Proceedings of the Canadian Society for Mechanical Engineering International Congress
31st Annual Conference of the Computational Fluid Dynamics Society of Canada
CSME/CFD2024
May 26–29, 2024, Toronto, Ontario, Canada

NUMBER

Nonlinear model reduction for rapid and reliable CFD: application to transonic RANS flows

Alireza H. Razavi[1*], Masayuki Yano[2]

[1,2]Institute for Aerospace Studies, University of Toronto, Toronto, Canada
[*]a.razavi@mail.utoronto.ca

February 12, 2024

***Abstract*—**We introduce a registration-based nonlinear model-order reduction (MOR) method for efficient solution of parametrized transonic aerodynamics problems. Our goal is to provide rapid and reliable prediction of outputs, such as lift and drag, in many-query scenarios, which require the solution of partial differential equations (PDEs) for many different values of parameters, such as angle of attack and Mach number. Traditional "linear" MOR methods yield accurate low-dimensional approximations of many PDEs; however, for PDEs with shocks whose location depends on the parameter, linear MOR methods fail due to a fundamental limitation of linear approximation spaces known as the Kolmogorov barrier. To mitigate this issue, we develop a nonlinear reduced-order model (NLROM), which constructs a nonlinear reduced approximation through a two-step procedure. The method first uses a dilation-based shock sensor to locate discontinuities and optimally transforms/warps the solutions to align the discontinuities. The method then applies linear MOR techniques to this transformed problem: proper orthogonal decomposition, which finds a low-dimensional reduced basis; empirical quadrature procedure, which provides hyperreduction that enables rapid residual evaluation; and a hyperreduced dual-weighted residual method, which estimates the error in the quantity of interest. We apply the framework to parametrized two-dimensional transonic Reynolds-averaged Navier–Stokes flow over the RAE2822 airfoil and demonstrate that the NLROM enables efficient reduction for the problem where traditional linear ROMs are ineffective.

***Keywords-component*—model reduction, nonlinear approximation, error estimation, hyperreduction, transonic aerodynamics**

I. Introduction

The objective of model-order reduction (MOR) is to provide rapid and reliable solution of parametrized partial differential equations (PDEs). We wish to evaluate engineering quantities of interests, such as lift and drag, as a function of configuration parameters, such as flight condition parameters (e.g., Mach number and angle of attack) and geometric parameters (e.g., airfoil shape). Our goal is to accelerate the solution of many-query problems, such as database generation, uncertainity quantification, and optimization, that require PDE solution for potentially thousands of different configurations and the use of expensive high-fidelity simulation (e.g., finite element methods (FEMs)) can be cost prohibitive. The purpose of MOR is to provide rapid predictions with minimial loss in the accuracy and also to equip the predictions with quantitative error estimates. In this work, we present a nonlinear MOR method for transonic flow problems modelled by the Reynolds-averaged Navier-Stokes (RANS) equations.

A. Full-order model (FOM)

Mathematically, our exact PDE problem takes the following form: given parameters $\mu \in \mathcal{D}$, find the solution $u(\mu) \in \mathcal{V}$ in some function space $\mathcal{V}$, and then compute the associated output $s(\mu)$. One way to provide an accurate approximation of the problem is to use an FEM. FEMs approximate $u(\mu)$ in an N_h-dimensional subspace $\mathcal{V}_h$ of $\mathcal{V}$; i.e. $u(\mu) \approx u_h(\mu) \in \mathcal{V}_h \subset \mathcal{V}$ where $N_h = \mathcal{O}(10^5\text{–}10^7)$. While there are many ways to evaluate $u_h(\mu)$, we use a discontinuous Galerkin (DG) FEM, where $\mathcal{V}_h \subset \mathcal{V}$ is a discontinuous space of piecewise polynomials [6]. To estimate the output error and reduce computational cost, we equip the DG-FEM with dual-weighted residual (DWR) error estimation [1] and adaptive mesh refinement (AMR), respectively; the combination has been shown to be effective for many aerodynamics problems [4].

B. Linear MOR and its limitations

In aerodynamics, DG-FEM typically yields solutions with $N_h = \mathcal{O}(10^5\text{–}10^7)$ degrees of freedom, which can be too expensive in many-query scenarios. We hence wish to provide a low-dimensional and accurate approximation of

the DG-FEM full-order model (FOM). Linear MOR is a popular class of reduction methods [2], which approximates the FOM solution in a N-dimensional subspace $\mathcal{V}_N$ for $N = \mathcal{O}(10\text{–}100) \ll N_h$; i.e., $u_h(\mu) \approx u_N(\mu) \in \mathcal{V}_N$. Like many other surrogate models, MOR appeals to an offline-online computational decomposition: in the offline stage, a reduced-order model (ROM) is constructed through a relatively expensive training process; in the online stage, the ROM is evaluated rapidly for many parameter values, in real time, and/or using limited computing resources. Specifically, in the offline stage, we first compute training snapshots $U_h \coloneqq \{u_h(\mu_i)\}_{i=1}^{N_{\text{train}}}$ associated with N_{train} parameter values using the FOM. We next use these snapshots to construct an N-dimensional reduced basis (RB) using, e.g., proper orthogonal decomposition (POD), which provides an optimal hierarchical orthogonal basis $\{\phi_i\}_{i=1}^{N \leq N_{\text{train}}}$. We then define the RB space $\mathcal{V}_N \coloneqq \text{span}\{\phi_i\}_{i=1}^{N} \subset \text{span}\{u_h(\mu_i)\}_{i=1}^{N_{\text{train}}}$ where $N \ll N_h$. In the online stage, we use the (Petrov-)Galerkin projection onto the subspace to provide an RB approximation $u_h(\mu) \approx u_N(\mu) \in \mathcal{V}_N$. The MOR procedure reduces the size of the problem from $N_h = \mathcal{O}(10^5\text{–}10^7)$ to $N = \mathcal{O}(10^1\text{–}10^2)$, and hence ROMs are often orders of magnitude more efficient than FOMs.

In our ROMs, we incorporate two additional ingredients to provide output error estimates and rapid solution of nonlinear PDEs. First is the DWR output error estimate, which has been extended to MOR methods [19]. Second is a so-called hyperreduction method to accelerate the nonlinear residual evaluation; we employ the point-wise variant of the empirical quadrature procedure (EQP) [21, 3] for hyperreduction.

Linear MOR method are effective for subsonic aerodynamics problems and typically provide $\mathcal{O}(100\text{–}1000\times)$ speedup relative to the FOM while incurring less than $\mathcal{O}(1\%)$ error [20]. However, linear MOR methods do not yield accurate approximations for transonic flow problems where the location of a shock depends on the parameter. This limitation is known as the Kolmogorov barrier [14, 20]: the Kolmogorov N-width

$$\epsilon_N = \inf_{\substack{\mathcal{V}_N \subset \mathcal{V}_h \\ \dim(\mathcal{V}_N)=N}} \sup_{\mu \in \mathcal{D}} \inf_{w_N \in \mathcal{V}_N} \|u_h(\mu) - w_N\|_{L^2(\Omega)}, \qquad (1)$$

which is the best-fit error associated with the best *linear* approximation space, decays as only $\mathcal{O}(N^{-1/2})$ where N is the size of RB. That is, even the best linear approximation space will require a large basis to yield accurate approximations. To obtain a more efficient approximation space, we must construct nonlinear approximations.

C. Nonlinear model reduction

Two classes of nonlinear MOR methods that have become popular in the last half decade are based on neural networks and registration (or transformation). Neural-network based methods can produce accurate approximations when a large number of training snapshots are available, such as in unsteady problems [9, 8]. However, for steady transonic flow problems, these methods are primarily limited by the size of the required snapshots (i.e. they need large training sets), in addition to extensive tuning needed to obtain accurate models.

On the other hand, registration-based methods obtain nonlinear approximations by constructing linear approximations of spaces in which discontinuities are aligned [13, 10]. Generally, registration-based methods use the following ingredients: a domain transformation $g(\gamma) : \Omega \to \Omega$ for $\gamma \in \mathcal{D}_\gamma$; a function $\xi : \mathcal{D} \to \mathcal{D}_\gamma$ to compute optimal transformation parameters; and a sensor $\mathcal{S}(u_h(\mu))$ to detect discontinuity locations and the distance between them. These ingredients are used to warp the domain Ω such that discontinuities are aligned. The transformed space $\mathcal{V}_h^g$ then becomes amenable to linear approximation using transformed training snapshots $U_h^g \coloneqq \{u_h^g(\mu)\}_{i=1}^{N_{\text{train}}} \subset \mathcal{V}_h^g$.

Registration-based methods have been successfully applied to textbook shock-dominated problems, including two-dimensional inviscid transonic flows modeled by the Euler equations (e.g., [12, 18, 10]). However, to the best of our knowledge, the methods have yet to be demonstrated for larger-scale problems and for the Reynolds-averaged Navier–Stokes (RANS) equations, which exhibit more complex flow features than the Euler equations. This requires the following: a scalable/parallelizable transformation formulation; online-efficient error estimates (to provide reliable predictions); and efficient hyperreduction (to provide significant online speedup).

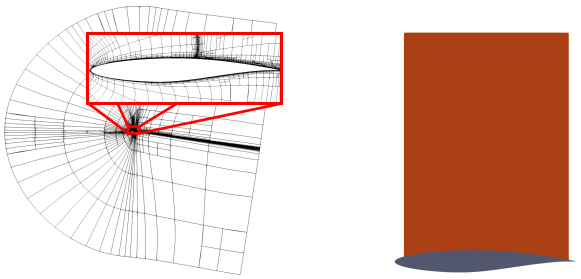
The contribution of this work is threefold. First, we develop a registration-based nonlinear ROM (NLROM) method that uses a dilation-based shock sensor and low-order statistics (rather than the $L^2(\Omega)$ norm of the solution field typically used) and is scalable/parallelizable. Second, we appeal to implicit transformation so that the method can readily incorporate online-efficient DWR error estimates and hyperreduction methods developed for linear MOR to provide rapid and reliable predictions. Third, we demonstrate the method for turbulent transonic flows over the RAE2822 airfoil modelled by the RANS equations.

II. Problem Statement

We first describe the mathematical problem. To this end, we introduce a parameter domain $\mathcal{D}$ and a physical domain $\Omega \subset \mathbb{R}^d$. We then state the system of m parametrized PDEs: given a parameter $\mu \in \mathcal{D}$, find $u(\mu) : \Omega \to \mathbb{R}^m$ such that

$$\begin{aligned} \nabla \cdot (F^c(u(\mu);\mu) - K(u(\mu);\mu)\nabla u(\mu)) \\ = S(u(\mu), \nabla u(\mu);\mu) \quad \text{in } \Omega, \\ B(u(\mu), \nabla u(\mu);\mu) = 0 \quad \text{on } \partial\Omega, \end{aligned} \qquad (2)$$

where $F^c : \mathbb{R}^m \times \mathcal{D} \to \mathbb{R}^{m \times d}$ is the convective flux, $K : \mathbb{R}^m \times \mathcal{D} \to \mathbb{R}^{m \times d \times m \times d}$ is the diffusion tensor, $S : \mathbb{R}^m \times \mathbb{R}^{m \times d} \times \mathcal{D} \to \mathbb{R}^m$ is a source term, and $B : \mathbb{R}^m \times \mathbb{R}^{m \times d} \times \mathcal{D} \to \mathbb{R}^m$ specifies the boundary conditions. We then evaluate the output $s(\mu) = J(u(\mu);\mu)$, where $J(\cdot;\mu)$ is some output functional which requires integration over the volume and boundaries. In aerodynamics problems, the parameter μ could describe the angle of attack, Mach number,



(a) Domain Ω (b) Subdomain ω

Figure. 1: Partitioning of domain Ω into ω and $\Omega \setminus \omega$ for the RAE2822 problem

and/or geometry, and the output $s(\mu)$ could be drag or lift, for example. Given the parameter space $\mathcal{D}$, (2) induces a parametric solution manifold $\mathcal{U}_\mathcal{D} := \{u(\mu)\}_{\mu \in \mathcal{D}}$. Our goal is to construct an NLROM for the manifold $\mathcal{U}_\mathcal{D}$.

III. Transformation framework

NLROM is composed of the following ingredients: parametrized geometry transformation, shock sensor, and a transformed PDE and FE model. In this section, we describe the transformation and sensor components, which are unique to our nonlinear approximation. We discuss the rest, which are common to linear MOR,—RB, output error estimation, and hyperreduction—in Sec. IV. A more detailed description of NLROM can be found in [15].

We wish to warp the domain Ω using geometry transformations such that discontinuities are aligned, which we detect using the shock sensor. We expect that when the discontinuities are aligned, the solutions will be amenable to approximation with linear MOR methods. The goal is to compute an aligned training snapshot set U_h^g given the untransformed snapshot set U_h. To this end, we align all the snapshots in U_h with a reference snapshot $u_h^{\text{ref}}(\mu^{\text{ref}}) \in U_h$, where μ^{ref} is chosen nearest to the centroid of $\mathcal{D}$.

A. Transformation

To define geometry transformations, we partition Ω into two subdomains: ω and $\Omega \setminus \omega$. The subdomain ω is chosen to encapsulate the shock for all parameter values. For example in transonic airfoil/wing problems, ω is the region on the top (or bottom) surface of the wing, as shown in Fig. 1.

Next, we introduce parametrized transformations $\mathcal{G} := \{g(\gamma) : \Omega \to \Omega\}$, which transform ω but do not transform $\Omega \setminus \omega$: i.e., $g(\gamma)|_\omega = g_\omega(\gamma)$ and $g(\gamma)|_{\Omega\setminus\omega} = \mathbf{I}$. We construct $g_\omega(\gamma)$ with sufficient regularity in the boundaries to avoid introducing a displacement, kink, or slip anywhere on the interface between ω and $\Omega\setminus\omega$, which would introduce instability and error in the DG-FEM and NLROM. However, we allow sliding on the regions of $\partial\omega$ which are on the boundary of the domain Ω; e.g., in Fig. 1, we allow transformations that "slide" the root of the shock along the surface.

We define the transformations $g_\omega(\gamma)$ as a composition of transfinite-interpolation (TFI) transformations [5] and parametrized splines. Composition of maps is adapted from Iollo and Taddei [7] and Taddei [17], and enables straightforward application of the method to complex geometries. The composite transformation is given by

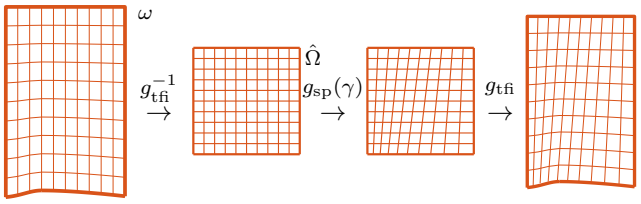


Figure. 2: Illustration of the composite transform $g_c(\gamma) = g_{\text{tfi}} \circ g_{\text{sp}}(\cdot;\gamma) \circ (g_{\text{tfi}})^{-1}$.

$$g(\cdot;\gamma) = g_{\text{tfi}} \circ g_{\text{sp}}(\cdot;\gamma) \circ g_{\text{tfi}}^{-1}. \tag{3}$$

An example of this transformation is illustrated in Fig. 2.

The first component of the composite transformation is a TFI map. TFI maps are invertible maps $g_{\text{tfi}} : \hat{\Omega} \to \omega$ which map the unit domain $\hat{\Omega} := (0,1)^d$ to the subdomain ω. In effect, $g_{\text{tfi}}(\cdot)$ is a bilinear interpolation from $\hat{\Omega}$ to ω.

The second component is a spline-based map $g_{\text{sp}}(\cdot;\gamma) : \hat{\Omega} \to \hat{\Omega}$. We define $g_{\text{sp}}(\cdot;\gamma)$ using a set of N_γ control points uniformly distributed in $\hat{\Omega}$ with coordinates $C = (c_1^{(j)}, c_2^{(j)})_{j=1}^{N_\gamma}$. We parametrize the transformation by the displacement at these control points, and interpolate between them using a sum of volumetric cubic B-splines $B_j(x)$ centered at $C^{(j)}$ with support over 5^d surrounding control points. We ensure that the spline-based maps meet the required continuity and regularity conditions at the interface between ω and $\Omega \setminus \omega$ by restricting the displacement of control points on the boundary $\partial\hat{\Omega}$ and their immediate neighbors. Note that the composition of TFI and spline-based maps allows us to avoid the direct construction of spline-based maps that meet the regularity conditions on arbitrary (and potentially complex) ω instead of the unit domain $\hat{\Omega}$.

B. Shock sensor and computing transformation parameters

We now describe how we compute optimal transformation parameters γ for the parametrized transformations $g(\cdot;\gamma)$. To this end, we use a truncated dilation-based shock indicator adapted from Moro et al. [11], given by

$$\mathcal{S}_{\text{dil}}(v) = \left(\frac{-L(\nabla \cdot v)}{c^*} \right)^+, \tag{4}$$

where c^* is the critical speed of sound, L is a tuned length scale, and $(z)^+ = \max\{z, 0\}$. We treat this shock indicator as a probability distribution, and compute the principal points $t(w) = \{t^{(i)}\}_{i=1}^{2d+1}$. An example of this sensor for the RAE2822 problem is illustrated in Fig. 3.

Our goal is to approximate the function $\xi : \mathcal{D} \to \mathcal{D}_\gamma$ which maps PDE parameter μ to transformation parameter γ: i.e.,

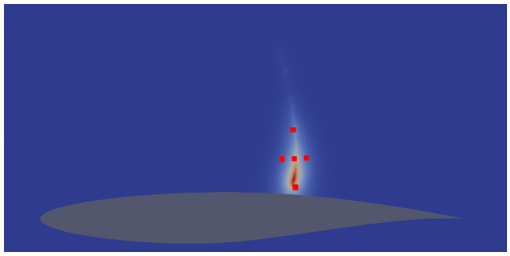


Figure. 3: Illustration of $\mathcal{S}_{\text{dil}}(u_h(\mu))$ and principal points for the RAE2822 problem at $M_\infty = 0.74$.

$\xi(\mu) = \gamma$. To this end, in the offline stage, we compute the optimal mapping $\mu_j \mapsto \gamma_j$ given by

$$\gamma_j = \underset{\gamma \in \mathcal{D}_\gamma}{\arg\min} \sum_{i=1}^{2d+1} |t^{(i)}(u_h^{\text{ref}}) - g^{-1}(t^{(i)}(u_h(\mu_j)); \gamma)| \quad (5)$$

for $j = 1, \ldots, N_{\text{train}}$. That is, we translate the principal points of the shock sensor (rather than the entire sensor field) to find the optimal transformation parameter; this significantly reduces the computational cost compared to aforementioned $L^2(\Omega)$-based sensors used in, e.g., [12, 18]. Then in the online stage, we use radial-basis functions (RBF) to interpolate the transformation parameters γ as a function of the PDE parameters μ: i.e., $\mu \mapsto \gamma \equiv \xi(\mu)$. We then use the optimal transformation: $g(\cdot; \gamma \equiv \xi(\mu))$.

IV. Transformed PDE and nonlinear ROM

Having outlined our procedure to construct an optimal parametrized domain transformation $g(\cdot; \gamma = \xi(\mu))$ in Sec. III, we now describe our procedure to construct NLROM that builds on the optimal transformation. To this end, we introduce a transformed system of PDEs in weak form: given PDE parameter $\mu \in \mathcal{D}$ and the optimal transformation parameter $\gamma \equiv \xi(\mu) \in \mathcal{D}_\gamma$, find $u^g(\mu, \gamma) \in \mathcal{V}$ such that

$$R^g(u^g(\mu,\gamma), v; \mu, \gamma) = \int_\Omega r^g_\Omega(u^g(\mu,\gamma), v; \mu, \gamma) dx + \int_{\partial\Omega} r^g_{\partial\Omega}(u^g(\mu,\gamma), v; \mu, \gamma) ds = 0 \quad \forall v \in \mathcal{V}, \quad (6)$$

where $r^g_\Omega(\cdot, \cdot; \mu)$ and $r^g_{\partial\Omega}(\cdot, \cdot; \mu)$ are transformed volume and boundary integrands, respectively, of the weak form of PDEs associated with (2). We then evaluate the output $s^g(\mu, \gamma) = J^g(u(\mu, \gamma); \mu, \gamma)$, where $J^g(\cdot; \mu, \gamma)$ is the transformed functional. The transformed terms incorporate a transformed gradient ∇_γ and transformation determinant $\det(Dg(\cdot; \gamma))$, which depend on the parametrized transformation $g(\cdot; \gamma)$. That is, the transformations are incorporated implicitly as a parametrization of the PDE, so that the solution to the original PDE (2) and the transformed PDE (6) are related by $u(\mu) = u^g(\mu, \gamma) \circ g(\cdot; \gamma)^{-1}$ and $s(\mu) = s^g(\mu, \gamma)$ for all $\mu \in \mathcal{D}$ and $\gamma \in \mathcal{D}_\gamma$. We hence solve (6) and compute a transformed output $s^g(\mu, \gamma)$.

We use a DG-FEM to find N_h-dimensional approximations of (6). Then, given a parameter space $\mathcal{D}$, a DG-FEM approximation of (6) yields a transformed parametric solution manifold $\mathcal{U}^g_{\mathcal{D},h} = \{u^g_h(\mu, \gamma \equiv \xi(\mu))\}_{\mu \in \mathcal{D}}$, where the shocks are aligned. As discussed in the Introduction, the parametric solution manifold associated with the transformed solutions, in which shocks are aligned, is ammenable to linear approximation and therefore is a way to mitigate the Kolmogorov barrier encountered by linear MOR methods on the original (untransformed) parametric solution manifold, in which the shocks are not aligned.

Having identified a transformed PDE amenable to linear compression, we now construct a ROM for this transformed PDE. To begin, we construct an N-dimensional RB approximation of this transformed manifold. We first introduce the transformed training snapshot set $U^g_h := \{u^g_h(\mu, \gamma \equiv \xi(\mu))\}_{i=1}^{N_{\text{train}}} \subset \mathcal{U}^g_{\mathcal{D},h}$. Unlike in linear ROM, where we compute the training snapshot set U_h using the DG-FEM on a single fixed/untransformed mesh, we compute U^g_h using a simultaneous adaptation and transformation strategy:

1) Compute untransformed ($\gamma = 0$) snapshots $U^g_{h,1}$ on a coarse mesh;
2) Compute optimal transformation parameters $\Gamma_1 = \{\gamma_i\}_{i=1}^{N_{\text{train}}}$ that (approximately) align discontinuities in all snapshots in $U^g_{h,1}$ using (5);
3) Perform output-error based AMR and compute more accurate snapshots $U^g_{h,2}$ on a finer mesh;
4) Update optimal transformation parameters Γ_2.

We repeat steps two to four for N_{iter} iterations until all snapshots meet the target error level of the DG-FEM method. This procedure also yields the optimal and online-efficient RBF-based transformation parameter function $\xi : \mu \mapsto \gamma$.

Once we have the transformed training snapshot set, the procedure to obtain an RB is identical to linear MOR. We apply POD to the transformed snapshots U^g_h to obtain an RB $\Phi := \{\phi_i\}_{i=1}^{N \le N_{\text{train}}}$ and construct an RB space $\mathcal{V}^g_N \subset \text{span}\{\phi_i\}_{i=1}^{N \le N_{\text{train}}}$. Since the discontinuities in the snapshot set U^g_h (and the manifold $\mathcal{U}^g_{\mathcal{D},h}$) are aligned, we expect POD to yield an RB which accurately approximates the manifold. Then we approximate DG-FEM solutions $u^g_h(\cdot; \mu, \gamma)$ by a linear combination of $\{\phi_i(\cdot)\}_{i=1}^N$ weighted by coefficients $\alpha(\mu, \gamma) \in \mathbb{R}^N$: i.e., $u^g_h(\cdot; \mu, \gamma) \approx u^g_N(\cdot; \mu, \gamma) = \sum_{i=1}^N \phi_i(\cdot)\alpha_i(\mu, \gamma)$. Finally, we apply Galerkin projection to find the coefficients $\alpha(\mu, \gamma)$ and hence the NLROM solution: given $\mu \in \mathcal{D}$ and $\gamma \equiv \xi(\mu) \in \mathcal{D}_\gamma$, find $u^g_N(\mu, \gamma) \in \mathcal{V}_h$ such that

$$R^g_h(u^g_N(\mu,\gamma), v_h; \mu, \gamma) = 0 \quad \forall v_N \in \mathcal{V}_N, \quad (7)$$

where $R^g_h(\cdot, \cdot; \mu, \gamma)$ is the DG-FEM approximation of (6).

We make several remarks on this NLROM. First, for $\gamma = 0$, (7) recovers linear ROM (for non-transformed PDEs). Indeed, even for $\gamma \neq 0$, although the PDE (6) is parametrized by μ and γ, due to the relation $\gamma \equiv \xi(\mu)$, NLROM is exactly a linear ROM for the μ-parametrized transformed PDE.

Second, note that although RB solutions $u^g_N(\cdot; \mu, \gamma)$ are represented by $\alpha(\mu, \gamma) \in \mathbb{R}^N$ coefficients, the evaluation of (7) remains expensive since we need to evaluate the integrals that appear in the DG-FEM approximation of (6) using piecewise Gauss(-like) quadrature rules with $Q_h = \mathcal{O}(N_h)$

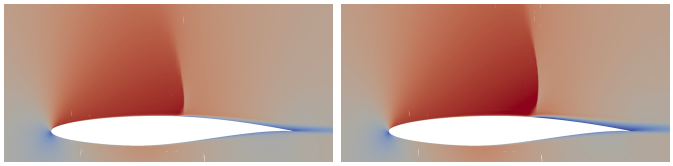


Figure. 4: RANS solutions of RAE2822 problem at $M_\infty = 0.73$ and $M_\infty = 0.75$. The field plotted is Mach number.

points. Fortunately, owing to the implicit incorporation of transformations in (6), we can readily employ hyperreduction methods to accelerate evaluation of the integrals. Namely, we use point-wise EQP hyperreduction [3] to accelerate evaluation of the residual form $R_h^g(\cdot,\cdot;\mu,\gamma)$. The procedure solves an optimization problem to find a small subset of the DG-FEM quadrature points to yield a sparse quadrature rule of the size $Q \ll Q_h$ that still provides an accurate evaluation of the DG-FEM residual. We remark that without the implicit transformation in (6), it would not be straightforward to extend EQP hyperreduction to (7).

Third, while the NLROM with hyperreduction yields rapid solutions, in order to have reliable solutions, we also need an output error estimate. Again, thanks to the implicit transformation in (6), we can readily incorporate the hyperreduced DWR-based output error estimate developed for linear MOR [19] to estimate $|s_h(\mu) - \tilde{s}_N(\mu)|$. The formulation provides an output error estimate for *any* parameter $\mu \in \mathcal{D}$, including those outside of the training set.

V. Application to transonic RANS flow

We assess the effectiveness of the NLROM using parametrized transonic turbulent flow over the RAE2822 airfoil. The governing equation is the RANS equations with the Spalart-Allmaras turbulence model [16], parametrized by the freestream Mach number M_∞ in the range $[0.73, 0.75]$. We fix angle of attack to $\alpha = 2.75°$ and Reynolds number to $\mathrm{Re}_c = 6.5 \times 10^6$. Despite the relatively narrow M_∞ range, the drag is sensitive to the input parameter and varies from $c_d = 169$ to 263 counts. Examples of the solutions are shown in Fig. 4.

We have introduced in Fig. 1 in Sec. III the computational domain Ω and transformation region ω. The transformation is defined by a composite transformation (3) associated with a grid of 5×5 uniformly distributed control points and the dilation-based shock sensor (4). We use 20 training snapshots from the untransformed manifold $\mathcal{U}_{\mathcal{D},h}$ to construct the linear ROM and 15 snapshots from the transformed manifold $\mathcal{U}^g_{\mathcal{D},h}$ to construct the NLROM. All snapshots are computed using the adaptive DG-FEM to achieve the 0.5% drag error level with respect to the (exact) PDE solutions.

Fig. 5 reports the maximum error of linear ROM and NLROM from randomly selected test parameters $\Xi_{\text{test}} = \{\mu_i\}_{i=1}^3$. We observe that the NLROM error converges to below 0.5% drag error using nine modes; the linear ROM error does not converge using 20 modes and fluctuates at the $\mathcal{O}(10\%)$ drag error level. In addition, we observe that the

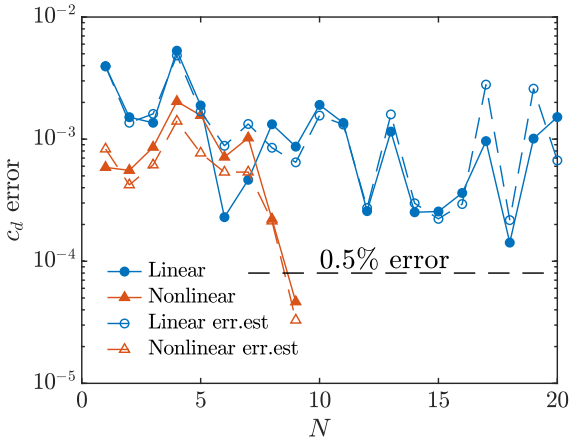


Figure. 5: Error convergence of linear ROM and NLROM for RANS-SA flow over the RAE2822 airfoil. Error is maximum over test set Ξ_{t}.

TABLE. I: Summary of linear ROM and NLROM for the RAE2822 RANS-SA problem. The times reported are in total core seconds.

Method	N_h N	Q_h Q	c_d error	t_{FEM} $\times 10^3$	$t_{\text{(NL)ROM}}$ $\times 10^{-2}$	Speedup
ROM	83820 20	137622 454	$> 10\%$	1.12	3.14	35500
NLROM	80220 9	131616 169	$< 0.5\%$	1.03	1.05	98000

DWR output error estimate is effective and closely predicts the true error for both approximations.

Table I summarizes the linear ROM and NLROM. First, we again observe that the linear ROM is unable to meet the target drag error of 0.5% despite using all $N = 20$ modes, while the NLROM meets the target using $N = 9$ modes. Second, we observe that the NLROM uses three times fewer reduced quadrature points than linear ROM ($Q = 169$ vs $Q = 454$); this is because the aligned shocks in the transformed space also makes the residual amenable to a sparser quadrature approximation. Third, the aligned snapshots also result in a slight reduction in the DG-FEM FOM size N_h; however, this effect is not as pronounced as observed in the Euler equations with the shock as the only dominant feature; see [15]. Finally, we observe that the three times sparser quadrature rule for the NLROM relative to the linear ROM yields a commensurate reduction in the online evaluation time; the NLROM achieves a significant online speedup of $\mathcal{O}(10^5)$ relative to the adaptive DG-FEM while controlling the drag error to $< 0.5\%$.

VI. Conclusion

We presented NLROM, a registration-based MOR framework, which builds on parametrized geometry transformation/warping, dilation-based shock sensor, transformation optimization, transformed PDE, RB, hyperreduction, and output error estimation. We applied the method to two-dimensional

transonic RANS flow over the RAE2822 airfoil and observed that NLROM with $N = 9$ modes provides rapid and reliable predictions that achieve 0.5% drag accuracy and are equipped with sharp error estimates; the linear ROM with $N = 20$ modes yields greater than 10% error.

Compared to existing registration-based methods, NLROM has three advantages, which leads to the threefold contributions of this work. First, NLROM uses the combination of shock sensor and principal points to enable scalable/parallizable computation of the optimal transform, which enables the application of NLROM to larger scale problems. Second, NLROM uses implicit transformation of the PDE, which enables a straightforward incorporation of DWR error estimates and EQP hyperreduction developed for linear ROMs. Third, NLROM has been demonstrated for transonic RANS flows, which exhibit much more complex flow features than the Euler equations.

References

[1] R. Becker and R. Rannacher. "A feed-back approach to error control in finite element methods: basic analysis and examples". In: *East-West J. Numer. Math* 4 (1996), pp. 237–264.

[2] P. Benner, S. Gugercin, and K. Willcox. "A survey of projection-based model reduction methods for parametric dynamical systems". In: *SIAM Review* 57.4 (2015), pp. 483–531. DOI: 10.1137/130932715.

[3] E. Du and M. Yano. "Efficient hyperreduction of high-order discontinuous Galerkin methods: Element-wise and point-wise reduced quadrature formulations". In: *Journal of Computational Physics* 466 (2022), p. 111399. DOI: 10.1016/j.jcp.2022.111399.

[4] K. Fidkowski and D. Darmofal. "Review of output-based error estimation and mesh adaptation in computational fluid dynamics". In: *AIAA Journal* 49.4 (2011), pp. 673–694. DOI: 10.2514/1.j050073.

[5] W. J. Gordon and C. A. Hall. "Construction of curvilinear co-ordinate systems and applications to mesh generation". In: *International Journal for Numerical Methods in Engineering* 7.4 (1973), pp. 461–477. DOI: 10.1002/nme.1620070405.

[6] J. S. Hesthaven and T. Warburton. *Nodal discontinuous Galerkin methods*. Springer New York, 2008. DOI: 10.1007/978-0-387-72067-8.

[7] A. Iollo and T. Taddei. "Mapping of coherent structures in parameterized flows by learning optimal transportation with Gaussian models". In: *Journal of Computational Physics* 471 (2022), p. 111671. DOI: 10.1016/j.jcp.2022.111671.

[8] Y. Kim, Y. Choi, D. Widemann, and T. Zohdi. "A fast and accurate physics-informed neural network reduced order model with shallow masked autoencoder". In: *Journal of Computational Physics* 451 (2022), p. 110841. DOI: 10.1016/j.jcp.2021.110841.

[9] K. Lee and K. T. Carlberg. "Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders". In: *Journal of Computational Physics* 404 (2020), p. 108973. ISSN: 0021-9991. DOI: 10.1016/j.jcp.2019.108973.

[10] M. A. Mirhoseini and M. J. Zahr. "Model reduction of convection-dominated partial differential equations via optimization-based implicit feature tracking". In: *Journal of Computational Physics* 473 (2023), p. 111739. DOI: 10.1016/j.jcp.2022.111739.

[11] D. Moro, N. C. Nguyen, and J. Peraire. "Dilation-based shock capturing for high-order methods". In: *International Journal for Numerical Methods in Fluids* 82.7 (2016), pp. 398–416. DOI: 10.1002/fld.4223.

[12] N. J. Nair and M. Balajewicz. "Transported snapshot model order reduction approach for parametric, steady-state fluid flows containing parameter-dependent shocks". In: *International Journal for Numerical Methods in Engineering* 117.12 (2019), pp. 1234–1262. DOI: 10.1002/nme.5998.

[13] M. Ohlberger and S. Rave. "Nonlinear reduced basis approximation of parameterized evolution equations via the method of freezing". In: *Comptes Rendus Mathematique* 351.23-24 (2013), pp. 901–906. DOI: 10.1016/j.crma.2013.10.028.

[14] M. Ohlberger and S. Rave. "Reduced basis methods: success, limitations and future challenges". In: *Proceedings of the Conference Algoritmy*. 2016, pp. 1–12.

[15] A. H. Razavi. "Adaptive Nonlinear Reduced-Order Models for Parametrized Partial Differential Equations of Transonic Flows". MASc thesis. Institute for Aerospace Studies: University of Toronto, 2024.

[16] P. R. Spalart and S. R. Allmaras. "A one-equation turbulence model for aerodynamics flows". In: *La Recherche Aérospatiale* 1 (1994), pp. 5–21.

[17] T. Taddei. "A registration method for model order reduction: data compression and geometry reduction". In: *SIAM Journal on Scientific Computing* 42.2 (2020), A997–A1027. DOI: 10.1137/19m1271270.

[18] T. Taddei and L. Zhang. "Registration-Based Model Reduction in Complex Two-Dimensional Geometries". In: *Journal of Scientific Computing* 88.3 (2021). ISSN: 1573-7691. DOI: 10.1007/s10915-021-01584-y.

[19] M. Yano. "Goal-oriented model reduction of parametrized nonlinear partial differential equations: Application to aerodynamics". In: *International Journal for Numerical Methods in Engineering* 121.23 (2020), pp. 5200–5226. DOI: 10.1002/nme.6395.

[20] M. Yano. "Model reduction in computational aerodynamics". In: *Model Order Reduction*. De Gruyter, 2020, pp. 201–236. DOI: 10.1515/9783110499001-006.

[21] M. Yano and A. T. Patera. "An LP empirical quadrature procedure for reduced basis treatment of parametrized nonlinear PDEs". In: *Computer Methods in Applied Mechanics and Engineering* 344 (2019), pp. 1104–1123. DOI: 10.1016/j.cma.2018.02.028.