 0, \tag{5}$$

where $\mathbf{E} \in \mathbb{R}^{n \times n}$ is the mass matrix, and $\mathbf{R}$ is the residual vector representing the final three terms of (3) plus the effects of boundary conditions. The vector $\mathbf{u}(t) \in \mathbb{R}^m$ contains m external forcing inputs that are applied through boundary conditions, such as prescribed motion of the domain boundary or incoming flow disturbances.

For steady-state flows, pseudo time-stepping is used to improve the initial transient behavior of the solver and the nonlinear system (5) is solved using a p-multigrid scheme with a line Jacobi smoother.[25,27] For unsteady computations, a second-order backward Euler temporal discretization is applied to (5). The resulting nonlinear equations are solved using a Newton solver. The motion of grid points on the domain boundary is prescribed according to the corresponding external input (e.g. prescribed motion of a blade). The resulting motion of internal grid points is computed using a Jacobi smoothing formulation.

For many practical applications, we are concerned with the prediction of a set of l output quantities of interest. We define these output quantities of interest to be contained in the vector $\mathbf{y} \in \mathbb{R}^l$ and defined by the nonlinear function $\mathbf{Y}$,

$$\mathbf{y}(t) = \mathbf{Y}(\bar{\mathbf{w}}(t)). \tag{6}$$

2. Linearized CFD model

In many cases of interest, the unsteady flow solution can be assumed to be a small perturbation from steady-state conditions. This allows the unsteady governing equations to be linearized around the steady-state flow, which reduces the computational cost of solution

considerably. The linearization of equations (5, 6) can be written in standard state-space form

$$\mathbf{E}\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \tag{7}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x}, \tag{8}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the state vector containing the n perturbations in flow unknowns from the steady-state solution $\bar{\mathbf{w}}_{ss}$, that is $\bar{\mathbf{w}}(t) = \bar{\mathbf{w}}_{ss} + \mathbf{x}(t)$. The matrices $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times p}$, and $\mathbf{C} \in \mathbb{R}^{q \times n}$ in (7) and (8) have constant coefficients evaluated at steady-state conditions and arise from the linearization of (5) and (6) as follows,

$$\mathbf{A} = \left.\frac{\partial \mathbf{R}}{\partial \bar{\mathbf{w}}}\right|_{\bar{\mathbf{w}}_{ss}}, \qquad \mathbf{B} = \left.\frac{\partial \mathbf{R}}{\partial \mathbf{u}}\right|_{\bar{\mathbf{w}}_{ss}}, \qquad \mathbf{C} = \left.\frac{\partial \mathbf{Y}}{\partial \bar{\mathbf{w}}}\right|_{\bar{\mathbf{w}}_{ss}}. \tag{9}$$

B. Linearized unsteady model with geometric variability

We next consider incorporating the effects of geometry variability into the linearized unsteady CFD model. Following Ref. 29, a general geometry, g, can be expressed as

$$g = g_n + \bar{g} + \sum_{i=1}^{n_s} \sigma_i z_i v_i, \tag{10}$$

where g_n is the nominal geometry, $\bar{g}$ is the average geometric variation, v_i are geometric mode shapes, and n_s is the number of mode shapes used to represent the variation in geometry. The geometric mode shapes could be computed, for example, by performing the principal component analysis (PCA) on a manufacturing sample of system geometries. In that case, the parameters z_i in (10) are random numbers normally distributed with zero mean and unity variance, $z_i \in N(0, 1)$, and σ_i is the standard deviation of the geometric data attributable to the ith mode; thus the product $\sigma_i z_i$ is the amount by which the mode v_i contributes to the geometry g. A detailed description of the methodology underlying this geometric model can be found in Ref. 29.

Using the model (10), a general geometry $g(\mathbf{z})$ is specified by the parameter vector $\mathbf{z} = [z_1, z_2, \ldots, z_{n_s}]^T$, which describes the geometry variability in terms of the geometry modes. The linearized CFD system corresponding to geometry $g(\mathbf{z})$ is given by

$$\mathbf{E}(\bar{\mathbf{w}}_{ss}(g(\mathbf{z})), g(\mathbf{z}))\dot{\mathbf{x}} = \mathbf{A}(\bar{\mathbf{w}}_{ss}(g(\mathbf{z})), g(\mathbf{z}))\mathbf{x} + \mathbf{B}(\bar{\mathbf{w}}_{ss}(g(\mathbf{z})), g(\mathbf{z}))\mathbf{u}, \tag{11}$$

$$\mathbf{y} = \mathbf{C}(\bar{\mathbf{w}}_{ss}(g(\mathbf{z})), g(\mathbf{z}))\mathbf{x}, \tag{12}$$

where the CFD system matrices $\mathbf{E}, \mathbf{A}, \mathbf{B}$ and $\mathbf{C}$ are in general both a nonlinear function

of the geometry, $g(\mathbf{z})$, and of the steady-state solution, $\bar{\mathbf{w}}_{ss}(g(\mathbf{z}))$, which is itself also a function of the geometry. To solve the CFD system (11), (12), for each geometry g we must firstly compute the steady-state solution, $\bar{\mathbf{w}}_{ss}(g(\mathbf{z}))$, secondly evaluate the linearized matrices $\mathbf{E}, \mathbf{A}, \mathbf{B}$ and $\mathbf{C}$, and thirdly solve the resulting large-scale linear system. This is a computationally prohibitive proposition for applications such as probabilistic analysis, where thousands of geometry perturbations may be analyzed over many random samples $\mathbf{z}$.

For convenience of notation, we write the dependence of the CFD matrices on the parameter $\mathbf{z}$ as $\mathbf{E}(\bar{\mathbf{w}}_{ss}(g(\mathbf{z})), g(\mathbf{z})) = \mathbf{E}(\mathbf{z})$, $\mathbf{A}(\bar{\mathbf{w}}_{ss}(g(\mathbf{z})), g(\mathbf{z})) = \mathbf{A}(\mathbf{z})$, $\mathbf{B}(\bar{\mathbf{w}}_{ss}(g(\mathbf{z})), g(\mathbf{z})) = \mathbf{B}(\mathbf{z})$, and $\mathbf{C}(\bar{\mathbf{w}}_{ss}(g(\mathbf{z})), g(\mathbf{z})) = \mathbf{C}(\mathbf{z})$. We use the expansion given by equation (10), which represents a general geometry as a perturbation about the average geometry $g_0 = g_n + \bar{g}$, to derive an approximate model for representing the effects of geometry variations. Instead of computing the linearized CFD matrices exactly for any random variability $\mathbf{z}$, we choose to linearize the relationships $\mathbf{E}(\mathbf{z})$, $\mathbf{A}(\mathbf{z})$, $\mathbf{B}(\mathbf{z})$, and $\mathbf{C}(\mathbf{z})$. We define the linearized unsteady CFD model for the average geometry $g_0 = g_n + \bar{g}$ by the matrices $\mathbf{E_0}, \mathbf{A_0}, \mathbf{B_0}$, and $\mathbf{C_0}$, with corresponding solution $\mathbf{x}_0$. That is, for $\mathbf{z} = 0$ we have

$$\mathbf{E}_0 \dot{\mathbf{x}}_0 = \mathbf{A}_0 \mathbf{x}_0 + \mathbf{B}_0 \mathbf{u}, \tag{13}$$

$$\mathbf{y}_0 = \mathbf{C}_0 \mathbf{x}_0. \tag{14}$$

Using a Taylor series expansion about $\mathbf{z} = 0$ for the matrix $\mathbf{A}(\mathbf{z})$ gives

$$\mathbf{A}(\mathbf{z}) = \mathbf{A}_0 + \left.\frac{\partial \mathbf{A}}{\partial z_1}\right|_{\mathbf{z}=0} z_1 + \ldots + \left.\frac{\partial \mathbf{A}}{\partial z_{n_s}}\right|_{\mathbf{z}=0} z_{n_s} + \ldots, \tag{15}$$

where the matrix partial derivatives denote componentwise derivatives, which can be evaluated through application of the chain rule. These derivatives are evaluated at nominal conditions, $\mathbf{z} = 0$. The matrices $\mathbf{E}(\mathbf{z}), \mathbf{B}(\mathbf{z})$ and $\mathbf{C}(\mathbf{z})$ can be expanded using formulae analogous to (15).

If the geometric variability (given by the product $\sigma_i z_i$) is sufficiently small, the constant and linear terms in the Taylor expansion (15) are sufficient to approximate the linearized matrices $\mathbf{A}(\mathbf{z})$ accurately, that is,

$$\mathbf{A}(\mathbf{z}) \approx \mathbf{A}_0 + \left.\frac{\partial \mathbf{A}}{\partial z_1}\right|_{\mathbf{z}=0} z_1 + \ldots + \left.\frac{\partial \mathbf{A}}{\partial z_{n_s}}\right|_{\mathbf{z}=0} z_{n_s}. \tag{16}$$

For $i = 1, 2, \ldots, n_s$, we define

$$\bar{\mathbf{E}}_i = \left.\frac{\partial \mathbf{E}}{\partial z_i}\right|_{\mathbf{z}=0}, \quad \bar{\mathbf{A}}_i = \left.\frac{\partial \mathbf{A}}{\partial z_i}\right|_{\mathbf{z}=0}, \quad \bar{\mathbf{B}}_i = \left.\frac{\partial \mathbf{B}}{\partial z_i}\right|_{\mathbf{z}=0}, \quad \bar{\mathbf{C}}_i = \left.\frac{\partial \mathbf{C}}{\partial z_i}\right|_{\mathbf{z}=0}, \tag{17}$$

where the matrices $\bar{\mathbf{E}}_i$, $\bar{\mathbf{A}}_i$, $\bar{\mathbf{B}}_i$, and $\bar{\mathbf{C}}_i$ can be computed, for example, using a finite difference approximation of the respective derivatives. The approximate linearized CFD model for any geometric variability $\mathbf{z}$ is then given by

$$\underbrace{\left(\mathbf{E}_0 + \sum_{i=1}^{n_s} \bar{\mathbf{E}}_i z_i\right)}_{\mathbf{E}(\mathbf{z})} \dot{\mathbf{x}} = \underbrace{\left(\mathbf{A}_0 + \sum_{i=1}^{n_s} \bar{\mathbf{A}}_i z_i\right)}_{\mathbf{A}(\mathbf{z})} \mathbf{x} + \underbrace{\left(\mathbf{B}_0 + \sum_{i=1}^{n_s} \bar{\mathbf{B}}_i z_i\right)}_{\mathbf{B}(\mathbf{z})} \mathbf{u}, \qquad (18)$$

$$\mathbf{y} = \underbrace{\left(\mathbf{C}_0 + \sum_{i=1}^{n_s} \bar{\mathbf{C}}_i z_i\right)}_{\mathbf{C}(\mathbf{z})} \mathbf{x}. \qquad (19)$$

It should be noted here that a number of large-scale steady-state CFD solves is required in order to determine the matrices $\mathbf{A}_0$, $\mathbf{B}_0$, $\mathbf{C}_0$, $\mathbf{E}_0$, $\bar{\mathbf{A}}_i$, $\bar{\mathbf{B}}_i$, $\bar{\mathbf{C}}_i$ and $\bar{\mathbf{E}}_i$. For example, if central difference approximations to the matrix derivatives are used, a total of $2\sum_{i=1}^{n_s} +1$ large-scale steady-state CFD solves is required. This is a one-time offline cost; once the matrices are computed, the approximate linearized system (18), (19) can be readily evaluated for an arbitrary geometry $g(\mathbf{z})$ without running the CFD steady solver.

It should also be noted that the model (18), (19) is valid only for small variations from the nominal geometry. Larger variations will incur larger errors, due to the neglect of the higher-order terms in the Taylor series expansion. Even with this restriction, the model is useful for many applications where small geometric variations are of interest; however, the approximate linearized model is still of high dimension, and thus is computationally too expensive for applications such as probabilistic analysis in which one needs to determine the unsteady aerodynamic response for many random geometries. In the next section we propose a model reduction method that enables us to further reduce the cost of solving the approximate linearized system. The key challenge that must be addressed is developing a reduced-order model that is accurate over both time and the geometric parameter space, described here by the vector $\mathbf{z}$.

III.   Model Reduction Methodology

A.   General projection framework

Most large-scale model reduction frameworks are based on a projection approach, which can be described in general terms as follows. Although the projection framework is described here for a general system that is linear in the state variables, it can also be applied to

nonlinear systems. Consider the general parameterized dynamical system

$$\mathbf{E}(\mathbf{z})\dot{\mathbf{x}} = \mathbf{A}(\mathbf{z})\mathbf{x} + \mathbf{B}(\mathbf{z})\mathbf{u}, \tag{20}$$

$$\mathbf{y} = \mathbf{C}(\mathbf{z})\mathbf{x}, \tag{21}$$

with initial condition

$$\mathbf{x}(0) = \mathbf{x}^0, \tag{22}$$

where $\mathbf{x}(\mathbf{z}, t) \in \mathbb{R}^n$ is the state vector, $\mathbf{u}(t) \in \mathbb{R}^m$ contains the m forcing inputs to the system, $\mathbf{y}(\mathbf{z}, \mathbf{x}, t) \in \mathbb{R}^l$, contains the l outputs of interest, and $\mathbf{x}^0$ is the specified initial state. The matrices $\mathbf{E} \in \mathbb{R}^{n \times n}$, $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times m}$, and $\mathbf{C} \in \mathbb{R}^{l \times n}$ in (20) and (21) may depend (possibly nonlinearly) on a set of n_s parameters contained in the vector $\mathbf{z} \in \mathbb{R}^{n_s}$. General dynamical systems of the form (20)–(22) often arise from discretization of PDEs. In that case, the dimension of the system, n, is large, and the parameters z_i could describe, for example, changes in the domain shape, boundary conditions, or PDE coefficients. The CFD model (18), (19) derived in the previous section is one example of a system of the form (20), (21); in that case the parameters z_i describe geometric shape variations.

A reduced-order model of (20)–(22) can be derived by assuming that the state $\mathbf{x}(\mathbf{z}, t)$ is represented as a linear combination of n_r basis vectors,

$$\tilde{\mathbf{x}} = \mathbf{\Phi}\mathbf{x}_r, \tag{23}$$

where $\tilde{\mathbf{x}}(\mathbf{z}, t)$ is the reduced model approximation of the state $\mathbf{x}(\mathbf{z}, t)$ and $n_r \ll n$. The projection matrix $\mathbf{\Phi} \in \mathbb{R}^{n \times n_r}$ contains as columns the basis vectors ϕ_i, i.e., $\mathbf{\Phi} = [\phi_1 \ \phi_2 \ \cdots \ \phi_{n_r}]$, and the vector $\mathbf{x}_r(\mathbf{z}, t) \in \mathbb{R}^{n_r}$ contains the corresponding modal amplitudes. Using the representation (23) together with a Petrov-Galerkin projection of the system (20)–(22) onto the space spanned by the left basis $\mathbf{\Psi} \in \mathbb{R}^{n \times n_r}$ yields the reduced-order model with state $\mathbf{x}_r(\mathbf{z}, t)$ and output $\mathbf{y}_r(\mathbf{z}, \mathbf{x}, t)$,

$$\mathbf{E}_r(\mathbf{z})\dot{\mathbf{x}}_r = \mathbf{A}_r(\mathbf{z})\mathbf{x}_r + \mathbf{B}_r(\mathbf{z})\mathbf{u}, \tag{24}$$

$$\mathbf{y}_r = \mathbf{C}_r(\mathbf{z})\mathbf{x}_r, \tag{25}$$

$$\mathbf{x}_r^0 = \mathbf{\Psi}^T \mathbf{x}(0), \tag{26}$$

where $\mathbf{E}_r(\mathbf{z}) = \mathbf{\Psi}^T \mathbf{E}(\mathbf{z})\mathbf{\Phi}$, $\mathbf{A}_r(\mathbf{z}) = \mathbf{\Psi}^T \mathbf{A}(\mathbf{z})\mathbf{\Phi}$, $\mathbf{B}_r(\mathbf{z}) = \mathbf{\Psi}^T \mathbf{B}(\mathbf{z})$, and $\mathbf{C}_r(\mathbf{z}) = \mathbf{C}(\mathbf{z})\mathbf{\Phi}$.

Projection-based model reduction techniques seek to find the bases $\mathbf{\Phi}$ and $\mathbf{\Psi}$ such that the reduced system (24)–(26) provides an accurate representation of the large-scale system (20)–(22) over the desired range of inputs $\mathbf{u}(t)$ and parameters $\mathbf{z}$. Here, we consider Galerkin projections, that is $\mathbf{\Phi} = \mathbf{\Psi}$, although our methodology holds in the general case.

B. Reduced basis for parametric input dependence

Using the general projection framework, our model reduction task becomes one of determining an appropriate reduced basis that spans both the parametric input space $\mathbf{z}$ and the space of unsteady inputs $\mathbf{u}(t)$. In the case of a linear time-invariant system, that is, a system of the form (20)–(22) with no dependence on parameters $\mathbf{z}$, a number of model reduction techniques can be used, such as Krylov-based methods and POD. To extend these techniques to the general case where the system matrices depend on the parameters $\mathbf{z}$, we require a systematic method of sampling the parametric input space.

In the case of the POD, the reduced basis is formed as the span of a set instantaneous state solutions, commonly referred to as snapshots. These snapshots are computed by solving the system (20)–(22) for selected values of the parameters $\mathbf{z}$ and selected forcing inputs $\mathbf{u}(t)$. The quality of the resulting reduced-order model is very dependent on the choice of parameters and inputs over which snapshots are computed. Two issues arise in selecting an appropriate sample set. First, choosing where and how many samples to generate is, in general, an ad-hoc process. One can use knowledge of the application at hand to determine representative inputs; however, there exist no guarantees on the quality of the resulting reduced-order model. Second, in the case that the parametric input space is of high dimension, the number of high-fidelity system solves required to generate the snapshots in an ad-hoc manner can become prohibitively large. Using standard sampling methods, a problem with just a few parameters can require a large number of samples to adequately cover the space, due to the combinatorial explosion of the number of possible parameter combinations.

Here, we use the greedy algorithm[21–24] to adaptively select snapshots, by finding the location in parameter space where the error between the full-order and reduced-order models is maximal, updating the basis with information gathered from this sample location, forming a new reduced model, and repeating the process. We formulate the greedy approach as an optimization problem that targets the error in reduced model output prediction, which is defined by introducing as constraints the systems of equations representing the full and reduced models. The optimization formulation treats the parameter space as continuous; that is, we do not require the *a priori* selection of a discrete parameter set. Further, since the optimization problem uses the true computed error between full and reduced outputs, our approach is applicable in cases for which error estimators are unavailable.

On each cycle of the greedy algorithm, the key step is to determine the location in parameter space where the error in the reduced model is maximal. We define the cost functional

$$\mathcal{G}(\mathbf{z}, \mathbf{x}, \mathbf{x}_r) = \frac{1}{2} \int_0^{t_f} \|\mathbf{y}(\mathbf{z}, \mathbf{x}, t) - \mathbf{y}_r(\mathbf{z}, \mathbf{x}_r, t)\|_2^2 \, dt = \frac{1}{2} \int_0^{t_f} \|\mathbf{C}\mathbf{x} - \mathbf{C}_r\mathbf{x}_r\|_2^2 \, dt, \qquad (27)$$

which describes the error between the full and reduced models over the parameter space $\mathbf{z}$, integrated over some time horizon of interest t_f. Given a current basis $\mathbf{\Phi}$, we find the location in parameter space of maximum error by solving the optimization problem

$$\max_{\mathbf{x},\mathbf{x}_r,\mathbf{z}} \mathcal{G} = \frac{1}{2} \int_0^{t_f} \|\mathbf{C}\mathbf{x} - \mathbf{C}_r\mathbf{x}_r\|_2^2 \, dt \tag{28}$$

subject to

$$\mathbf{E}(\mathbf{z})\dot{\mathbf{x}} = \mathbf{A}(\mathbf{z})\mathbf{x} + \mathbf{B}(\mathbf{z})\mathbf{u}, \tag{29}$$
$$\mathbf{x}^0 = \mathbf{x}(0), \tag{30}$$
$$\mathbf{E}_r(\mathbf{z})\dot{\mathbf{x}}_r = \mathbf{A}_r(\mathbf{z})\mathbf{x}_r + \mathbf{B}_r(\mathbf{z})\mathbf{u}, \tag{31}$$
$$\mathbf{x}_r^0 = \mathbf{\Phi}^T x(0), \tag{32}$$
$$\mathbf{z}_{min} \leq \mathbf{z} \leq \mathbf{z}_{max}, \tag{33}$$

where $\mathbf{z}_{min}$ and $\mathbf{z}_{max}$ are respectively lower and upper bounds on the parameter vector $\mathbf{z}$. We denote the parameter vector that solves the maximization problem (28)–(33) by $\mathbf{z}^*$. Next, we compute the solution $\mathbf{x}(\mathbf{z}^*, t)$ of the full system at the worst-case parameter value $\mathbf{z}^*$. This solution information is added to the basis $\mathbf{\Phi}$, for example using the POD. (Note that once the sample location has been found, other model reduction methods could also be employed.) The procedure is then repeated by solving the optimization problem (28)–(33) with the updated basis $\mathbf{\Phi}$. Thus, we are using a systematic, adaptive error metric based on the ability of the reduced-order model to capture the outputs of interest in order to choose the snapshot locations. This model reduction approach is summarized in the following algorithm.

Algorithm 1 *Adaptive Sampling Procedure*

1. *Given a reduced basis $\mathbf{\Phi}$, solve the optimization problem (28)–(33) to find the location in parameter space at which the error is maximized, i.e. find $\mathbf{z}^* = \arg\max \mathcal{G}(\mathbf{z})$.*

2. *If $\mathcal{G}(\mathbf{z}^*) < \epsilon$, where ϵ is the desired level of accuracy, then terminate the algorithm. If not, go to the next step.*

3. *With $\mathbf{z} = \mathbf{z}^*$, solve the full system (20)–(22) to compute the state solutions $\mathbf{x}(\mathbf{z}^*, t)$, $t = (0, t_f)$. Use the span of these state solutions to update the basis $\mathbf{\Phi}$. Go to step 1.*

In Step 3 of Algorithm 1, the basis can be updated using many of the existing model reduction methods. For example, the POD could be used to compute the span of the updated snapshot set, which would comprise the existing basis vectors and the new state solutions

$\mathbf{x}(\mathbf{z}^*, t)$. As an alternative approach, one could also solve an (inner) optimization problem to find the basis that minimizes the output error at the sample points.[30] Algorithm 1 is initialized by choosing the initial basis as the empty set, $\boldsymbol{\Phi} = \emptyset$; thus the reduced model is initially a zero-order approximation of the full model.

The optimization problem (28)–(33) that must be solved in each adaptive cycle (i.e. Step 1 of Algorithm 1) is large scale; in particular, note that the large-scale system equations appear as constraints in (29). The determination of each sample point $\mathbf{z}^*$ via solution of this optimization problem therefore requires some number of solves of the system (29), which for the large-scale problems of interest here ($n > 10^5$) is the dominant computational cost. It is therefore critical to use an efficient optimization method; that is, one that exploits the structure of the problem to offer rapid convergence. We employ recent advances in scalable algorithms for large-scale optimization of systems governed by PDEs, which have permitted solution of problems with millions of state and optimization variables.[31, 32]

In order to solve the constrained optimization problem (28)–(33), we choose to solve an equivalent bound-constrained optimization problem in the $\mathbf{z}$ variables by eliminating the state variables $\mathbf{x}$ and $\mathbf{x}_r$. That is, we replace $\min_{\mathbf{x}, \mathbf{x}_r, \mathbf{z}} \mathcal{G}(\mathbf{x}, \mathbf{x}_r, \mathbf{z})$ with $\min_{\mathbf{z}} \mathcal{G}(\mathbf{x}(\mathbf{z}), \mathbf{x}_r(\mathbf{z}), \mathbf{z})$, where the dependence of $\mathbf{x}$ and $\mathbf{x}_r$ on $\mathbf{z}$ is implicit through the full and reduced state equations (29)–(32). To deal with the bound constraints, we use the Coleman-Li approach (see, for example Ref. 33). This enables us to use the subspace trust-region interior-reflective Newton framework, proposed in Ref. 33, to solve the resulting bound-constrained optimization problem efficiently. That is, we use the conjugate gradient (CG) method to determine the subspace over which the linear system of equations arising at each Newton step is solved, and globalize by a trust region scheme (see, for example, Ref. 34). This method combines the rapid locally-quadratic convergence rate properties of Newton's method, the effectiveness of trust region globalization for treating ill-conditioned problems, and the Eisenstat-Walker idea of preventing oversolving.[35]

The gradient of $\mathcal{G}$ with respect to $\mathbf{z}$, as required by Newton's method, can be computed efficiently by an adjoint method. The Hessian-vector product as required by CG is computed on-the-fly; because it is a directional derivative of the gradient its computation similarly involves solution of state-like and adjoint-like equations. Therefore, the optimization algorithm requires solution of a pair of state and adjoint systems at each CG iteration.

Since the system dependence on the parameter $\mathbf{z}$ is nonlinear, in the general case the optimization problem (28)–(33) is non-convex. In particular, as the greedy algorithm progresses we expect the cost functional to become increasingly multi-modal, since the error function will be close to zero (below the tolerance ϵ) at each of the previous parameter sample locations. It should be noted that, while finding the global maximum is obviously preferred, convergence to a local maximum is not necessarily an adverse result. Solving the

greedy optimization problem is a heuristic to systematically find "good" sample points; at a local maximum the error is (locally) large. To avoid convergence to a local maximum close to a previous sample location, and thus explore the parameter space more widely, a random initialization of the optimization variables $\mathbf{z}$ is used for each cycle of Algorithm 1. An initial guess is accepted only if it is sufficiently far from previous sample locations, measured using a tolerance that is set relative to the parameter ranges. The stopping criterion applied in Step 2 of Algorithm 1 monitors $\mathcal{G}(\mathbf{z}^*)$, the reduced model error associated with the optimal solution $\mathbf{z}^*$. It is important to note that if $\mathcal{G}(\mathbf{z}^*)$ falls below the desired error level, this guarantees only that the local error between full and reduced model is sufficiently small. Due to the non-convexity of the optimization problem, it is possible that larger errors may exist elsewhere in the parameter space.

C.   Reduced-order linearized aerodynamic model with geometric variability

Combining the linearized unsteady model with geometric variability from Section II together with the reduced basis model reduction methodology based on adaptive sampling, we now have a method to create efficient reduced-order models that capture the effects of small geometric variations.

Using the projection framework, and a basis $\mathbf{\Phi}$ computed using Algorithm 1, the reduced-order model of (18), (19) is

$$\underbrace{\left(\mathbf{E}_{r_0} + \sum_{i=1}^{n_s} \bar{\mathbf{E}}_{r_i} z_i\right)}_{\mathbf{E}_r(\mathbf{z})} \dot{\mathbf{x}}_r = \underbrace{\left(\mathbf{A}_{r_0} + \sum_{i=1}^{n_s} \bar{\mathbf{A}}_{r_i} z_i\right)}_{\mathbf{A}_r(\mathbf{z})} \mathbf{x}_r + \underbrace{\left(\mathbf{B}_{r_0} + \sum_{i=1}^{n_s} \bar{\mathbf{B}}_{r_i} z_i\right)}_{\mathbf{B}_r(\mathbf{z})} \mathbf{u}, \qquad (34)$$

$$\mathbf{y}_r = \underbrace{\left(\mathbf{C}_{r_0} + \sum_{i=1}^{n_s} \bar{\mathbf{C}}_{r_i} z_i\right)}_{\mathbf{C}_r(\mathbf{z})} \mathbf{x}_r, \qquad (35)$$

where the reduced-order matrices are given by

$$\mathbf{E}_{r_0} = \mathbf{\Phi}^T \mathbf{E}_0 \mathbf{\Phi}, \ \mathbf{A}_{r_0} = \mathbf{\Phi}^T \mathbf{A}_0 \mathbf{\Phi}, \ \mathbf{B}_{r_0} = \mathbf{\Phi}^T \mathbf{B}_0, \ \mathbf{C}_{r_0} = \mathbf{C}_0 \mathbf{\Phi}, \qquad (36)$$

$$\bar{\mathbf{E}}_{r_i} = \mathbf{\Phi}^T \bar{\mathbf{E}}_i \mathbf{\Phi}, \ \ \bar{\mathbf{A}}_{r_i} = \mathbf{\Phi}^T \bar{\mathbf{A}}_i \mathbf{\Phi}, \ \ \bar{\mathbf{B}}_{r_i} = \mathbf{\Phi}^T \bar{\mathbf{B}}_i, \ \ \bar{\mathbf{C}}_{r_i} = \bar{\mathbf{C}}_i \mathbf{\Phi}, \qquad i = 1, 2, \ldots, n_s. \ \ (37)$$

The key enabling feature of the adaptive sampling approach is that it allows the basis $\mathbf{\Phi}$ to be computed in an efficient systematic manner, even when the dimension of the parameter space is large. The methodology also gives us a means to monitor the (local) error between reduced-order and full-order outputs. The key advantage of the linearized variability model is that it leads to an affine parameter dependence; thus the reduced-order matrices in (36)

and (37) can be evaluated offline, and the online cost of solving the reduced-order model (34), (35) does not depend on the large-scale state dimension n.

IV.   Probabilistic analysis application

The model reduction methodology is applied to probabilistic analysis of a subsonic rotor blade that moves in unsteady rigid motion. The analysis seeks to quantify the effects on the blade forced response of small variations in blade geometry. Mistuning, or blade-to-blade variation, is an important consideration for aeroelastic analysis of bladed disks, since even small variations among blades can have a large impact on the forced response and consequently the high-cycle fatigue properties of the engine. The effects of blade structural mistuning (variations in mass and stiffness properties) have been extensively studied, see for example Refs. 36–39; however, due to the prohibitively high computational cost of performing probabilistic analysis with a CFD model, the aerodynamic effects due to variations in geometry are less understood.

Geometric mistuning effects have been incorporated into structural responses of bladed disks using a mode-acceleration method to convert the effect of geometric mistuning to that of external forces of the tuned disks.[40] Truncated sets of tuned system modes compensated by static modes—generated by external forces that were constructed from mistuning—were then used to obtain efficient and accurate structural reduced models. Several studies have also found that found that a small number of PCA geometric modes can capture manufactured variability in bladed disks accurately.[29, 41, 42] Such reduced geometric variability models have been used to investigate the impact of geometric variability on axial compressor steady aerodynamic performance using Monte Carlo simulation based on a large-scale nonlinear CFD model.[29] Using Monte Carlo simulation of a CFD model to quantify the impact of geometric variability on *unsteady* performance is a computationally prohibitive proposition. For example, if the unsteady analysis for one geometry takes one minute to compute (a conservative estimate), the $O(50,000)$ such analyses that would be required for a Monte Carlo simulation would take roughly one month of CPU time. Therefore, we desire to obtain a reduced-order model that captures both unsteady response and variation over blade geometries. Our method combines the reduced geometric variability model and the adaptive model reduction methodology of Algorithm 1 to obtain a reduced-order model that is valid over a range of forcing frequencies, aerodynamic damping, and small perturbations in blade geometries, and thus enables fast and accurate probabilistic analysis.

A.   Blade forced response example

For the example presented here, the flow is modeled using the two-dimensional Euler equations written at the blade mid-section. The average geometry of the blade is shown in Figure 1 along with the unstructured grid for a single blade passage, which contains 4292 triangular elements. The Euler equations are discretized in space with the DG method described in Section II. For the case considered here, the incoming steady-state flow has a Mach number of $M = 0.113$ and a flow angle of $\beta = 59°$. Flow tangency boundary conditions are applied on the blade surfaces. To compute the steady-state flow for the nominal case, we exploit the fact that the rotor is cyclically symmetric; thus, the steady flow in each blade passage is the same and the steady-state solution can be computed on a computational domain that describes just a single blade passage. Periodic boundary conditions are applied on the upper and lower boundaries of the grid to represent the effects of neighboring blade passages.
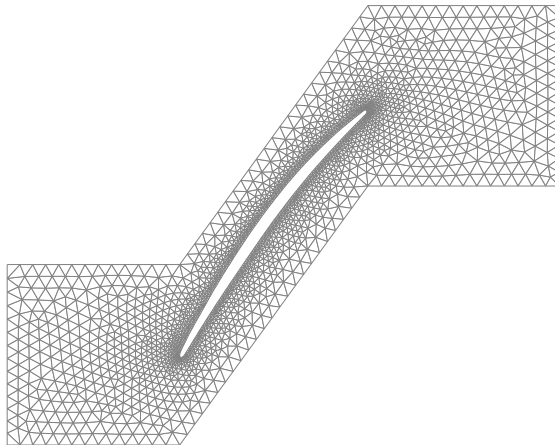

Figure 1.   Geometry and CFD mesh for a single blade passage.

A linearized model is derived for unsteady flow computations by assuming that the unsteady flow is a small deviation from steady state as described in Section IIA. An affine dependence of the linearized system matrices on the blade geometries is derived using the method described in Section IIB. This leads to a system of the form (18), (19), where the state vector, $\mathbf{x}(t)$, contains the unknown perturbation flow quantities (density, Cartesian momentum components and energy). For the DG formulation, the states are the coefficients corresponding to each nodal finite element shape function. Using linear elements, there are 12 degrees of freedom per element, giving a total state-space size of $n = 51504$ states per blade passage. For the problem considered here, the forcing input, $\mathbf{u}(t)$, describes the unsteady motion of each blade, which in this case is assumed to be rigid plunging motion

(vertical motion with no rotation). The outputs of interest, $\mathbf{y}(t)$, are the unsteady lift forces and pitching moments generated on each blade. The initial perturbation flow is given by $\mathbf{x}^0 = 0$.

B. Geometric variability model

Geometric modes were computed using a PCA of data modified from 145 actual blades, measured at thirteen sections along the radial direction. The mid-section geometries were then extracted. Thus the parameter vector $\mathbf{z}$ contains the normally distributed random variables that describe perturbations in the geometry of each blade according to the model (10). In Figure 2, we consider a geometric model that uses the two dominant variability modes, $n_s = 2$. The figure shows the lift coefficient, C_L, and moment coefficient, C_M, of a blade in response to a pulse input in plunge for a particular geometry that corresponds to $z_1 = 1.59$, $z_2 = 1.59$. The response is computed using the exact linearized CFD model, i.e. the system (11), (12) and the approximate linearized model (18), (19) with $n_s = 2$ geometry modes. For reference, the response of the nominal blade is also shown in the figure. It can be seen that despite the small perturbation in geometry, the change in lift and moment coefficient responses is significant. The approximate linearized geometric model captures the unsteady response accurately.
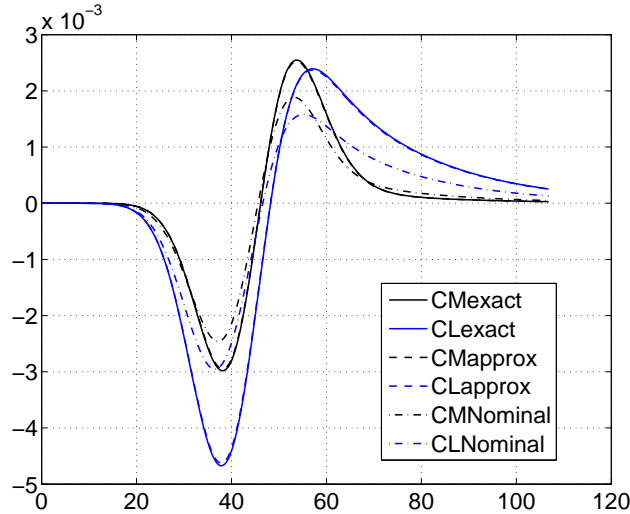

Figure 2. Lift coefficient, C_L, and moment coefficient, C_M, in response to a pulse input in blade plunge displacement for the nominal geometry and a perturbed geometry described by two geometric PCA modes with $z_1 = 1.59$, $z_2 = 1.59$. Perturbed geometry results are computed with both the exact and approximate linearized CFD model.

Table 1 shows the error in lift and moment outputs due to the linearized geometry approximation for several different blade geometries with a pulse input in plunge. The error e is defined as the 2-norm of the difference between the approximate and the exact linearized

output as a percentage of the change between the exact and the nominal output,

$$e = \frac{\sqrt{\int_0^{t_f} \|\mathbf{y}_e - \mathbf{y}_a\|_2^2 dt}}{\sqrt{\int_0^{t_f} \|\mathbf{y}_e - \mathbf{y}_o\|_2^2 dt}} \times 100\%, \tag{38}$$

where $\mathbf{y}_e$, $\mathbf{y}_a$, and $\mathbf{y}_o$ are respectively the exact, approximate, and nominal outputs. In the table, e_{C_M} denotes the error in moment coefficient response, while e_{C_L} denotes the error in lift coefficient response. In general, we expect the quality of the approximate model to be compromised as the size of the geometric perturbation increases. The errors shown in Table 1 for blade geometries in the tails of the distribution, i.e. those with large geometry variation, are deemed to be acceptable for the probabilistic application of interest here. For applications where greater accuracy for large geometry variations is important (for example, determining the probability of failure would require the tail of the distribution to be resolved accurately), the results suggest that the approximate linearized CFD system is not appropriate. In such cases, one might consider including more terms in (16), the Taylor series expansion of the CFD matrices.

Table 1. **Error in approximate linearized model predictions for a pulse input in blade displacement for several different geometries.**

| Variability amplitudes	$e_{C_M}(\%)$	$e_{C_L}(\%)$
$z_1 = 1.59, z_2 = 1.59$	5.04	2.6
$z_1 = 1.59, z_2 = -1.59$	0.3	0.1
$z_1 = -1.59, z_2 = -1.59$	2.0	0.8
$z_1 = 3.0, z_2 = 3.0$	16.6	9.2
$z_1 = 3.0, z_2 = -3.0$	4.1	2.3
$z_1 = -3.0, z_2 = -3.0$	12.4	4.7

C.   Model reduction

To create a reduced-order model for use in probabilistic analysis, the adaptive model reduction methodology of Algorithm 1 is applied to this problem. Results are shown here for the case of two blades moving with an interblade phase angle of 180°. Each blade geometry is represented by two variability modes, giving $n_s = 4$ geometric parameters in this example. Applying the adaptive model reduction methodology with $\epsilon = 10^{-6}$ and with the lift and moment coefficients as the outputs of interest yields a reduced-order model of size $m = 438$ (for two blades) after five adaptive cycles with a total of 26 high-fidelity unsteady solves. For

each high-fidelity unsteady solve, 401 snapshots are uniformly generated over the nondimensional time horizon $[0, 200]$. It took one day seven hours twenty-four minutes and fifty-two seconds CPU time to obtain a reduced model with $m = 438$ states. This reduced model accurately captures the unsteady response of the original two-blade system with 103008 states over the range of geometries described by the four geometric parameters.

As an example of an application for which this reduced model is useful, we consider probabilistic analysis of the system. Specifically, we consider the impact of blade geometry variabilities on the work per cycle, which is defined as the integral of the blade motion times the lift force over one unsteady cycle. A Monte Carlo simulation was performed in which 5000 blade geometries were selected randomly from the given distributions for each blade. The same 5000 geometries were analyzed using the approximate linearized CFD model and the reduced model. Figure 3 shows the resulting probability density functions (PDFs) of work per cycle for the first blade, computed using the approximate linearized CFD model and the reduced-order model. Figure 4 shows the PDFs of work per cycle for the second blade. Table 2 shows that the reduced model Monte Carlo simulation can be performed in 4 hours 43 minutes 16 seconds, as compared to 14 days 11 hours 47 minutes 28 seconds of computation time for the approximate linearized CFD runs. These results were obtained using Matlab[43] on a dual core personal computer with 3.2GHz Pentium processor. The table also compares the statistics of the two distributions. It can be seen from Figure 4 and Table 2 that the reduced-order model predicts the mean, variance and shape of the distribution of work per cycle very accurately. To further verify the quality of the reduced model, we use the *kstest2* Matlab function, which uses the Kolmogorov-Smirnov method,[44] to test whether the reduced WPCs and the full WPCs are drawn from a same distribution. The results show that we cannot reject the hypothesis that the distribution is the same at 5% significance level. More explicitly, the probability that the hypothesis is true is 0.9563 for the first blade and 0.9999 for the second blade.

To further compare the reduced-order and CFD results, we pick four particular geometries corresponding to the left tail, right tail, mid-left and mid-right locations on the PDF of the first blade as indicated by the circles in Figure 3(a). In Table 3 the work per cycle is given for these four blade geometries as computed by the exact CFD model, the approximate linearized CFD model, and the reduced-order model. The table shows that again the approximate linearized CFD is in very good agreement with the exact one, especially for the mid-left and mid-right cases, which have smaller variability. In addition, the accuracy of the adaptive model reduction methodology of Algorithm 1 can be clearly seen from the very good match between the approximate linearized CFD and the reduced results.
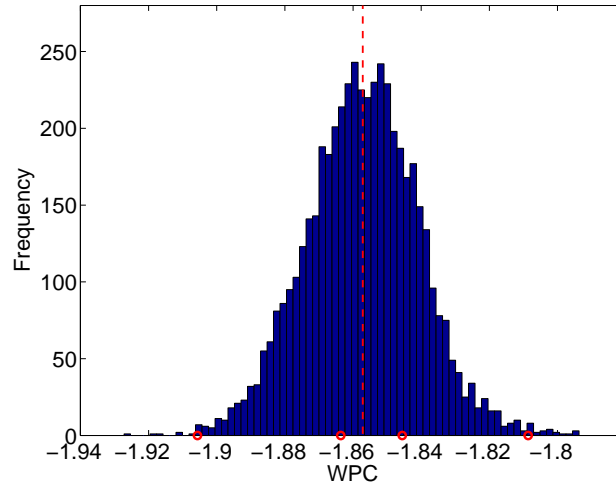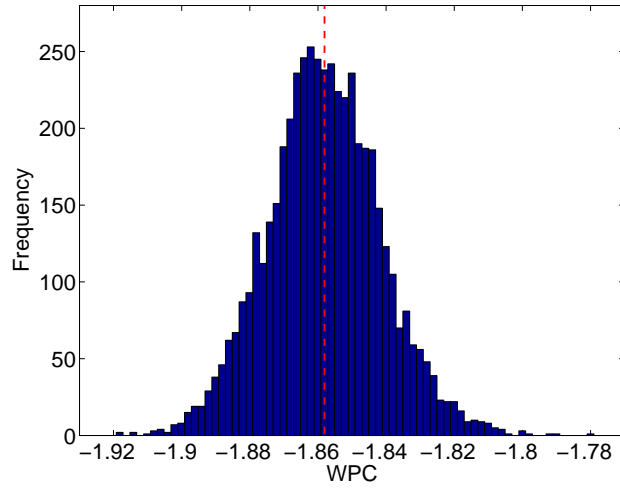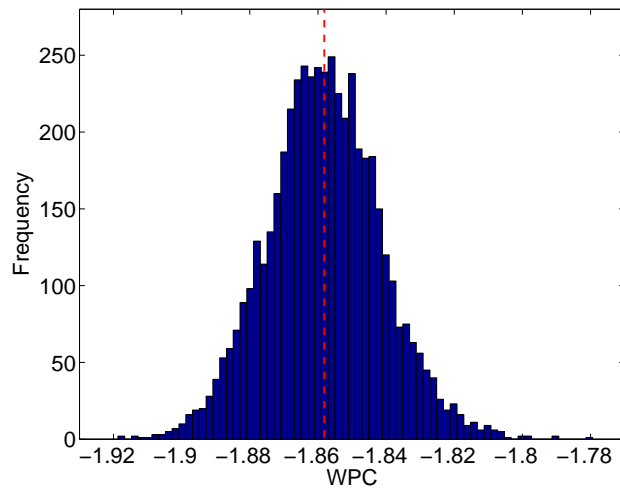
(a) Full WPC for Blade 1
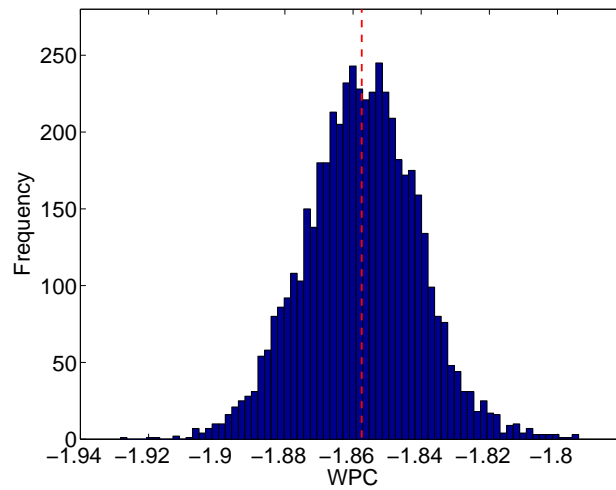
(b) Reduced WPC for Blade 1

Figure 3. Comparison of linearized CFD and reduced-order model predictions of work per cycle for Blade 1. Monte Carlo simulation results are shown for 5000 blade geometries. The same geometries were analyzed in each case. Dashed line denotes the mean.

(a) Full WPC for Blade 2

(b) Reduced WPC for Blade 2

Figure 4. Comparison of linearized CFD and reduced-order model predictions of work per cycle for Blade 2. Monte Carlo simulation results are shown for 5000 blade geometries. The same geometries were analyzed in each case. Dashed line denotes the mean.

Table 2. Linearized CFD model and reduced-order model Monte Carlo simulation results. Work per cycle (WPC) is predicted for blade plunging motion at an interblade phase angle of 180° for 5000 randomly selected blade geometries.

	CFD	Reduced
Model size	103008	438
Computation time	14 days 11 hours 47 minutes 28 seconds	4 hours 43 minutes 16 seconds
Blade 1 WPC mean	-1.8572	-1.8574
Blade 1 WPC variance	2.7484e-4	2.778e-4
Blade 2 WPC mean	-1.8581	-1.8581
Blade 2 WPC variance	2.7887e-4	2.8044e-4

Table 3. Exact CFD, approximate CFD, and reduced-order model work per cycle prediction for the four geometries indicated in Figure 3(a).

	Exact	Approximate	Reduced
Left tail	-1.8973	-1.9056	-1.9061
Mid-left	-1.8637	-1.8636	-1.8639
Mid-right	-1.8459	-1.8455	-1.8457
Right tail	-1.8014	-1.8086	-1.8085

V.   Conclusions

The key contributions of this paper are the derivation of a linearized CFD model that permits the effects of geometry variations to be represented with an explicit affine function and the development of an adaptive sampling method to derive a reduced-order basis that spans both forcing input and parameter space. Together, these contributions lead to a computationally tractable formulation for analyzing the effects of small variations in geometry on unsteady aerodynamic response. The methodology was demonstrated here for a problem that is linear in the state variables and affine in the parameter variables; however, the adaptive greedy sampling approach provides a general framework that is applicable to nonlinear problems. In the general nonlinear case, however, one must address the challenge of carrying out online reduced-order model computations in an efficient manner that does not depend on the large-scale system dimension.

Acknowledgments

References

[1]Sirovich, L., "Turbulence and the dynamics of coherent structures. Part 1: Coherent structures," *Quarterly of Applied Mathematics*, Vol. 45, No. 3, October 1987, pp. 561–571.

[2]Romanowski, M., "Reduced Order Unsteady Aerodynamic and Aeroelastic Models using Karhunen-Loéve Eigenmodes," AIAA Paper 96-194, 1996.

[3]Dowell, E. and Hall, K., "Modeling of fluid-structure interaction," *Annual Review of Fluid Mechanics*, Vol. 33, 2001, pp. 445–90.

[4]Lieu, T., Farhat, C., and Lesoinne, M., "Reduced-Order Fluid/Structure Modeling of a Complete Aircraft Configuration," *Computer Methods in Applied Mechanics and Engineering*, Vol. 195, 2006, pp. 5730–5742.

[5]Hinze, M. and Volkwein, S., "Proper orthogonal decomposition surrogate models for nonlinear dynamical systems: Error estimates and suboptimal control," *Dimension Reduction of Large-Scale Systems*, edited by P. Benner, V. Mehrmann, and D. Sorensen, Lecture Notes in Computational and Applied Mathematics, 2005, pp. 261–306.

[6]Kunisch, K. and Volkwein, S., "Control of Burgers' equation by reduced order approach using proper orthogonal decomposition," *Journal of Optimization Theory and Applications*, Vol. 102, 1999, pp. 345–371.

[7]Gugercin, S. and Antoulas, A., "A survey of model reduction by balanced truncation and some new results," *International Journal of Control*, Vol. 77, 2004, pp. 748–766.

[8]Li, J. and White, J., "Low rank solution of Lyapunov equations," *SIAM Journal on Matrix Analysis and Applications*, Vol. 24, No. 1, 2002, pp. 260–280.

[9]Penzl, T., "Algorithms for model reduction of large dynamical systems," *Linear Algebra and its Applications*, Vol. 415, No. 2–3, June 2006, pp. 322–343.

[10]Sorensen, D. and Antoulas, A., "The Sylvester equation and approximate balanced reduction," *Linear Algebra and its Applications*, Vol. 351–352, 2002, pp. 671–700.

[11]Feldmann, P. and Freund, R., "Efficient Linear Circuit Analysis by Padé Approximation via the Lanczos Process." *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 14, 1995, pp. 639–649.

[12]Gallivan, K., Grimme, E., and Van Dooren, P., "Padé Approximation of Large-Scale Dynamic Systems with Lanczos Methods," Proceedings of the 33rd IEEE Conference on Decision and Control, December 1994.

[13]Grimme, E., *Krylov Projection Methods for Model Reduction*, Ph.D. thesis, Coordinated-Science Laboratory, University of Illinois at Urbana-Champaign, 1997.

[14]Deane, A., Kevrekidis, I., Karniadakis, G., and Orszag, S., "Low-dimensional models for complex geometry flows: Application to grooved channels and circular cylinders," *Phys. Fluids*, Vol. 3, No. 10, 1991, pp. 2337–2354.

[15]Holmes, P., Lumley, J., and Berkooz, G., *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press, Cambridge, UK, 1996.

[16]Afanasiev, K. and Hinze, M., "Adaptive control of a wake flow using proper orthogonal decomposition," *Lecture Notes in Pure and Applied Mathematics*, Vol. 216, Marcel Dekker, 2001, pp. 317–332.

[17]Arian, E., Fahl, M., and Sachs, E., "Trust-Region Proper Orthogonal Decomposition for Optimal Flow Control," Tech. Rep. ICASE 2000-25, Institute for Computer Applications in Science and Engineering, May 2000.

[18]Daniel, L., Siong, O., Chay, L., Lee, K., and White, J., "Multiparameter Moment Matching Model Reduction Approach for Generating Geometrically Parameterized Interconnect Performance Models," *Transactions on Computer Aided Design of Integrated Circuits*, Vol. 23, No. 5, May 2004, pp. 678–693.

[19]Gunzburger, M., Peterson, J., and Shadid, J., "Reduced-order modeling of time-dependent PDEs with multiple parameters in the boundary data," *Comput. Methods Appl. Mech. Engrg.*, Vol. 196, 2007, pp. 1030–1047.

[20]Gugercin, S., Antoulas, A., and Beattie, C., "A rational Krylov iteration for optimal H2 model reduction," Proceedings of MTNS 2006, Japan, 2006.

[21]Veroy, K., Prud'homme, C., Rovas, D., and Patera, A., "A posteriori error bounds for reduced-basis approximation of parametrized noncoercive and nonlinear elliptic partial differential equations," AIAA Paper 2003-3847, Proceedings of the 16th AIAA Computational Fluid Dynamics Conference, Orlando, FL, 2003.

[22]Veroy, K. and Patera, A., "Certified real-time solution of the parametrized steady incompressible Navier-Stokes equations: Rigorous reduced-basis *a posteriori* error bounds," *International Journal for Numerical Methods in Fluids*, Vol. 47, 2005, pp. 773–788.

[23]Grepl, M. and Patera, A., "*A posteriori* error bounds for reduced-basis approximations of parametrized parabolic partial differential equations," *ESAIM-Mathematical Modelling and Numerical Analysis (M2AN)*, Vol. 39, No. 1, 2005, pp. 157–181.

[24]Grepl, M., *Reduced-Basis Approximation and A Posteriori Error Estimation for Parabolic Partial Differential Equations*, Ph.D. thesis, MIT, Cambridge, MA, June 2005.

[25]Darmofal, D. and Haimes, R., "Towards the next generation in computational fluid dynamics," AIAA Paper 2005-0087, presented at 43rd AIAA Aerospace Sciences Meeting and Exhibit, Reno, NV, January, 2005.

[26]Cockburn, B. and Shu, C.-W., "Runge-Kutta discontinuous Galerkin methods for convection-dominated problems," *Journal of Scientific Computing*, Vol. 16, No. 3, 2001, pp. 173–261.

[27]Fidkowski, K. and Darmofal, D., "Development of a higher-order solver for aerodynamic applications," AIAA Paper 2004-0436, presented at the 42nd AIAA Aerospace Sciences Meeting and Exhibit, Reno, NV January 5–8, 2004.

[28]Roe, P., "Approximate Riemann solvers, parametric vectors, and difference schemes," *Journal of Computational Physics*, Vol. 43, 1981, pp. 357–372.

[29]Garzon, V. and Darmofal, D., "Impact of geometric variability on axial compressor performance," *Journal of Turbomachinery*, Vol. 125, No. 4, October 2003, pp. 692–703.

[30]Bui-thanh, T., Willcox, K., Ghattas, O., and van Bloemen Waanders, B., "Goal-Oriented, Model-Constrained Optimization for Reduction of Large-Scale Systems," *Journal of Computational Physcis*, 2006, to appear.

[31]Akçelik, V., Biros, G., and Ghattas, O., "Parallel multiscale Gauss-Newton-Krylov methods for inverse wave propagation," *Proceedings of SC2002*, 2002.

[32]Akçcelik, V., Bielak, J., Biros, G., Epanomeritakis, I., Fernandez, A., Ghattas, O., Kim, E., Lopez, J., O'Hallaron, D., Tu, T., and Urbanic, J., "Terascale forward and inverse earthquake modeling," *Proceedings of SC2003*, 2003.

[33]Branch, M., Coleman, T., and Li, Y., "A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems," *SIAM J. Sci. Comput*, Vol. 21, No. 1, October 1999, pp. 1–23.

[34]Nocedal, J. and Wright, S., *Numerical Optimization*, Springer, New York, 1999.

[35]Eisenstat, S. and Walker, H., "Choosing the forcing terms in an inexact Newton method," *SIAM Journal on Scientific Computing*, Vol. 17, 1996, pp. 16–32.

[36]Castanier, M. P., Ottarsson, G., and Pierre, C., "A Reduced Order Modeling Technique for Mistuned Bladed Disks," *Journal of Vibration and Acoustics*, Vol. 119, No. 3, 1997, pp. 439–447.

[37]Yang, M.-T. and Griffin, J. H., "A Reduced Order Model of Mistuning Using a Subset of Nominal Modes," *Journal of Engineering for Gas Turbines and Power*, Vol. 123, No. 4, 2001, pp. 893–900.

[38]Petrov, E., Sanliturk, K., Ewins, D., and Elliott, R., "Quantitative Prediction of the Effects of Mistuning Arrangement on Resonant Response of a Practical Turbine Bladed Disk," *5th National Turbine Engine High Cycle Fatigue Conference, Chandler, Arizona*.

[39]Feiner, D. and Griffin, J. H., "Mistuning Identification of Bladed Disks Using a Fundamental Mistuning Model – Part II: Application," *ASME Journal of Turbomachinery*, Vol. 126, No. 1.

[40]Lim, S., Castanier, M., and Pierre, C., "Vibration Modeling of Bladed Disks Subject to Geometric Mistuning and Design Changes," *45th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*.

[41]Brown, J., Slater, J., and Grandhi, R., "Probabilistic Analysis of Geometric Uncertainty Effects of Blade Modal Response," *ASME GT2003-38557, Proceedings of ASME Turbo Expo, Atlanta, GA.*.

[42]Sinha, A., Hall, B., Cassenti, B., and Hilbert, G., "Vibratory Parameters of Blades from Coordinate

Measurement Machine (CMM) Data," *Proceedings of 9th National Turbine Engine High Cycle Fatigue (HCF) Conference, Pinehurst, North Carolina.*

[43] *MATLAB is licensed by The MathWorks, Inc., Cochituate Place, 24 Prime Park Way, Natick, MA 01760, (508)653-1415, Fax: (508)653-2997, Email: info@mathworks.com.*

[44] DeGroot, M., *Probability and Statistics*, Addison-Wesley, MA, 1991.