On the Robustness of Residual Minimization for Constructing POD-Based Reduced-Order CFD Models

David Amsallem*, Charbel Farhat†, and Matthew Zahr ‡
Stanford University, Stanford, CA, 94305-3035, USA

Residual minimization is often used in constructing nonlinear reduced-order models (ROMs) for engineering computations. In this paper, it is shown that for unsteady CFD applications, the performance of this method strongly depends on the chosen definition of the residual. To this effect, two common residual definitions are considered first. One is the standard residual associated with the governing high-dimensional discrete equations. The other is obtained by scaling the standard residual with the inverses of the volumes of the cells of the given CFD mesh. This second definition of the residual is common in high-dimensional CFD computations as it maximizes the accuracy of the computed results in the boundary layer regions where the spatial discretization is the finest. Using the proper orthogonal decomposition (POD) method for constructing a reduced-order basis and residual minimization for computing a reduced-order approximation of a CFD solution using this basis, it is shown that both aforementioned residual definitions lead to nonlinear CFD ROMs that may perform poorly. For this reason, a new definition of the residual is proposed for the purpose of nonlinear model reduction. Unlike the two previous definitions, this one is not biased by mesh spacing considerations. More importantly, using Burger's equation as a model CFD problem with shocks, and the Ahmed body problem as a representative of three-dimensional turbulent flow problems, the proposed definition of the residual is shown to lead to nonlinear ROMs that perform significantly better than their counterparts based of the first two aforementioned definitions of the CFD residual.

I. Introduction

The discretization of the unsteady Navier-Stokes equations around full aircraft configurations typically leads to large-scale systems of equations. Solving these systems is usually computationally intensive. In some cases — for example, when the flow is turbulent — it may require days of CPU time. When this happens, CFD-based parameteric studies become cost prohibitive. In this case, reduced-order CFD models [1–8] are desirable. Indeed, because they involve a much smaller number of degrees of freedom related to pre-computed reduced-order bases (ROBs) and known as generalized coordinates, well-constructed ROMs can reduce significantly computational cost while maintaining accuracy.

For compressible flows, the nonlinear functions associated with the systems of equations expressed in conservation form are non polynomials. As a result, those nonlinear model reduction approaches which are based on the expansion of the reduced nonlinear functions into monomials and are popular for incompressible flows [9] are not applicable. Instead, model reduction is typically pursued in this case at the fully discrete level [7,10,11]. The governing high-dimensional discrete equations are often obtained using an implicit scheme — except of course when the time step is dictated by accuracy rather than numerical stability conditions. This is because such a scheme typically allows larger time steps and often delivers a better computationally efficiency. At each time instance, it also leads to a nonlinear system of algebraic equations that is usually solved using the Newton-Raphson method or a variant.

*Engineering Research Associate, Department of Aeronautics and Astronautics, William F. Durand Building, Room 028A, Stanford University, Stanford, CA 94305-3035; AIAA Member

†Vivian Church Hoff Professor of Aircraft Structures, Department of Aeronautics and Astronautics, William F. Durand Building, Room 257, Stanford University, Stanford, CA 94305-3035; AIAA Fellow

‡Graduate Student, Institute of Computational and Mathematical Engineering, William F. Durand Building, Room 028, Stanford University, Stanford, CA 94305-3035; AIAA Member

Projection-based methods are natural for constructing linear and nonlinear ROMs for CFD as well as many other scientific computing applications. This is because these methods mimic the discretization process that was responsible in the first place for generating the governing high-dimensional discrete equations. In general, projection-based methods have two distinct components. The first one constructs a matrix $\mathbf{V}$ representing a ROB that is suitable for approximating the solution $\mathbf{w}$ of the problem of interest in the subspace defined by this ROB — that is, $\mathbf{w} \approx \mathbf{V}\mathbf{w}_r$. The second one computes the generalized coordinates of this approximation, $\mathbf{w}_r$. In CFD, proper orthogonal decomposition (POD) by the method of snapshots [12] is perhaps the most popular procedure for constructing a ROB matrix $\mathbf{V}$. Residual minimization, which is ubiquitous in scientific computing and particularly in linear algebra, has recently gained popularity for computing the generalized coordinates $\mathbf{w}_r$ [2, 7, 10, 11].

In general, for a given set of discrete systems, the definition of a residual is not unique. For example, multiple residuals can be defined by scaling differently the "standard" residual. As a result, whereas the concept of residual minimization is independent of the definition of the residual, there is evidence that the performance of this procedure for determining the generalized coordinates $\mathbf{w}_r$ associated with a model reduction method can strongly depend on the chosen definition of the residual [13].

For the time dependent Navier-Stokes equations written in conservation form, the standard definition of the residual is the instantaneous unbalance of the quantities that must be conserved. When the problem of interest contains boundary layers, it is common practice however to scale the standard residual with the inverses of the cell volumes of the CFD mesh. This leads to a second definition of the residual that is useful for the following reasons:

1. Scaling the standard residual with the inverses of the cell volumes of the CFD mesh magnifies the residual in the small cells. Given a stopping criterion, a convergence tolerance or a finite number of iterations not to exceed, the solution of the governing high-dimensional discrete equations delivered by an iterative procedure is in this case most accurate in the flow regions where the cells are the smallest. This is highly desirable because the smallest cells are typically located in the flow regions where accuracy is most desired.

2. A pseudo time marching solution procedure is usually equipped with local time stepping. Hence for a steady-state problem or an unsteady problem where each system of equations arising from implicit time discretization is solved by dual time stepping, scaling the entries of the standard residual by the inverses of the cell volumes accelerates convergence to the desired solution.

For linearized CFD problems, it was recently shown [13] that scaling the standard residual by the inverses of the cell volumes before performing model reduction has however an adverse effect on the numerical stability of the resulting linearized CFD ROM, whereas working with the standard residual for the same purpose does not. For nonlinear CFD problems, it is shown in this paper that both aforementioned residual definitions may lead to nonlinear CFD ROMs that perform poorly. For this reason, a third and weighted residual definition is proposed. More importantly, it is shown that the new definition of the CFD residual proposed in this paper leads to nonlinear CFD ROMs that perform significantly better than their counterparts based of the first two aforementioned residual definitions.

II. High-Dimensional Discrete CFD Models

Consider a computational fluid domain Ω discretized by a CFD mesh with N non-overlapping cells or control volumes $\{\Omega_i\}_{i=1}^N$. Using this mesh, the semi-discretization of the Naviver-Stokes equations by a finite volume method [a] leads to a system of N ordinary differential equations (ODEs) of the form

$$\frac{d\mathbf{w}_i}{dt} + \frac{1}{|\Omega_i|}\mathbf{f}_i(\mathbf{w}) = \mathbf{0}, \tag{1}$$

where t denotes time, $\mathbf{w}_i \in \mathbb{R}^m$ is the local fluid state vector of m conservative variables defined at vertex i associated with cell Ω_i, $|\Omega_i|$ denotes the volume of this cell, and $\mathbf{f}_i$ denotes the vector of semi-discrete

[a] Whereas a finite volume semi-discretization method is used in this work, the methodologies, results, and conclusions formulated in this paper also hold for finite difference and finite element spatial discretizations.

convective and diffusive fluxes at vertex i.

$$\mathbf{w} = \begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_N \end{bmatrix} \in \mathbb{R}^{mN}$$

denotes the global fluid state vector.

An equivalent and more convenient form of equation (1) above is

$$|\Omega_i| \frac{d\mathbf{w}_i}{dt} + \mathbf{f}_i(\mathbf{w}) = \mathbf{0}. \tag{2}$$

Equation (2) can be written in global vector form as

$$\mathbf{A} \frac{d\mathbf{w}}{dt} + \mathbf{f}(\mathbf{w}) = \mathbf{0}, \tag{3}$$

and equation (1) can be written in a similar global vector form as

$$\frac{d\mathbf{w}}{dt} + \mathbf{A}^{-1}\mathbf{f}(\mathbf{w}) = \mathbf{0}, \tag{4}$$

where

$$\mathbf{f}(\mathbf{w}) = \begin{bmatrix} \mathbf{f}_1(\mathbf{w}) \\ \vdots \\ \mathbf{f}_N(\mathbf{w}) \end{bmatrix} \in \mathbb{R}^{mN},$$

$$\mathbf{A} = \begin{bmatrix} |\Omega_1|\mathbf{I}_m & & & (0) \\ & |\Omega_2|\mathbf{I}_m & & \\ & & \ddots & \\ (0) & & & |\Omega_N|\mathbf{I}_m \end{bmatrix} \mathbb{R}^{mN \times mN},$$

and $\mathbf{I}_m$ denotes the identity matrix of dimension m.

Equations (3) and (4) can be discretized in time, for example, using the implicit backward difference formula (BDF) which is popular in CFD. For simplicity but without any loss of generality, the two-point BDF scheme, which is equivalent to the backward Euler scheme, is considered here. Hence, if $t^0 = 0 < t^1 < \cdots < t^{N_t} = T$ is a discretization of the time interval $[0, T]$ and $\mathbf{w}^n \approx \mathbf{w}(t^n), n \in \{1, \cdots, N_t\}$, a discrete counterpart of equation (3) is

$$\mathbf{A} \frac{\mathbf{w}^n - \mathbf{w}^{n-1}}{\Delta t^n} + \mathbf{f}(\mathbf{w}^n) = \mathbf{0}, \tag{5}$$

where $\Delta t^n = t^n - t^{n-1}$, and a discrete counterpart of equation (4) is

$$\frac{\mathbf{w}^n - \mathbf{w}^{n-1}}{\Delta t^n} + \mathbf{A}^{-1}\mathbf{f}(\mathbf{w}^n) = \mathbf{0}. \tag{6}$$

III. Residuals in Descriptor and Non-Descriptor Forms

Equation (3) is known as the *descriptor form* of the governing system of ODEs. The residual associated with its time-discretization by the two-point BDF scheme (5) is

$$\mathbf{r}_D(\mathbf{w}^n) = \mathbf{A} \frac{\mathbf{w}^n - \mathbf{w}^{n-1}}{\Delta t^n} + \mathbf{f}(\mathbf{w}^n) \in \mathbb{R}^{mN}. \tag{7}$$

This residual, which is referred to in the remainder of this paper as the residual in descriptor form, corresponds to what was referred to in the introduction of this paper as the standard residual. It represents the instantaneous unbalance of the fluid quantities that must be conserved.

On the other hand, equation (4) is known as the *non-descriptor form* of the governing system of ODEs. The residual associated with its time-discretization by the two-point BDF scheme (6) is

$$\mathbf{r}_{ND}(\mathbf{w}^n) = \frac{\mathbf{w}^n - \mathbf{w}^{n-1}}{\Delta t^n} + \mathbf{A}^{-1}\mathbf{f}(\mathbf{w}^n) \in \mathbb{R}^{mN}. \tag{8}$$

It is referred to in this work as the residual in non-descriptor form.

From equation (7) and equation (8), it follows that

$$\mathbf{r}_{ND}(\mathbf{w}^n) = \mathbf{A}^{-1}\mathbf{r}_D(\mathbf{w}^n), \tag{9}$$

which shows that the residual in non-descriptor form is obtained by scaling the standard residual with the inverses of the cell volumes. Again, this residual is popular in CFD analysis for the reasons outlined in the introduction section of this paper.

Here, the Newton-Raphson method is used to solve both nonlinear algebraic systems of equations $\mathbf{r}_D(\mathbf{w}^n) = \mathbf{0}$ and $\mathbf{r}_{ND}(\mathbf{w}^n) = \mathbf{0}$. Since the matrix $\mathbf{A}$ is invertible, the same solution can be expected in both cases.

IV. POD/Minimum Residual Based Nonlinear Model Reduction

A. POD-based reduced-order approximation

The dimension of the high-dimensional CFD model (CFD HDM) outlined in Section II can be reduced by searching at each time instance $n \in \{1, \cdots, N_t\}$ for a solution increment in a subspace of dimension k defined by a ROB represented by a matrix $\mathbf{V} \in \mathbb{R}^{mN \times k}$. This leads to a solution of the problem of interest that can be written as each n-th time instance as follows

$$\mathbf{w}^n \approx \mathbf{w}^{n-1} + \mathbf{V}\mathbf{w}_k^n, \tag{10}$$

where $\mathbf{w}_k^n \in \mathbb{R}^k$ and $k \ll mN$. The ROB matrix $\mathbf{V}$ can be constructed by a number of methods. In this work, the POD procedure using the method of snapshots [12] is used for this purpose.

B. Residual minimization

At each time instance t^n, both residuals $\mathbf{r}_D(\mathbf{w}^n)$ and $\mathbf{r}_{ND}(\mathbf{w}^n)$ associated with the approximation (10) are non-zero. Consequently, $\mathbf{w}_k^n$ can be determined by minimizing whichever residual is considered in the least-squares sense

$$\min_{\mathbf{w}_k^n \in \mathbb{R}^k} \left\| \mathbf{r}(\mathbf{w}^{n-1} + \mathbf{V}\mathbf{w}_k^n) \right\|_2^2, \tag{11}$$

where $\mathbf{r}$ is either $\mathbf{r}_D$ or $\mathbf{r}_{ND}$. Equation (11) above is typically solved using the Gauss-Newton method [14].

In principle, the solutions of the equations $\mathbf{r}_D(\mathbf{w}^n) = \mathbf{0}$ and $\mathbf{r}_{ND}(\mathbf{w}^n) = \mathbf{0}$ are identical. However, when the CFD mesh is graded — that is, when $\mathbf{A}$ is not a uniform scaling matrix (multiple of the identity matrix) — the solutions of the minimization problems $\min_{\mathbf{w}^n} \|\mathbf{r}_D(\mathbf{w}^n)\|_2^2$ and $\min_{\mathbf{w}^n} \|\mathbf{r}_{ND}(\mathbf{w}^n)\|_2^2$ and therefore the two obtained sets of generalized coordinates $\{\mathbf{w}_k^n\}_{n=1}^{N_t}$ may differ.

As emphasized in Section I and recalled in Section II, the residual in non-descriptor form $\mathbf{r}_{ND}$ favors the minimization of the entries of the residual vector corresponding to the small cells of the CFD mesh. Therefore, one could expect *a priori* that the approximate solutions produced by the resulting CFD ROM will tend be more accurate in these small cells than in the large ones. On the other hand, the residual in descriptor form $\mathbf{r}_D$ favors the minimization of the entries of the residual vector associated with the large cells. For this reason, one could also expect that the approximate solutions delivered by the CFD ROM constructed using this residual will tend to be more accurate in the large cells of the CFD mesh than in the small ones. Unfortunately, it will be shown in Section V of this paper that the minimum residual approach outlined above may produce in both cases nonlinear CFD ROMs that perform poorly. For this reason, a third, new residual definition is proposed and discussed in Section IV.B.3. However, before introducing this new definition of the CFD residual, important results presented in [13] for *linear* model reduction using projection methods and both descriptor and non-descriptor forms of governing equations are recalled. Because these results have motivated the design of the residual definition proposed in Section IV.B.3, connections between residual minimization and projection methods are also established.

1. *Results for linear model reduction using projection methods*

Unlike in the nonlinear case, CFD problems can be reduced in the linear (or linearized) case at the semi-discrete level. This can be more convenient as the result is a CFD ROM that is independent of the time-discretization algorithm, and can be used for other purposes than unsteady flow analysis. For example, it can be used for flow stability studies using eigenvalue analysis. Furthermore, whereas residual minimization can still be used in this case for determining the generalized coordinates $\mathbf{w}_r$ of the approximation $\mathbf{w} \approx \mathbf{V}\mathbf{w}_r$, projection of the residual on a left basis $\mathbf{U}$ of the same dimensions as $\mathbf{V}$ is more common for this purpose. When $\mathbf{U} = \mathbf{V}$, the resulting projection is referred to as a Galerkin projection. When $\mathbf{U} \neq \mathbf{V}$, it is referred to as a Petrov-Galerkin projection.

For a linearized CFD problem with a source term or forcing function, the governing semi-discrete linear time invariant (LTI) equations can be written in descriptor form as follows

$$\mathbf{A}\frac{d\mathbf{w}}{dt} + \mathbf{H}\mathbf{w} + \mathbf{b}\mathbf{u}(t) = \mathbf{0}, \tag{12}$$

where $\mathbf{H} \in \mathbb{R}^{mN \times mN}$ is the Jacobian matrix of the semi-discrete fluxes $\mathbf{f}$ with respect to the semi-discrete state vector $\mathbf{w}$, $\mathbf{b}\mathbf{u}(t) \in \mathbb{R}^{mN}$ is the source term or forcing function, $\mathbf{b} \in \mathbb{R}^{mN \times l}$ is some given matrix, and $\mathbf{u}(t) \in \mathbb{R}^{l}$ is a given input vector. The reduction of (12) by an approximation of the form $\mathbf{w}(t) \approx \mathbf{V}\mathbf{w}_r(t)$, where $\mathbf{V} \in \mathbb{R}^{mN \times k}$ is a ROB, and projection of the resulting residual in descriptor form, $\mathbf{r}_D$, results in the semi-discrete Galerkin ROM

$$\left(\mathbf{V}^T\mathbf{A}\mathbf{V}\right)\frac{d\mathbf{w}_r}{dt} + \left(\mathbf{V}^T\mathbf{H}\mathbf{V}\right)\mathbf{w}_r + \left(\mathbf{V}^T\mathbf{b}\right)\mathbf{u}(t) = \mathbf{0}, \tag{13}$$

where the superscript T denotes the matrix transpose.

On the other hand, the semi-discrete LTI equations governing a linearized CFD problem can be written in non-descriptor form as follows

$$\frac{d\mathbf{w}}{dt} + \left(\mathbf{A}^{-1}\mathbf{H}\right)\mathbf{w} + \left(\mathbf{A}^{-1}\mathbf{b}\right)\mathbf{u}(t) = \mathbf{0}. \tag{14}$$

Similarly, the reduction of the above equation by a Galerkin projection onto a subspace represented by the same ROB $\mathbf{V}$ can be written as

$$\left(\mathbf{V}^T\mathbf{V}\right)\frac{d\mathbf{w}_r}{dt} + \left(\mathbf{V}^T\mathbf{A}^{-1}\mathbf{H}\mathbf{V}\right)\mathbf{w}_r + \left(\mathbf{V}^T\mathbf{A}^{-1}\mathbf{b}\right)\mathbf{u}(t) = \mathbf{0}. \tag{15}$$

In [13], it was shown that for CFD-based LTI systems, the Galerkin ROMs of the form (13) emanating from the residual in descriptor form $\mathbf{r}_D$ are typically stable, whereas their counterparts of the form (15) emanating from the residual in non-descriptor form $\mathbf{r}_{ND}$ tend to be unstable.

As stated above, the objective here is to establish some connections between the results recalled above and residual minimization, in order to identify a definition of the CFD residual that is suitable for nonlinear model reduction using this technique. This is performed next by connecting the Galerkin projections of the time-discrete versions of equations (12) and (14) and the minimization of their respective residuals in the least-squares sense.

2. *Connections between the residual minimization method and classical projection methods*

For nonlinear compressible CFD problems, model reduction is often performed, as already mentioned above, at the discrete level and using residual minimization for determining the generalized coordinates. Applying this model reduction approach to linearized CFD problems in view of establishing connections with classical projection methods begins with the selection of a residual definition and proceeds as follows.

The residual in descriptor form associated with the time-discretization of (12) by the implicit two-point BDF scheme at time t^n can be written as

$$\mathbf{r}_D(\mathbf{w}^n) = \mathbf{A}\frac{\mathbf{w}^n - \mathbf{w}^{n-1}}{\Delta t^n} + \mathbf{H}\mathbf{w}^n + \mathbf{b}\mathbf{u}^n.$$

Similarly, the residual in non-descriptor form associated with the time-discretization of (14) by the implicit two-point BDF scheme at time t^n can be written as

$$\mathbf{r}_{ND}(\mathbf{w}^n) = \frac{\mathbf{w}^n - \mathbf{w}^{n-1}}{\Delta t^n} + \left(\mathbf{A}^{-1}\mathbf{H}\right)\mathbf{w}^n + \left(\mathbf{A}^{-1}\mathbf{b}\right)\mathbf{u}^n.$$

For a given residual definition denoted here by $\mathbf{r}$, solving the minimization problem $\min_{\mathbf{w}_r} \|\mathbf{r}(\mathbf{V}\mathbf{w}_r)\|_2^2$ leads to the ROM $\mathbf{V}^T\mathbf{J}(\mathbf{V}\mathbf{w}_r)^T\mathbf{r}(\mathbf{V}\mathbf{w}_r) = \mathbf{0}$, where $\mathbf{J}$ denotes the Jacobian of $\mathbf{r}$ with respect to $\mathbf{w} \approx \mathbf{V}\mathbf{w}_r$. For example for $\mathbf{r} = \mathbf{r}_D$, it leads to

$$\mathbf{V}^T \left(\frac{1}{\Delta t^n}\mathbf{A} + \mathbf{H} \right)^T \left((\mathbf{A}\mathbf{V}) \frac{\mathbf{w}_r^n - \mathbf{w}_r^{n-1}}{\Delta t^n} + (\mathbf{H}\mathbf{V})\,\mathbf{w}_r^n + \mathbf{b}\mathbf{u}^n \right) = \mathbf{0}, \tag{16}$$

which is a reduced order version of the CFD model implied by $\mathbf{r}_D = \mathbf{0}$. Connecting this ROM to its counterpart (13) obtained by Galerkin projection requires first constructing its semi-discrete version. This is performed here as follows.

Multiplying equation (16) by Δt^n and taking the limit when $\Delta t^n \to 0$ using the expansion

$$\frac{\mathbf{w}_r^n - \mathbf{w}_r^{n-1}}{\Delta t^n} = \frac{d\mathbf{w}_r}{dt} + \mathcal{O}(\Delta t^n) \tag{17}$$

leads to

$$\mathbf{V}^T\mathbf{A}^T \left((\mathbf{A}\mathbf{V}) \frac{d\mathbf{w}_r}{dt} + (\mathbf{H}\mathbf{V})\,\mathbf{w}_r + \mathbf{b}\mathbf{u}(t) + \mathcal{O}(\Delta t) \right) = \mathbf{0}.$$

Keeping in the above expression only the lowest order terms delivers the minimum residual based semi-discrete linear CFD ROM

$$\left(\mathbf{V}^T\mathbf{A}^T\mathbf{A}\mathbf{V}\right) \frac{d\mathbf{w}_r}{dt} + \left(\mathbf{V}^T\mathbf{A}^T\mathbf{H}\mathbf{V}\right)\mathbf{w}_r + \left(\mathbf{V}^T\mathbf{A}^T\mathbf{b}\right)\mathbf{u}(t) = \mathbf{0}. \tag{18}$$

Similarly for $\mathbf{r} = \mathbf{r}_{ND}$, solving the minimization problem $\min_{\mathbf{w}_r} \|\mathbf{r}(\mathbf{V}\mathbf{w}_r)\|_2^2$ and performing the same asymptotic analysis as above leads to the minimum residual based semi-discrete linear CFD ROM

$$\left(\mathbf{V}^T\mathbf{V}\right) \frac{d\mathbf{w}_r}{dt} + \left(\mathbf{V}^T\mathbf{A}^{-1}\mathbf{H}\mathbf{V}\right)\mathbf{w}_r + \left(\mathbf{V}^T\mathbf{A}^{-1}\mathbf{b}\right)\mathbf{u}(t) = \mathbf{0}. \tag{19}$$

From (18) and (13), it follows that for linear CFD problems, determining the generalized coordinates of a reduced-order solution of the form $\mathbf{w} \approx \mathbf{V}\mathbf{w}_r$ by minimizing the residual in descriptor form leads to a linear Petrov-Galerkin semi-discrete ROM with $\mathbf{U} = \mathbf{A}\mathbf{V}$. Hence, this ROM is biased by the large cells of the CFD mesh. Furthermore, because of the Petrov-Galerkin rather than Galerkin nature of this ROM, no conclusion about its stability tendency can be directly anticipated from the results published in [13].

In the same manner, from (19) and (15), it follows that for linear CFD problems, minimizing the residual in non-descriptor form leads to the same linear Galerkin semi-discrete ROM as that obtained by Galerkin projection. In [13], it was shown that such a ROM, which emanates from the residual in non-descriptor form, $\mathbf{r}_{ND}$, tends to be unstable.

In summary, the analysis performed above for the reduction of linear CFD problems suggests that in the nonlinear case, the residual minimization approach based on the residual in non-descriptor form can be expected to lead to nonlinear CFD ROMs that are unreliable, and that based on the residual in descriptor form to lead to ROMs whose performance is at least mesh dependent. Hence, both anticipated conclusions — which are shown in Section V to hold — suggest finding an alternative definition of the CFD residual that is more suitable for nonlinear model reduction using residual minimization.

3. Residual in hybrid form

Consider again the time-discrete equations (5) and (6). Let

$$\mathbf{r}_H(\mathbf{w}^n) = \mathbf{A}^{\frac{1}{2}} \frac{\mathbf{w}^n - \mathbf{w}^{n-1}}{\Delta t^n} + \mathbf{A}^{-\frac{1}{2}}\mathbf{f}(\mathbf{w}^n). \tag{20}$$

From equation (7), equation (8) and the above definition, it follows that

$$\mathbf{r}_H(\mathbf{w}^n) = \mathbf{A}^{-\frac{1}{2}}\mathbf{r}_D(\mathbf{w}^n) = \mathbf{A}^{\frac{1}{2}}\mathbf{r}_{ND}(\mathbf{w}^n) = \frac{1}{2}\left(\mathbf{A}^{-\frac{1}{2}}\mathbf{r}_D(\mathbf{w}^n) + \mathbf{A}^{\frac{1}{2}}\mathbf{r}_{ND}(\mathbf{w}^n)\right),$$

which implies that $\mathbf{r}_H$ is the residual associated with the combination of the descriptor form (5) and non-descriptor form (6) of the governing discrete equations using the matrix coefficients $\mathbf{A}^{-\frac{1}{2}}$ and $\mathbf{A}^{\frac{1}{2}}$, respectively. This residual is referred to in this work as the residual in *hybrid form*.

Qualitatively, the residual in hybrid form (20) balances the effects of the large and small cells of the CFD mesh. To assess *a priori* its potential for robust model reduction using residual minimization, the same analysis performed in Section IV.B.2 using $\mathbf{r}_D$ and $\mathbf{r}_{ND}$ is repeated here using $\mathbf{r}_H$.

For linear CFD problems time-discretized using the two-point BDF scheme, the residual in hybrid form is

$$\mathbf{r}_H(\mathbf{w}^n) = \mathbf{A}^{\frac{1}{2}} \frac{\mathbf{w}^n - \mathbf{w}^{n-1}}{\Delta t^n} + \left(\mathbf{A}^{-\frac{1}{2}}\mathbf{H}\right)\mathbf{w}^n + \left(\mathbf{A}^{-\frac{1}{2}}\mathbf{b}\right)\mathbf{u}^n. \tag{21}$$

Solving in this case the minimization problem $\min_{\mathbf{w}_r} \|\mathbf{r}_H(\mathbf{V}\mathbf{w}_r)\|_2^2$ leads to

$$\mathbf{V}^T\left(\frac{1}{\Delta t^n}\mathbf{A}^{\frac{1}{2}} + \mathbf{A}^{-\frac{1}{2}}\mathbf{H}\right)^T\left(\left(\mathbf{A}^{\frac{1}{2}}\mathbf{V}\right)\frac{\mathbf{w}_r^n - \mathbf{w}_r^{n-1}}{\Delta t^n} + \left(\mathbf{A}^{-\frac{1}{2}}\mathbf{H}\mathbf{V}\right)\mathbf{w}_r^n + \left(\mathbf{A}^{-\frac{1}{2}}\mathbf{b}\right)\mathbf{u}^n\right) = \mathbf{0}, \tag{22}$$

which is a reduced order version of the CFD model implied by $\mathbf{r}_H = \mathbf{0}$. Next, multiplying equation (22) by Δt^n and taking the limit when $\Delta t^n \to 0$ using the expansion (17) leads to

$$\mathbf{V}^T\mathbf{A}^{\frac{1}{2}^T}\left(\left(\mathbf{A}^{\frac{1}{2}}\mathbf{V}\right)\frac{d\mathbf{w}_r}{dt} + \left(\mathbf{A}^{-\frac{1}{2}}\mathbf{H}\mathbf{V}\right)\mathbf{w}_r + \left(\mathbf{A}^{-\frac{1}{2}}\mathbf{b}\right)\mathbf{u}(t) + \mathcal{O}(\Delta t)\right) = \mathbf{0}.$$

Retaining in the above expression only the lowest order terms and noting that $\mathbf{A}^T = \mathbf{A}$ yields the following linear CFD ROM

$$\left(\mathbf{V}^T\mathbf{A}\mathbf{V}\right)\frac{d\mathbf{w}_r}{dt} + \left(\mathbf{V}^T\mathbf{H}\mathbf{V}\right)\mathbf{w}_r + \left(\mathbf{V}^T\mathbf{b}\right)\mathbf{u}(t) = \mathbf{0}. \tag{23}$$

From (23) and (13), it follows that for linear CFD problems, determining the generalized coordinates of a reduced-order solution of the form $\mathbf{w} \approx \mathbf{V}\mathbf{w}_r$ by minimizing the residual in hybrid form leads to the same ROM as that obtained by performing a Galerkin projection of the residual in descriptor form. Given that it was shown in [13] that such a ROM is typically stable, and given that the residual $\mathbf{r}_H$ balances the effects of the large and small cells of the CFD mesh instead of favoring one of them, one can reasonably expect residual minimization using $\mathbf{r}_H$ to produce nonlinear ROMs that perform better than those that can be obtained by the same approach but using $\mathbf{r}_D$ or $\mathbf{r}_{ND}$. This expectation, which is justified by all analyses performed so far, is numerically verified next using a model CFD problem with shocks, and a benchmark turbulent flow problem.

V. Applications

A. One-dimensional model problem with shocks

Consider the following one-dimensional initial boundary value problem (IBVP) based on Burger's equation ($m = 1$) and previously discussed in [15]

$$\frac{\partial \mathcal{W}}{\partial t}(x,t) + 0.5\frac{\partial(\mathcal{W}^2(x,t))}{\partial x} = 0.02e^{0.02x},\ (x,t) \in [0,100]\times[0,70] \tag{24}$$

$$\mathcal{W}(x,0) = 1,\ x \in [0,100]$$

$$\mathcal{W}(0,t) = 3,\ t > 0. \tag{25}$$

The exact solution of its steady-state version is

$$\mathcal{W}^{ex}(x) = \sqrt{2e^{0.02x} + 7},\ x \in [0,100]. \tag{26}$$

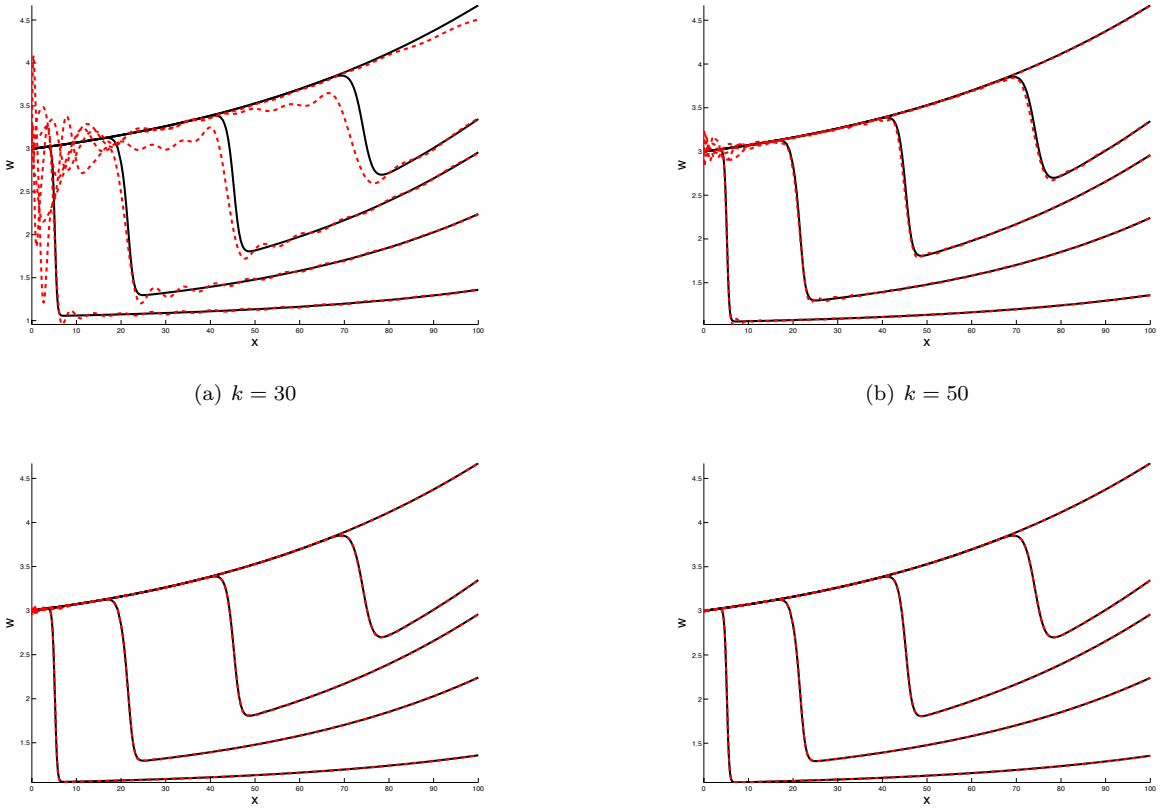
First, the above IBVP problem is discretized on a mesh with 1000 cells using a finite volume method. This mesh is referred to in the captions of the corresponding figures as the coarse mesh. It is designed to be very fine in the vicinity of $x = 0$ and very coarse in the vicinity of $x = 100$, in order to highlight the effect of mesh grading on the various residual definitions considered in this paper. The ratio between the lengths of the largest cell and the smallest cell is 1.9×10^{13}. Using the two-point BDF scheme, 1000 snapshots of the discrete solution $\mathbf{w}$ are collected using the CFD HDM of dimension 1000 associated with this mesh. Then, the POD method is applied to these snapshots to construct various ROBs of dimension $30 \leq k \leq 100$, and residual minimization is applied using all three forms of the residual discussed in this paper to construct for each considered ROB the corresponding nonlinear ROM. The performance of each generated CFD ROM

is assessed by comparing the results it delivers to their counterparts delivered by the CFD HDM. To this effect, in all figures shown below, these results are reported as functions of space, for various time instances ranging from $t = 0$ to $t = 50$ where the steady-state regime is attained, one curve per time instance. In each case, the reference high-dimensional solutions are shown using solid black lines, and the reduced-order solutions are shown using red dashed lines.

Figure 1 reports the solutions computed using the ROMs based on the residual in descriptor form, $\mathbf{r}_D$, for various values of $30 \leq k \leq 100$. As expected, the computed solutions are more accurate in the vicinity of $x = 100$ where the cells of the CFD mesh are the largest and therefore the entries of $\mathbf{A}$ are the largest. For $k \geq 70$, accuracy is achieved however throughout the computational domain.

Similarly, Figure 2 reports the counterpart solutions computed using the ROMs based on the residual in non-descriptor form, $\mathbf{r}_{ND}$. As expected, these solutions are found to be more accurate in the subdomain $x \in [0, 10]$ where the cells of the CFD mesh are the smallest, and therefore the entries of $\mathbf{A}^{-1}$ are the largest. For $k < 70$, large discrepancies can be observed away from $x = 0$, but for $k \geq 70$, accuracy is achieved in the entire computational domain.

Figure 3 reports the solutions computed using the ROMs based on the residual in hybrid form, $\mathbf{r}_H$. The reader can observe that as expected, accuracy is achieved uniformly in the computational domain. Furthermore, even when the dimension of the constructed ROM is as small as $k = 30$, a good level of accuracy is obtained.



(a) $k = 30$

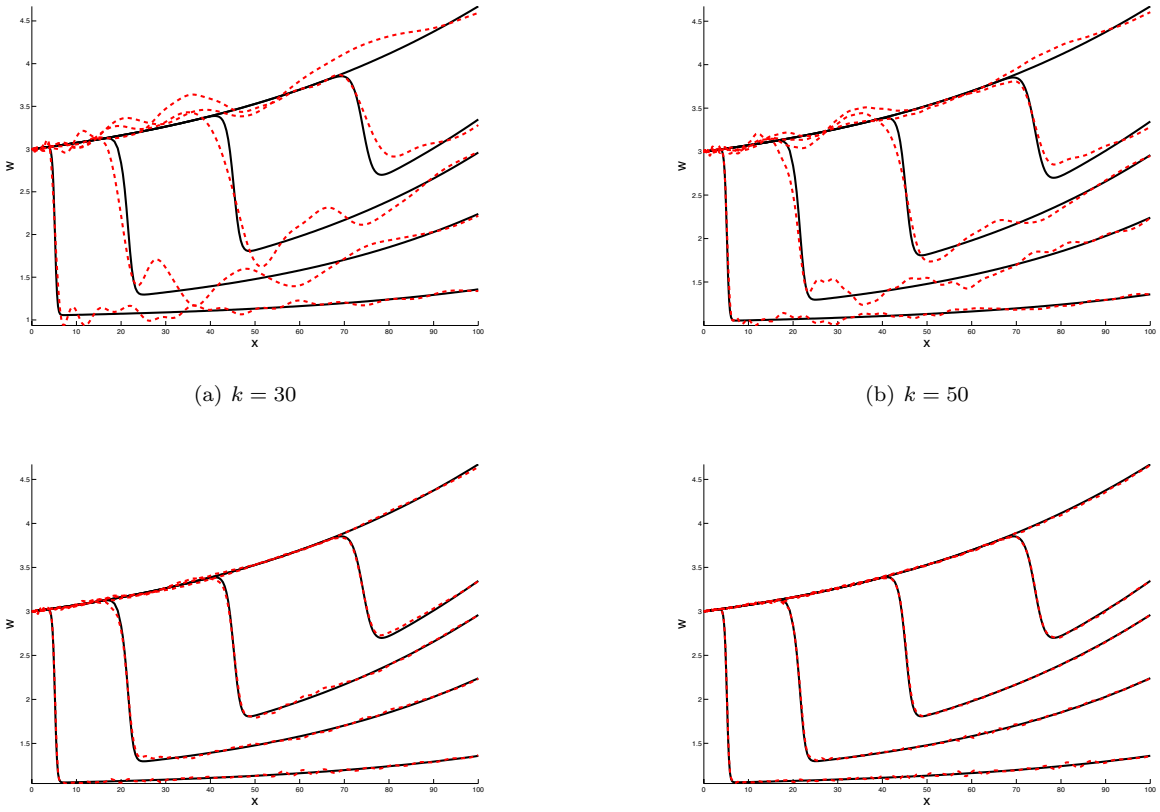
(b) $k = 50$

(c) $k = 70$

(d) $k = 90$

Figure 1. Burger's problem: performance of the residual minimization based nonlinear ROMs emanating from the residual in descriptor form (coarse mesh)

At the steady-state, the $\mathcal{L}_2$-norms of the relative errors of the computed CFD HDM and ROM solutions (measured with respect to the exact solution) are reported in Figure 4. The reader can observe that when the CFD ROM emanates from the residual in hybrid form, the solution it predicts converges to that predicted by the HDM for $k = 50$. Similarly, the reader can observe that in this sense, the CFD ROM emanating from the residual in descriptor form converges to the HDM for $k = 80$, and that emanating from the residual in



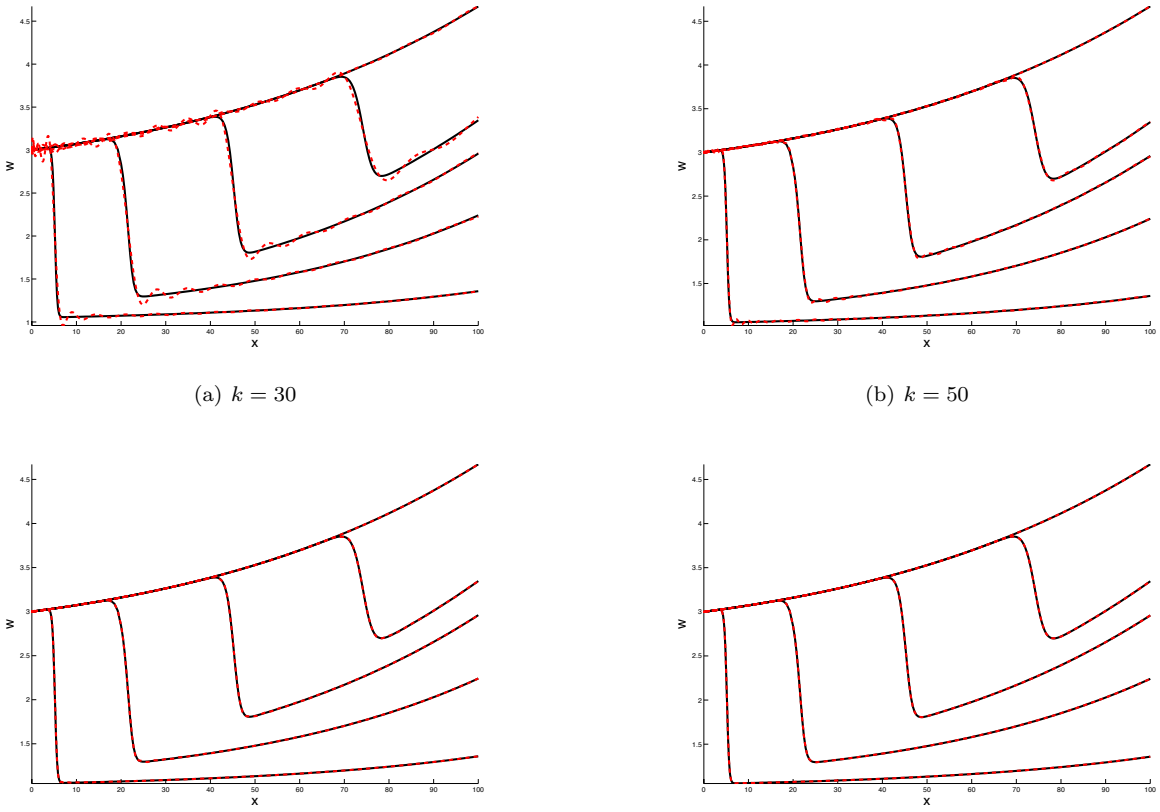
(a) $k = 30$

(b) $k = 50$

(c) $k = 70$

(d) $k = 90$

Figure 2. Burger's problem: performance of the residual minimization based nonlinear ROMs emanating from the residual in non-descriptor form (coarse mesh)



(a) $k = 30$

(b) $k = 50$

(c) $k = 70$

(d) $k = 90$

Figure 3. Burger's problem: performance of the residual minimization based nonlinear ROMs emanating from the residual in hybrid form (coarse mesh)

non-descriptor form has not yet converged for $k = 100$.

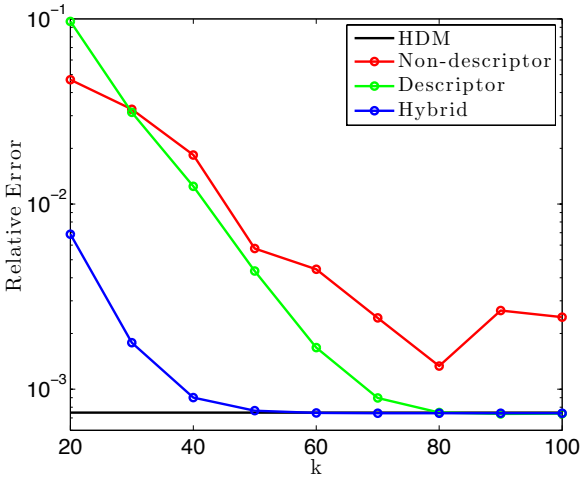


Figure 4. Burger's problem: convergence of the ROM solutions (coarse mesh)

Next, a similar set of numerical experiments is carried out using however a CFD HDM based on a mesh with 10000 cells. This mesh is referred to in the captions of the corresponding figures as the fine mesh. Like its coarser version, it is designed so that it is finer in the vicinity of $x = 0$, and coarser near $x = 100$. The ratio between the lengths of the largest and smallest cells is 2×10^{13}. Using the two-point BDF scheme on this mesh, 1000 snapshots of the HDM solution $\mathbf{w}$ are collected. Various ROMs of dimension $25 \leq k \leq 250$ are also constructed using the same methodology as before. The performance of all constructed ROMs is assessed and graphically depicted in Figures 5, 6 and 7 in the same manner and using the same format as in the case of the coarse mesh. The reader can verify that all results illustrated in these figures lead to the same conclusions as previously.

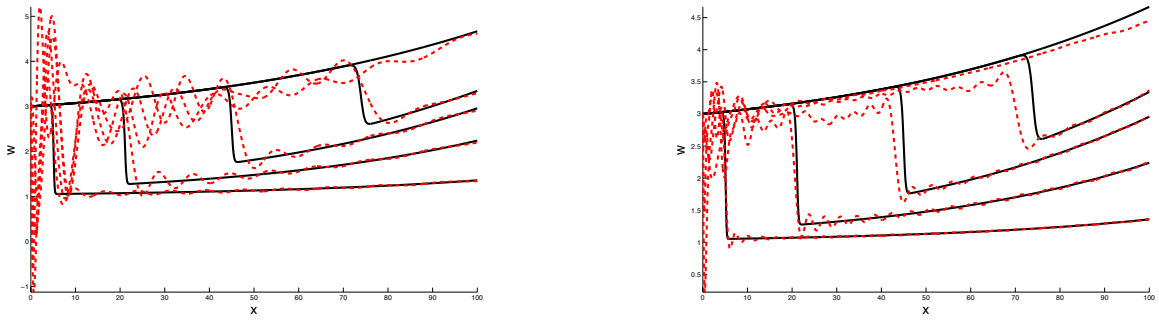
Finally, the $\mathcal{L}_2$-norms of the relative errors of the steady-state solutions computed using the HDM and various constructed ROMs are reported in Figure 8. The reader can observe that in this case, when the CFD ROM emanates from the residual in hybrid form, the solution it predicts converges to that predicted by the HDM for $k \approx 150$. The CFD ROM emanating from the residual in descriptor form converges to the HDM for $k = 250$, but that emanating from the residual in non-descriptor form is far from convergence even for $k = 250$.

B. Three-dimensional turbulent flow problem

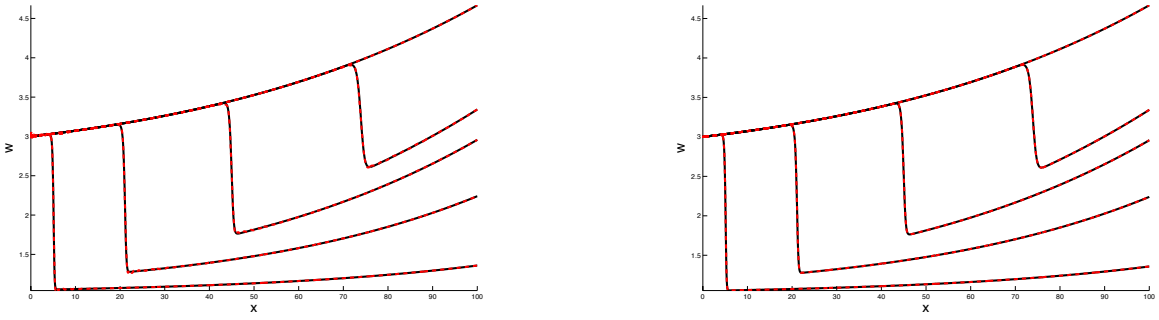
Consider next the simulation of the three-dimensional, unsteady, turbulent flow past the Ahmed body [16] using a compressible flow solver. The free-stream velocity is set to $V_\infty = 60$ m/s, in which case the Reynolds number is 4.29×10^6. As in [7] where this problem was also considered for nonlinear model reduction based on residual minimization, a Detached Eddy Simulation (DES) turbulence model is used together with Reichardt's wall law. The HDM constructed for this problem is based on a CFD mesh with $N = 2,890,434$ cells. Because of the one-parameter turbulence model used in this case, $m = 6$, and therefore the dimension of the HDM is $m \times N = 17,342,604$.

The geometry of the body and CFD mesh constructed for computing the flow around it are shown in Figure 9. The ratio between the volumes of the largest and smallest cells is 0.97×10^8. The governing equations are semi-discretized using a second-order finite volume method and time discretized using the three-point BDF scheme. First, the HDM simulation is performed and snapshots are collected. The time-averaged drag coefficient predicted by the HDM is $\bar{C}_D = 0.258$. Next, two ROBs of dimension $k = 150$ and $k = 200$ are built and used to generate a series of nonlinear ROMs by minimizing all three residuals considered in this work. The obtained ROMs are then applied to the computation of the turbulent flow associated with the same configuration presented above.

Figures 10 and 11 compare the time histories of the drag coefficient obtained using the HDM and the

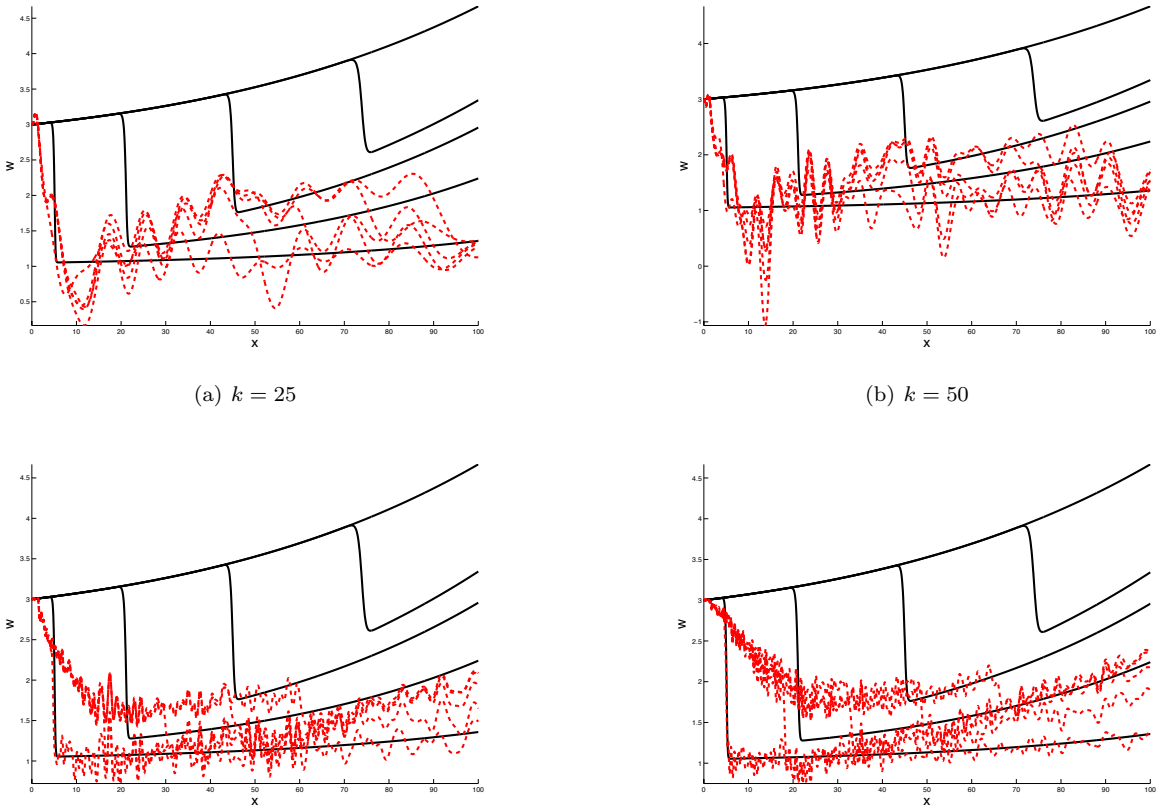


(a) $k = 25$ (b) $k = 50$



(c) $k = 150$ (d) $k = 250$

Figure 5. Burger's problem: performance of the residual minimization based nonlinear ROMs emanating from the residual in descriptor form (fine mesh)



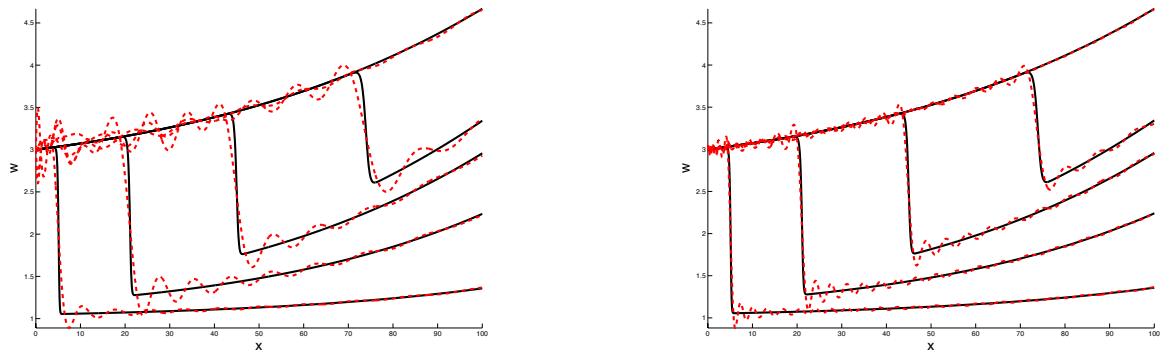
(a) $k = 25$

(b) $k = 50$

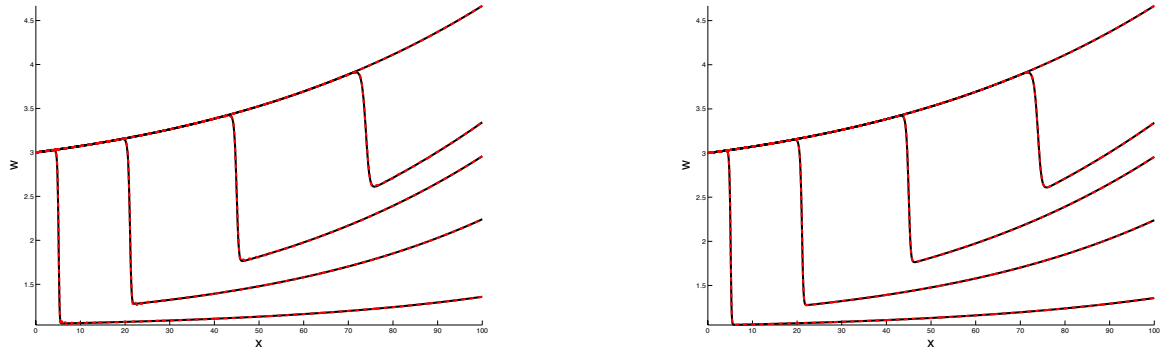
(c) $k = 150$

(d) $k = 250$

Figure 6. Burger's problem: performance of the residual minimization based nonlinear ROMs emanating from the residual in non-descriptor form (fine mesh)



(a) $k = 25$ (b) $k = 50$



(c) $k = 150$ (d) $k = 250$

Figure 7. Burger's problem: performance of the residual minimization based nonlinear ROMs emanating from the residual in hybrid form (fine mesh)

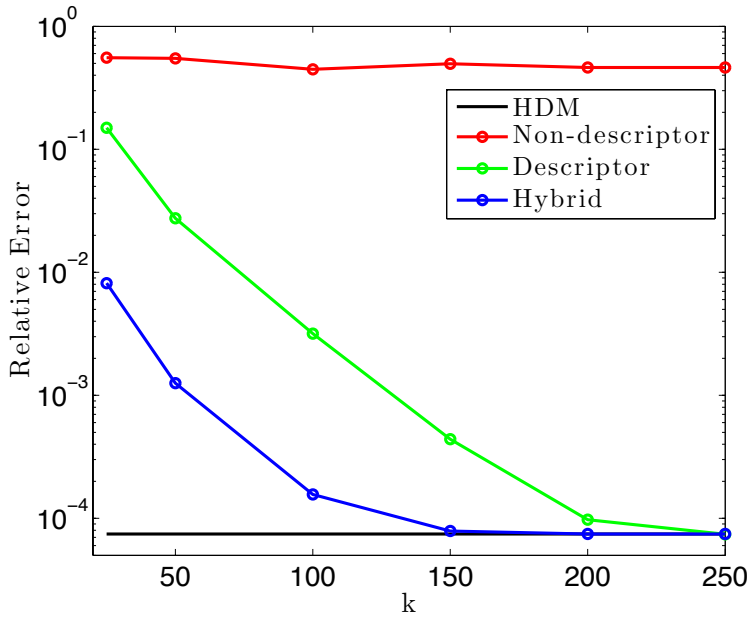
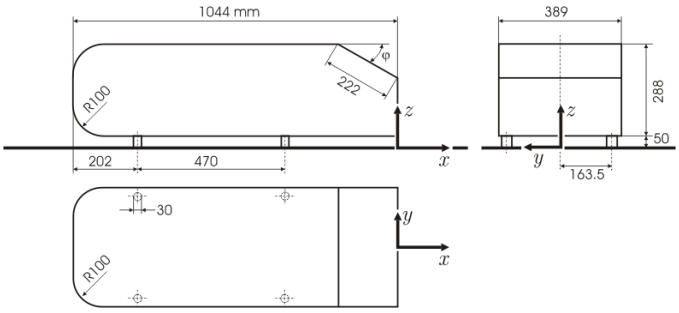
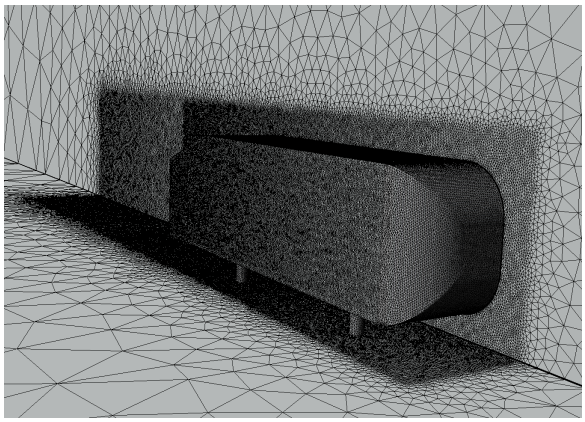


Figure 8. Burger's problem: convergence of the ROM solutions (fine mesh)



(a) Geometry (from Ref. 16)



(b) CFD mesh near the body

Figure 9. The Ahmed body problem

nonlinear ROMs based on all three considered residual definitions. Table 1 reports the corresponding time averaged value of this coefficient. The ROM solutions computed using the residuals in hybrid and non-descriptor forms lead to drag time histories that match well their HDM counterpart, with the predictions associated with the residual in hybrid form being always the most accurate. In particular, for a ROB of dimension $k = 150$, the minimization of the residual in non-descriptor form leads to erroneous predictions of the drag coefficient for $t \geq 0.03$ s, while that of the residual in hybrid form leads to predictions that match much better the HDM counterpart. On the other hand, minimizing the residual in descriptor form leads to very inaccurate results and occasionally to nonlinear instabilities that cause the ROM computations to end prematurely. This is because this form of the residual focuses on the large cells of the computational domain which are typically located far from the body. As a result, the obtained solution is inaccurate near the body and delivers an erroneous prediction of the drag.

Residual form	Non-descriptor		Descriptor		Hybrid	
ROM size	$\bar{C}_D$	Relative error	$\bar{C}_D$	Relative error	$\bar{C}_D$	Relative error
$k = 150$	0.249	3.0 %	0.236	8.5 %	0.254	1.5 %
$k = 200$	0.255	0.84 %	0.236	8.5 %	0.256	0.69 %

Table 1. Ahmed body problem: time averaged drag coefficient predicted by three nonlinear ROMs based on three different residual definitions (HDM reference value $\bar{C}_D = 0.258$)

VI. Conclusions

Nonlinear model reduction often relies on residual minimization for computing the generalized coordinates of the solution of interest in a given reduced-order basis. The residual minimization process itself depends

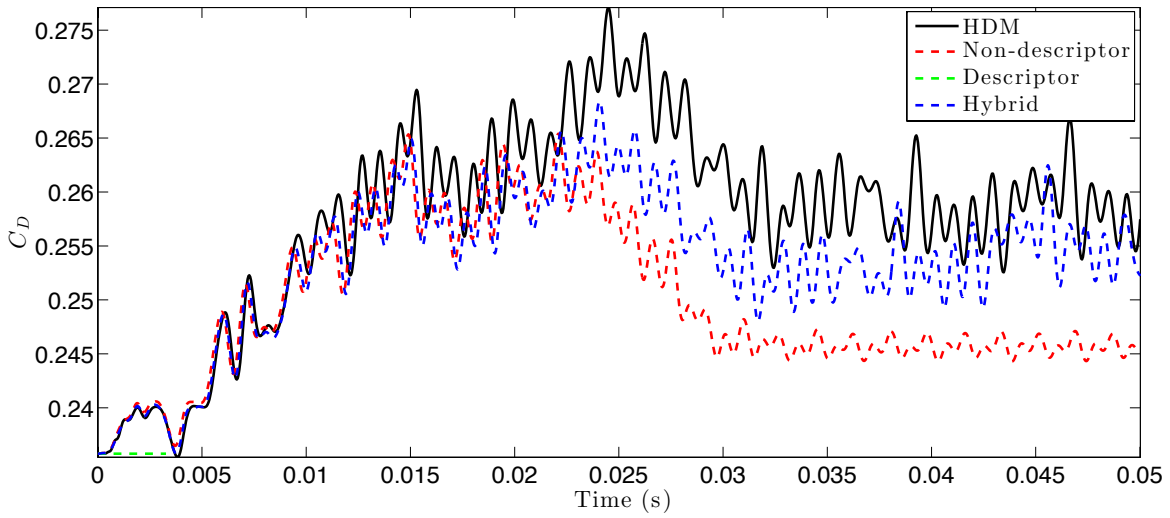


Figure 10. The Ahmed body problem: time histories of the drag coefficient predicted by the HDM and its three reduced-order counterparts of dimension $k = 150$

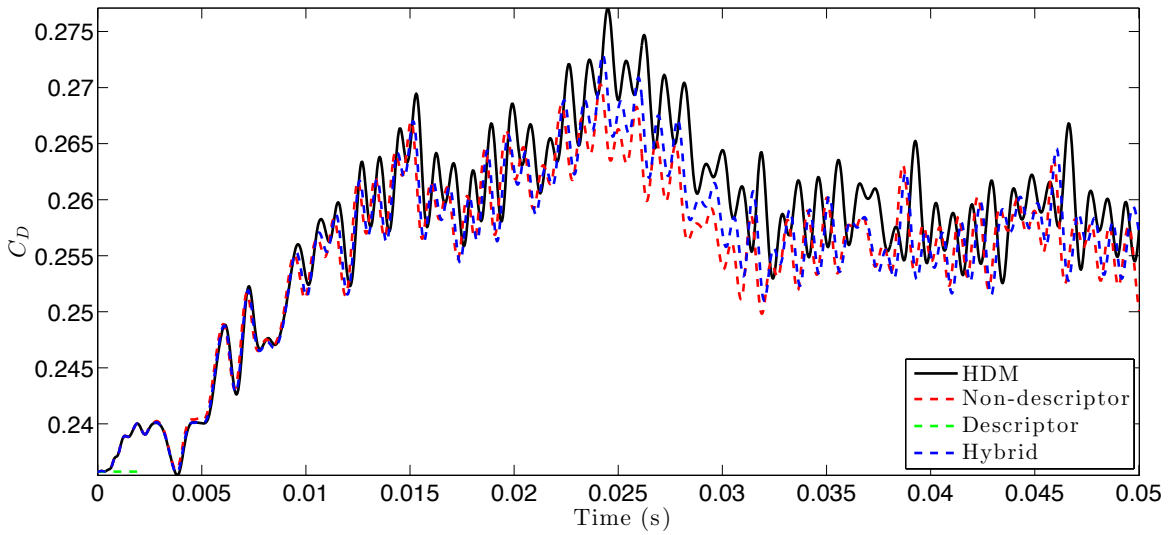


Figure 11. The Ahmed body problem: time histories of the drag coefficient predicted by the HDM and its three reduced-order counterparts of dimension $k = 150$

on the chosen definition of the residual. For a given set of governing equations, this definition is not unique. Consequently, the performance of residual minimization based reduced-order models (ROMs) can strongly depend on the adopted definition of the residual, particularly before the convergence of their size. For example, two residuals are commonly used in computational fluid dynamics (CFD). The first one is the standard residual which can be interpreted as the unbalance of the quantities that must be conserved, and is referred to in ordinary differential equation (ODE) vocabulary as the residual in descriptor form. The second residual is obtained by scaling the residual in descriptor form with the inverses of the volumes of the cells of the CFD mesh. In ODE parlance, this residual is a residual in non-descriptor form. When the reduction of a CFD model is performed using residual minimization based on the residual in descriptor form, the resulting ROMs are biased by the largest cells of the mesh. Therefore, before their size is increased to reach convergence, they tend to deliver accurate results primarily in the regions of the computational domain where the cells are the largest. On the other hand, when residual minimization is based on the residual in non-descriptor form, the resulting ROMs are biased by the smallest cells. Therefore, before their size is increased to reach convergence, they tend to perform well primarily in the regions of the computational domain where the cells are the smallest. For these reasons, a third residual definition is introduced in this paper and labeled as the residual in hybrid form. This residual balances the effects of the small and large cells in a graded mesh and therefore does not bias the model reduction process by mesh spacing considerations. For linearized CFD problems, ROMs emanating from the proper orthogonal decomposition (POD) method and the minimization of this proposed residual are identical to their counterparts constructed using the classical Galerkin projection method. Therefore, they tend to be stable. For nonlinear CFD problems with shocks and turbulent flows, ROMs constructed using POD and the minimization of the residual in hybrid form deliver a bette performance than their counterparts based on POD and the minimization of the residual in descriptor or non-descriptor form.

Acknowledgments

The first and second authors acknowledge partial support by the Army Research Laboratory through the Army High Performance Computing Research Center under Cooperative Agreement W911NF-07-2-0027, and partial support by the Office of Naval Research under Grant N00014-11-1-0707. The third author acknowledges the support of the Department of Energy Computational Science Graduate Fellowship. The content of this publication does not necessarily reflect the position or policy of any of these sponsors, and no official endorsement should be inferred.

References

[1]Hall, K. C., Thomas, J. P., and Dowell, E. H., "Proper orthogonal decomposition technique for transonic unsteady aerodynamic flows," *AIAA Journal*, Vol. 38, No. 10, 2000, pp. 1853–1862.

[2]LeGresley, P. and Alonso, J., "Airfoil design optimization using reduced order models based on proper orthogonal decomposition," *AIAA Paper 2000-2545 Fluids 2000 Conference and Exhibit, Denver, CO*, 2000.

[3]Lieu, T. and Farhat, C., "Adaptation of aeroelastic reduced-order models and application to an F-16 configuration," *AIAA Journal*, Vol. 45, No. 6, 2007, pp. 1244–1257.

[4]Amsallem, D., Farhat, C., and Lieu, T., "Aeroelastic Analysis of F-16 and F-18/A Configurations Using Adapted CFD-Based Reduced-Order Models," *48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, 23-26 April 2007, Honolulu, HI*, 2007.

[5]Amsallem, D. and Farhat, C., "Interpolation method for adapting reduced-order models and application to aeroelasticity," *AIAA Journal*, Vol. 46, No. 7, 2008, pp. 1803–1813.

[6]Amsallem, D., Cortial, J., and Farhat, C., "Toward real-time computational-fluid-dynamics-based aeroelastic computations using a database of reduced-order information," *AIAA Journal*, Vol. 48, No. 9, 2010, pp. 2029–2037.

[7]Carlberg, K., Farhat, C., Cortial, J., and Amsallem, D., "The GNAT method for nonlinear model reduction: effective implementation and application to computational fluid dynamics and turbulent flows," *Journal of Computational Physics*, Vol. 242, 2013, pp. 623–647.

[8]Washabaugh, K., Amsallem, D., Zahr, M., and Farhat, C., "Nonlinear Model Reduction for CFD Problems Using Local Reduced Order Bases," *AIAA-2012-2686, AIAA Fluid Dynamics and Co-located Conferences and Exhibit, New-Orleans, LA*, June 2012, pp. 1–16.

[9]Balajewicz, M., Dowell, E. H., and Noack, B., "A Novel Model Order Reduction Approach for Navier-Stokes Equations at High Reynolds Number," *arXiv.org*, Nov. 2012.

[10]Bui-Thanh, T., Willcox, K., and Ghattas, O., "Parametric Reduced-Order Models for Probabilistic Analysis of Unsteady Aerodynamic Applications," *AIAA Journal*, Vol. 46, No. 10, 2008, pp. 2520–2529.

[11]Carlberg, K., Bou-Mosleh, C., and Farhat, C., "Efficient nonlinear model reduction via a least-squares Petrov-Galerkin

projection and compressive tensor approximations," *nternational Journal for Numerical Methods in Engineering*, Vol. 86, 2011, pp. 155–181.

[12]Sirovich, L., "Turbulence and the dynamics of coherent structures. Part I: Coherent structures," *Quarterly of applied mathematics*, Vol. 45, No. 3, 1987, pp. 561–571.

[13]Amsallem, D. and Farhat, C., "On the stability of reduced-order linearized computational fluid dynamics models based on POD and Galerkin projection: descriptor vs non-descriptor forms," *Modeling, Simulation and Applications*, 2013.

[14]Farhat, C., Tezaur, R., and Djellouli, R., "On the solution of three-dimensional inverse obstacle acoustic scattering problems by a regularized Newton method," *Inverse Problems*, 2002, pp. 1229–1246.

[15]Rewienski, M. and White, J., "Model order reduction for nonlinear dynamical systems based on trajectory piecewise-linear approximations," *Linear Algebra and its Applications*, Vol. 415, No. 2-3, 2006, pp. 426–454.

[16]Ahmed, S. R., Ramm, G., and Faitin, G., "Some salient features of the time - averaged ground vehicle wake," *Society of Automotive Engineers*, Jan. 1984.