Goal-oriented model reduction for parametrized time-dependent nonlinear partial differential equations

Michael K Sleeman[a,1], Masayuki Yano[a,2,*]

[a]*Institute for Aerospace Studies, University of Toronto, 4925 Dufferin Street, Toronto, M3H 5T6, ON, Canada*

Abstract

We present a projection-based model reduction formulation for parametrized time-dependent nonlinear partial differential equations (PDEs). Our approach builds on the following ingredients: reduced bases (RB), which provide rapidly convergent approximations of the parameter-temporal solution manifold; reduced quadrature (RQ) rules, which provide hyperreduction of the nonlinear residual; and the dual-weighted residual (DWR) method, which provides an error representation formula for the quantity of interest. To find the RQ rules, we develop an empirical quadrature procedure (EQP) for time-dependent problems; we analyze the output error due to hyperreduction using a space-time DWR framework and identify appropriate constraints so that the output error due to hyperreduction is controlled. We in addition equip our reduced model with an online-efficient DWR *a posteriori* error estimate for the output; we again analyze the error in the hyperreduced dual problem and DWR expression to find appropriate constraints for the EQP so that the error in the error estimate is controlled. In the offline stage, the RBs and RQs, as well as the finite element mesh, are simultaneously constructed using a POD-greedy algorithm that leverages the online-efficient output error estimate. We demonstrate the framework for parametrized unsteady flows in a lid-driven cavity and over a NACA0012 airfoil. Reduced models achieve over two orders of magnitude reduction in the number of degrees of freedom, number of quadrature points, and wall-clock computational time, while achieving less than 0.5% output error and providing efficient error estimates in predictive settings.

Keywords: reduced basis method, parametrized nonlinear PDEs, time-dependent PDEs, hyperreduction, empirical quadrature, compressible flow

1. Introduction

We consider rapid and reliable solution of parametrized time-dependent nonlinear partial differential equations (PDEs). Our goal is to evaluate quantities of interest (i.e., outputs), such as the time-averaged drag on an aerodynamic body, for various configurations (i.e., input parameters), such as the Mach number and Reynolds number. Our interest is in many-query scenarios, which require the input-output evaluation for many different configurations, and in real-time scenarios, which demand rapid output predictions. Our approach to this problem is projection-based model reduction with an efficient offline-online computational decomposition: in the offline stage, which is expensive but performed only once, we construct a reduced order model (ROM) using a systematic training algorithm; in the online stage, we invoke the ROM for many different parameter configurations to provide rapid output predictions, coupled with *a posteriori* error estimates to ensure reliability. This work focuses on PDEs that are parametrized, time-dependent, and nonlinear.

Model reduction techniques for time-dependent PDEs, or more generally dynamical systems, can be categorized by the class of problems considered: reproduction problems and parametrized problems. In a reproduction problem, we are given a "truth" solution $\{u(\mu,t)\}_{t\in[0,T]}$ associated with (a discrete approximation of) the PDE for some fixed configuration (i.e., parameter value) μ; we then construct a ROM that approximates the "truth" solution more rapidly for the same fixed configuration. There is a vast body of literature on projection-based model reduction of time-dependent PDEs in the context of reproduction problems; we refer to a review paper [10] for an overview of the existing methods. We note that arguably the most common approach is to use proper orthogonal decomposition (POD) to construct a reduced basis (RB) and then to apply (Petrov-)Galerkin projection to construct

*Corresponding author

Email addresses: msleeman@caltech.edu (Michael K Sleeman), myano@utias.utoronto.ca (Masayuki Yano)

[1]Graduate Student, Institute for Aerospace Studies, University of Toronto

[2]Assistant Professor, Institute for Aerospace Studies, University of Toronto

a ROM [50, 38, 39]; the approach has been successfully applied to large-scale simulations of incompressible flows (e.g., [54, 35]) and compressible flows (e.g., [13, 14]). These works consider reproduction problems, in which the "truth" solution is known.

The focus of the present work is parametrized problems. Unlike a reproduction problem that constructs a ROM for a single (parameter) configuration, in a parametrized problem we wish to construct a ROM that works for all parameter values μ in a parameter domain $\mathcal{D}$. We wish to use the ROM in a "predictive" setting and invoke it for parameter values for which the "truth" solution may not be available. This poses two additional challenges. First, to provide reliable predictions, the ROM must provide an *a posteriori* error estimate that works in the "predictive" setting. This is unlike the reproduction case, in which the quality of the ROM can be verified for the fixed configuration at the time of construction. Second, to enable efficient construction of a ROM, we must judiciously select the parameter values to evaluate the "truth" solutions used to construct the RB. One way to approach this problem is to use a greedy algorithm [53]: given a ROM that provides an online-efficient *a posteriori* error estimate, we evaluate the ROM solution and error estimate for many candidate parameter values, solve the "truth" problem at the worst-approximated parameter value, and successively update the ROM until the user-prescribed error tolerance is met over the entire $\mathcal{D}$.

We review here some of the seminal model reduction works on *a posteriori* error estimation and the greedy algorithm for time-dependent PDEs. (For reviews of the methods for steady PDEs, we refer to, e.g., [46, 40, 33].) Grepl and Patera [29] introduce a greedy framework for linear time-dependent PDEs. Haasdonk and Ohlberger [31] then introduce the POD-greedy framework, which combines greedy sampling over the parameter space with POD temporal compression at each parameter point. Drohmann et al. [20] and Grepl [27] extend the POD-greedy framework to nonlinear PDEs using (variants of) empirical interpolation method (EIM) [5] for hyperreduction. Wirtz et al. [55] apply a similar approach to dynamical systems. Haasdonk [30] provides theoretical analysis of the convergence rate of the POD-greedy algorithm. Urban and Patera [52] develop a space-time formulation that uses space-time reduced bases; the formulation has been extended to quadratically nonlinear problems [60, 56]. These works provide formulations and theoretical foundations to systematically construct ROMs with *a posteriori* error estimates for parametrized time-dependent (nonlinear) problems; however, the numerical demonstrations have been limited to relatively small-scale problems compared to the aforementioned work on reproduction problems (e.g., [54, 35, 13, 14]). One notable exception is the recent work of Fick et al. [24], which applies the POD-greedy framework to direct numerical simulation of turbulent flows; however, the work does not consider general (non-quadratic) nonlinearities or provide an *a posteriori* error estimate for the quantity of interest.

In this work, we extend the above work on model reduction for parametrized time-dependent nonlinear PDEs in three distinct ways: (i) we develop a versatile hyperreduction method that provides quantitative control of the output error due to hyperreduction for parametrized time-dependent problems with general nonlinearities; (ii) we develop an effective and online-efficient output *a posteriori* error estimate for the problems; and (iii) we develop a systematic ROM training procedure that controls the output error due to the finite element approximation, the reduced basis approximation, and hyperreduction. We demonstrate the framework for time-dependent problems with applications to fluids and aerodynamics. We discuss the three contributions in detail in the next three paragraphs

Our first contribution is the development of a hyperreduction method that provides quantitative control of output error due to hyperreduction for parametrized time-dependent PDEs with general (non-polynomial) nonlinearities. As briefly discussed, model reduction of PDEs with general nonlinearities requires hyperreduction to enable online-efficient evaluation of the nonlinear residual. In the context of time dependent problems, Grepl et al. [28] as well as the aforementioned works [20, 27, 55] employ the EIM, Carlberg et al. [13, 14] develop the Gauss-Newton with approximated tensors (GNAT) method based on the Gappy POD [21], and Farhat et al. [22, 23] develop the energy-conserving sampling and weighting (ECSW) method for second-order dynamical systems. In this work, we employ the empirical quadrature procedure (EQP) [42], which belongs to a family of quadrature-based hyperreduction methods that includes [47, 3, 22, 32]. EQP can control the hyperreduction error in either the norm of the solution [59] or the quantity of interest [58]. This feature distinguishes EQP from many other other hyperreduction methods, which control the error in the nonlinear residual operator, so that the hyperreduction tolerance must be "tuned" to achieve the desired error tolerance. To extend the goal-oriented EQP framework in [58] to time-dependent problems, we appeal to the space-time formalism and the dual-weighted residual (DWR) method [9, 8]. This combination allows us to analyze the output error due to hyperreduction and then identify an appropriate set of constraints to use in the EQP so that the hyperreduced ROM meets the user-prescribed output error tolerance. This eliminates the need for problem-specific and user-intensive tuning of hyperreduction parameters. Our formulation is versatile in the sense that it can be used with any quadrature-based spatial discretization and any time-marching scheme.

Our second contribution is the development of an online-efficient output *a posteriori* error estimate for parametrized time-dependent PDEs with general nonlinearities. Our formulation is based on the DWR framework [9, 8], which provides a general means to relate the residual to the output error using a dual (or adjoint) solution. This method is a popular choice for adaptive finite element approximation of fluid dynamics problems [25], including time-dependent problems [7, 49, 26]. The DWR method provides only an error estimate unlike the aforementioned bound formulations for linear PDEs (e.g., [29, 45, 31, 52]) and quadratically nonlinear PDEs (e.g., [37, 56]). However, the estimate applies to general nonlinearities in predictive settings and are effective in practice; see, e.g., [41, 11, 58] for previous applications of the DWR method in model reduction. In this work, we develop an online-efficient DWR error estimate for time-dependent PDEs with general nonlinearities. To this end, we approximate the dual solution in an RB space and then hyperreduce the dual problem and the DWR expression using EQPs. We present an approach to hyperreduce the dual problem in time-dependent contexts that is different from the original approach considered for steady problems in [58]. We appeal to the space-time formalism and the *dual-of-the-dual* problem to analyze the error in the error estimate due to hyperreduction and to identify appropriate constraints to use in the EQPs. In the dual-of-the-dual approach, we apply the DWR framework to the dual problem so that we can control the error in the DWR error estimate.

Building on the above two contributions to hyperreduction and *a posteriori* error estimation, we finally develop a POD-greedy algorithm that simultaneously constructs the finite element mesh, the RBs, and the reduced quadrature rules to meet the user-prescribed output error tolerance in a fully automated manner. Drohmann et al. [20] explores the idea of simultaneous RB and hyperreduction for time-dependent nonlinear PDEs. In this work, we in addition perform adaptive mesh refinement to control the finite element error using the DWR error estimate, an idea that has been explored for steady nonlinear PDEs in [58]; adaptive mesh refinement is arguably a necessity to reliably solve complex aerodynamics problems [25]. We demonstrate the automated training framework using parametrized unsteady flows in a lid-driven cavity and over the NACA0012 airfoil. The reduced models achieve over two orders of magnitude reduction in the number of degrees of freedom, the number of quadrature points, and the wall-clock computational time relative to the FE model, while controlling the relative output error to less than 0.5%; the model in addition provides effective output *a posteriori* error estimates.

The remainder of this paper is organized as follows. In Section 2, we introduce the parametrized time-dependent nonlinear problem considered throughout this work and develop EQPs that control the output error. In Section 3, we present our error estimation formulation, with an emphasis on EQPs to construct an online-efficient DWR error estimate. In Section 4, we present the simultaneous finite-element, RB, and hyperreduction training procedure. In Section 5, we demonstrate the framework for parametrized unsteady flows in a lid-driven cavity and over the NACA0012 airfoil.

2. Output prediction

2.1. Problem statement

We first present the output prediction problem that we consider throughout this work. To this end, we introduce a P-dimensional parameter space $\mathcal{D} \subset \mathbb{R}^P$, a d-dimensional spatial domain $\Omega \subset \mathbb{R}^d$, and a time interval $I \equiv (0, T]$, where $T \in \mathbb{R}_{>0}$ is the terminal time. We next introduce a vector-valued Hilbert space $\mathcal{V}$ on Ω. We consider the following weak form of parametrized time-dependent nonlinear PDEs: given $\mu \in \mathcal{D}$, find $u(\mu, t) \in \mathcal{V}$ such that

$$\begin{aligned} m(\frac{\partial u}{\partial t}(\mu, t), v; \mu) + r(u(\mu, t), v; \mu, t) = 0 \quad \forall v \in \mathcal{V},\ \forall t \in I, \\ m(u^0 - u(\mu, t = 0), v) = 0 \quad \forall v \in \mathcal{V}, \end{aligned} \tag{1}$$

where $m : \mathcal{V} \times \mathcal{V} \times \mathcal{D} \to \mathbb{R}$ is the bilinear form associated with the mass operator, $r : \mathcal{V} \times \mathcal{V} \times \mathcal{D} \times I \to \mathbb{R}$ is the semi-linear form associated with the spatial residual, and $u^0 \in \mathcal{V}$ is the initial condition. Given the solution $u(\mu, \cdot)$, we evaluate the output

$$s(\mu) = J(u(\mu, \cdot), \mu) \equiv \int_I f(u(\mu, t); \mu, t) dt + g(u(\mu, T); \mu),$$

where $f : \mathcal{V} \times \mathcal{D} \times I \to \mathbb{R}$ is the time-dependent output functional, and $g : \mathcal{V} \times \mathcal{D} \to \mathbb{R}$ is the terminal-time output functional. Our goal is to rapidly and accurately approximate the input-output map $\mu \mapsto u(\mu, \cdot) \mapsto s(\mu)$ in many-query and real-time scenarios; we demonstrate this framework in the many-query context in Section 5.

2.2. Finite element approximation

We now consider a finite element (FE) approximation of the problem (1). To begin, we introduce a piecewise polynomial approximation space $\mathcal{V}_h$ of dimension N_h. We next introduce an FE bilinear form $m_h : \mathcal{V}_h \times \mathcal{V}_h \times \mathcal{D} \to \mathbb{R}$ and an FE semi-linear form $r_h : \mathcal{V}_h \times \mathcal{V}_h \times \mathcal{D} \times I \to \mathbb{R}$, which arise from the approximation of the integrals in $m(\cdot,\cdot;\cdot)$ and $r(\cdot,\cdot;\cdot,\cdot)$, respectively, using a piecewise Gauss-like quadrature rule with $Q_h = O(N_h)$ points. The FE approximation of (1) is the following: given $\mu \in \mathcal{D}$, find $u_h(\mu,t) \in \mathcal{V}_h$ such that

$$\begin{aligned} \bar{r}_h(u_h(\mu,t), v_h; \mu, t) = 0 \quad \forall v_h \in \mathcal{V}_h, \quad \forall t \in I, \\ m_h(u^0 - u_h(\mu, t=0), v_h; \mu) = 0 \quad \forall v_h \in \mathcal{V}_h, \end{aligned} \tag{2}$$

where the semi-linear form $\bar{r}_h : \mathcal{V}_h \times \mathcal{V}_h \times \mathcal{D} \times I \to \mathbb{R}$ associated with the time-dependent residual is given by

$$\bar{r}_h(w_h, v_h; \mu, t) \equiv m_h(\frac{\partial w_h}{\partial t}, v_h; \mu) + r_h(w_h, v_h; \mu, t). \tag{3}$$

Given the FE approximation of the true solution, we evaluate the FE output

$$s_h(\mu) = J_h(u_h(\mu); \mu) \equiv \int_I f_h(u_h(\mu,t); \mu, t)dt + g_h(u_h(\mu, T); \mu), \tag{4}$$

where $f_h : \mathcal{V}_h \times \mathcal{D} \times I \to \mathbb{R}$ and $g_h : \mathcal{V}_h \times \mathcal{D} \to \mathbb{R}$ are the Q_h-point quadrature approximations of $f(\cdot;\cdot,\cdot)$ and $g(\cdot;\cdot)$, respectively.

We now introduce the quadrature-point-wise decomposition of the FE forms to prepare for our discussion of the model reduction formulation:

$$m_h(w, v; \mu) = \sum_{q=1}^{Q_h} \rho_q m_q(w, v; \mu) \quad \text{and} \quad r_h(w, v; \mu, t) = \sum_{q=1}^{Q_h} \rho_q r_q(w, v; \mu, t), \tag{5}$$

where Q_h is the total number of quadrature points inside the domain Ω and on the boundary $\partial\Omega$. For example, if $r(w, v; \mu, t) \equiv \int_\Omega \nabla v \cdot c(\mu, t) \nabla w dx$ and we approximate the integral using quadrature points $\{x_q\}_{q=1}^{Q_h}$ and the associated weights $\{\rho_q\}_{q=1}^{Q_h}$, then $r_h(w, v; \mu, t) \equiv \sum_{q=1}^{Q_h} \rho_q [\nabla v \cdot c(\mu, t) \nabla w]_{x_q}$ and hence $r_q(w, v; \mu, t) \equiv [\nabla v \cdot c(\mu, t) \nabla w]_{x_q}$. The time-dependent residual (3) inherits the quadrature-point-wise decomposition of $m_h(\cdot,\cdot;\cdot)$ and $r_h(\cdot,\cdot;\cdot,\cdot)$ so that

$$\bar{r}_h(w, v; \mu, t) = \sum_{q=1}^{Q_h} \rho_q \bar{r}_q(w, v; \mu, t),$$

where $\bar{r}_q(w, v; \mu, t) = \frac{d}{dt} m_q(w, v; \mu) + r_q(w, v; \mu, t)$. Similarly, the output functionals admit a quadrature-point-wise decomposition so that

$$f_h(w; \mu, t) = \sum_{q=1}^{Q_h} \rho_q f_q(w; \mu, t), \quad g_h(w; \mu) = \sum_{q=1}^{Q_h} \rho_q g_q(w; \mu), \quad \text{and} \quad J_h(w; \mu) = \sum_{q=1}^{Q_h} \rho_q J_q(w; \mu), \tag{6}$$

where $J_q(w; \mu) \equiv \int_I f_q(w(t); \mu, t)dt + g_q(w(T); \mu)$.

We conclude this section with a few remarks about the FE problem.

Remark 1. In the quadrature-point-wise decompositions (5) and (6), we assume that each decomposed form depends only on the parameter value, the time, the quadrature point location, and the function and gradient evaluated at the quadrature point. For instance, $r_q(w, v; \mu, t)$ depends only on the parameter value μ, current time t, the coordinates of the q-th quadrature point x_q, function values $w(x_q, t)$ and $v(x_q, t)$, and gradient values $\nabla w(x_q, t)$ and $\nabla v(x_q, t)$. This assumption is satisfied for most finite element methods.

Remark 2. In practice, to compute the FE solution, we apply a time-marching scheme to the semi-discrete residual (3) to obtain a fully discrete residual. We then approximate the temporal integral in the output functional (4) using the same time-marching scheme. However, as our model reduction formulation admits all multi-step and multi-stage time marching schemes, we will develop the formulation for the semi-discrete form of the equation. We refer to Appendix B for one realization of fully discrete approximations using diagonally implicit Runge-Kutta (DIRK) methods.

Remark 3. Following the application of a time-marching scheme, the FE system (2) is solved using a Newton-like solver. The cost to compute the solution can be decomposed into two parts: the cost for the residual and Jacobian evaluations, which scales linearly with the number of quadrature points Q_h, and the cost for linear solves, which scales superlinearly with the number of degrees of freedom N_h. Since the number of quadrature points Q_h is $O(N_h)$ for a typical FE method, the overall cost to compute $\mu \mapsto u_h(\mu,\cdot) \mapsto s_h(\mu)$ is $O(N_h^\bullet)$, where the exponent $^\bullet$ is between 1 and 3 depending on the linear solver.

2.3. Reduced-basis and reduced-quadrature approximation

We now introduce the reduced-basis (RB) approximation. We first introduce an N-dimensional RB approximation space $\mathcal{V}_N^{\text{pr}} \subset \mathcal{V}_h$ spanned by a reduced basis $\{\phi_i^{\text{pr}}\}_{i=1}^{N^{\text{pr}}}$, where formally $N^{\text{pr}} \leq N_h$ but in practice $N^{\text{pr}} \ll N_h$. The superscript "pr" indicates that the basis is associated with the primal problem, as opposed to the dual problem for the error estimate discussed in Section 3. We will discuss a systematic procedure to construct $\{\phi_i^{\text{pr}}\}_{i=1}^{N^{\text{pr}}}$ in Section 4.1; for now we assume an RB is given. Our RB approximation of (1) is as follows: given $\mu \in \mathcal{D}$, find $u_N(\mu, t) \in \mathcal{V}_N^{\text{pr}}$ such that

$$\begin{aligned} \overline{r}_h(u_N(\mu,t), v_N; \mu, t) = 0 \quad \forall v_N \in \mathcal{V}_N^{\text{pr}},\ \forall t \in I, \\ m_h(u^0 - u_N(\mu, t=0), v_N; \mu) = 0 \quad \forall v_N \in \mathcal{V}_N^{\text{pr}}, \end{aligned} \tag{7}$$

and evaluate the RB output

$$s_N(\mu) = J_h(u_N(\mu); \mu) = \int_I f_h(u_N(\mu,t); \mu, t) dt + g_h(u_N(\mu, T); \mu).$$

While the RB solution $u_N(\mu, t) \in \mathcal{V}_N^{\text{pr}}$ is approximated in the RB space of dimension $N^{\text{pr}} \ll N_h$, the evaluation of the FE residual (and the Jacobian) requires the application of the FE quadrature rule with $Q_h = O(N_h)$ points. Hence, the cost to compute $\mu \mapsto u_N(\mu, \cdot) \mapsto s_N(\mu)$ is still $O(N_h)$.

To evaluate the residual and the output with a cost independent of N_h, we apply hyperreduction to FE forms $\overline{r}_h(\cdot, \cdot; \cdot, \cdot)$ and $J_h(\cdot; \cdot)$. Namely, we approximate $\overline{r}_h(\cdot, \cdot; \cdot, \cdot)$ using a reduced quadrature (RQ) rule that uses $\tilde{Q}^r \ll Q_h$ points: i.e.,

$$\tilde{\overline{r}}_h(w, v; \mu, t) \equiv \sum_{\tilde{q}=1}^{\tilde{Q}^r} \tilde{\rho}_{\tilde{q}}^r \overline{r}_{q^r(\tilde{q})}(w, v; \mu, t), \tag{8}$$

where $q^r : \{1, \ldots, \tilde{Q}^r\} \to \{1, \ldots, Q_h\}$ is the mapping from the RQ indices to the FE quadrature indices, and $\{\tilde{\rho}_{\tilde{q}}^r\}_{\tilde{q}=1}^{\tilde{Q}^r}$ are the RQ weights. We indicate all quantities associated with RQ approximations with a tilde $(\tilde{\cdot})$, and the superscript "r" denotes that the RQ rule is associated with the (time-dependent) residual. We discuss an approach to find a RQ rule $\{q^r(\tilde{q}), \tilde{\rho}_{\tilde{q}}^r\}_{\tilde{q}=1}^{\tilde{Q}^r}$ in Section 2.5; for now we assume the indices and weights are given. Similarly, we approximate output functionals $f_h(\cdot; \cdot, \cdot)$ and $g_h(\cdot; \cdot)$ using an RQ rule that uses $\tilde{Q}^J \ll Q_h$ points; i.e.,

$$\tilde{f}_h(w; \mu, t) \equiv \sum_{\tilde{q}=1}^{\tilde{Q}^J} \tilde{\rho}^J f_{q^J(\tilde{q})}(w; \mu, t) \quad \text{and} \quad \tilde{g}_h(w; \mu) \equiv \sum_{\tilde{q}=1}^{\tilde{Q}^J} \tilde{\rho}^J g_{q^J(\tilde{q})}(w; \mu)$$

where $q^J : \{1 \ldots, \tilde{Q}^J\} \to \{1, \ldots, Q_h\}$ is the mapping from the RQ indices to the FE quadrature indices of the output functionals. We discuss an approach to find the RQ rule $\{q^J(\tilde{q}), \tilde{\rho}_{\tilde{q}}^J\}_{\tilde{q}=1}^{\tilde{Q}^J}$ in Section 2.5.

Given RB and RQ rules for the (time-dependent) residual and the output functionals, our RB-RQ approximation of (1) is as follows: given $\mu \in \mathcal{D}$, find $\tilde{u}_N(\mu, t) \in \mathcal{V}_N^{\text{pr}}$ such that

$$\begin{aligned} \tilde{\overline{r}}_h(\tilde{u}_N(\mu,t), v; \mu, t) = 0 \quad \forall v_N \in \mathcal{V}_N^{\text{pr}},\ \forall t \in I, \\ m_h(u^0 - u_N(\mu, t=0), v_N; \mu) = 0 \quad \forall v_N \in \mathcal{V}_N^{\text{pr}}, \end{aligned} \tag{9}$$

and evaluate the RB-RQ output

$$\tilde{s}_N(\mu) = \tilde{J}_h(\tilde{u}_N(\mu); \mu) \equiv \int_I \tilde{f}_h(u(\mu,t); \mu, t) dt + \tilde{g}_h(u(\mu, T); \mu). \tag{10}$$

Assuming $\tilde{Q}^r \ll Q_h$ and $\tilde{Q}^J \ll Q_h$ and in particular $O(\tilde{Q}^r) = O(\tilde{Q}^J) = O(N^{\text{pr}})$, the RB-RQ approximation permits the evaluation $\mu \mapsto \tilde{u}_N(\mu, \cdot) \mapsto \tilde{s}_N(\mu)$ in $O((N^{\text{pr}})^3)$ cost, which is independent of N_h and Q_h.

Remark 4. Similarly to Remark 2 on the solution of the FE problem, we apply a time-marching scheme to (7) and (9) to obtain a fully discrete system to find the RB and RB-RQ solutions, respectively. We refer to Appendix B.2 for fully discrete approximations of the RB problems using DIRK methods.

Remark 5. Some formulations of RB output include the so-called dual correction term so that the output is given by $s_N(\mu) = J_h(u_N(\mu); \mu) - \int_I \overline{r}_h(u_N(\mu,t), z_N^{\text{du}}(\mu,t); \mu, t) dt$, where $z_N^{\text{du}}(\mu)$ is an RB approximation of the dual solution computed in an RB space $\mathcal{V}_N^{\text{du}}$ that is different from $\mathcal{V}_N^{\text{pr}}$ [29]; the correction yields a more accurate RB output, where the output error is bounded by the product of the errors due to the approximation of the primal solution $u_h(\mu)$ in $\mathcal{V}_N^{\text{pr}}$ and the dual solution $z_h(\mu)$ in $\mathcal{V}_N^{\text{du}}$ (instead of in $\mathcal{V}_N^{\text{pr}}$). In this work, we do not include the dual correction term in the RB output and instead use the correction term to estimate the error in the output; see (25) and (26). This is analogous to the approach used for steady problems in [58].

2.4. Error in the RB approximation of output due to RQ

Our goal is to find RQ rules $\{q^r(\tilde{q}), \rho^r_{\tilde{q}}\}_{\tilde{q}=1}^{\tilde{Q}^r}$ and $\{q^J(\tilde{q}), \rho^J_{\tilde{q}}\}_{\tilde{q}=1}^{\tilde{Q}^J}$ that are (i) sparse (to facilitate rapid evaluation of the residual and the output functional), and (ii) accurate (to control the error in the output $\tilde{s}_N(\mu)$ due to the use of the RQ). To devise a systematic procedure to identify RQ rules with these properties, we first analyze the error due to the use of the RQ rules. To this end, we decompose the output error into two parts:

$$|s_N(\mu) - \tilde{s}_N(\mu)| = |J_h(u_N(\mu);\mu) - \tilde{J}_h(\tilde{u}_N(\mu);\mu)| \le \underbrace{|J_h(u_N(\mu);\mu) - J_h(\tilde{u}_N(\mu);\mu)|}_{\text{OE1}} + \underbrace{|J_h(\tilde{u}_N(\mu);\mu) - \tilde{J}_h(\tilde{u}_N(\mu);\mu)|}_{\text{OE2}}.$$

We see that "OE1" is associated with the error in the RB-RQ solution $\tilde{u}_N(\mu)$ given by (9), relative to the RB solution $u_N(\mu)$ given by (7), while "OE2" is associated with the approximation of $J(\cdot;\cdot)$ by $\tilde{J}(\cdot;\cdot)$. Here "OE" stands for output error. The analysis of OE2 is relatively straightforward, so we focus on OE1 for the remainder of this section.

We appeal to the space-time framework to facilitate the analysis of OE1, but we make one cautionary remark before we begin:

Remark 6. In this work, we appeal to the space-time framework to enable formal manipulations that are systematic, but without rigorous justifications. Given the present setting of general time-dependent nonlinear PDEs, we do not attempt to rigorously verify conditions under which these manipulations hold or under which the space-time problems are well-posed. We refer to, e.g., [52] for a rigorous space-time treatment of RB formulations for linear PDEs.

We first introduce a space-time space $\mathcal{Y}_N^{\text{pr}} = H^1(I; \mathcal{V}_N^{\text{pr}})$. We next introduce a space-time operator associated with (7): $A : \mathcal{Y}_N^{\text{pr}} \times \mathcal{D} \to \mathcal{Y}_N^{\text{pr}\prime}$ such that

$$\langle A(w;\mu), v\rangle \equiv \int_I m_h(\frac{\partial w}{\partial t}, v;\mu)dt + \int_I r_h(w, v;\mu, t)dt + m_h(u^0 - w(t=0), v(t=0);\mu) \quad \forall w, v \in \mathcal{Y}_N^{\text{pr}}. \tag{11}$$

We also recall the space-time output functional is $J : \mathcal{Y}_N^{\text{pr}} \times \mathcal{D} \to \mathbb{R}$ such that $J(w;\mu) \equiv \int_I f_h(w;\mu,t)dt + g_h(w(\mu,T);\mu)$. We in addition introduce the derivative $DA : \mathcal{Y}_N^{\text{pr}} \times \mathcal{D} \to \mathcal{L}(\mathcal{Y}_N^{\text{pr}}, \mathcal{Y}_N^{\text{pr}\prime})$ such that

$$\begin{aligned}\langle DA(y;\mu)w, v\rangle &= \int_I m_h(\frac{\partial w}{\partial t}, v;\mu)dt + \int_I r_h'(y; w, v;\mu, t)dt - m_h(w(t=0), v(t=0);\mu) \\ &= -\int_I m_h(w, \frac{\partial v}{\partial t};\mu)dt + \int_I r_h'(y; w, v;\mu, t)dt + m_h(w(t=T), v(t=T);\mu),\end{aligned} \tag{12}$$

and the derivative $DJ : \mathcal{Y}_N^{\text{pr}} \times \mathcal{D} \to \mathcal{Y}_N^{\text{pr}\prime}$ such that

$$\langle DJ(y;\mu), w\rangle = \int_I f_h'(y; w;\mu, t)dt + g_h'(y; w(t=T);\mu), \tag{13}$$

where $r_h'(y; w, v;\mu, t)$, $f_h'(y; w;\mu, t)$, and $g_h'(y; w;\mu)$ are the Gateaux derivatives of $r_h(\cdot, v;\mu, t)$, $f_h(\cdot;\mu, t)$, and $g_h(\cdot;\mu)$, respectively, about $y \in \mathcal{Y}_N^{\text{pr}}$ in the direction $w \in \mathcal{Y}_N^{\text{pr}}$.

In the space-time framework, we can concisely state the RB problem (7) as follows: given $\mu \in \mathcal{D}$, find $u_N(\mu) \in \mathcal{Y}_N^{\text{pr}}$ such that

$$\langle A(u_N(\mu);\mu), v_N\rangle = 0 \quad \forall v_N \in \mathcal{Y}_N^{\text{pr}},$$

and then evaluate the output $s_N(\mu) \equiv J(u_N(\mu);\mu)$. We also introduce the associated dual problem: given $\mu \in \mathcal{D}$ and $u_N(\mu) \in \mathcal{Y}_N^{\text{pr}}$, find $z_N^{\text{pr}}(\mu) \in \mathcal{Y}_N^{\text{pr}}$ such that

$$\langle DA(u_N(\mu);\mu)v_N, z_N^{\text{pr}}(\mu)\rangle = \langle DJ(u_N(\mu);\mu), v\rangle \quad \forall v_N \in \mathcal{Y}_N^{\text{pr}}.$$

(The superscript "pr" on $z_N^{\text{pr}}(\mu)$ indicates that the RB dual problem is solved in the space $\mathcal{V}_N^{\text{pr}}$. In Section 3, we will introduce a dual solution computed in a different space, which will bear a different superscript.) We may appeal to the expressions for the space-time operators (12) and (13) to obtain a more explicit form of the dual problem: given $\mu \in \mathcal{D}$ and $u_N(\mu) \in \mathcal{Y}_N^{\text{pr}}$, find $z_N^{\text{pr}}(\mu, t) \in \mathcal{V}_N^{\text{pr}}$ for all $t \in [0, T)$ such that

$$\begin{aligned}\overline{r}_h^{\text{du}}(u_N(\mu,t); w_N, z_N^{\text{pr}}(\mu,t);\mu,t) = 0 \quad \forall w_N \in \mathcal{V}_N^{\text{pr}},\ \forall t \in [0,T), \\ m(w_N, z_N^{\text{pr}}(\mu,T);\mu) = g_h'(u_N(\mu,T); w_N;\mu) \quad \forall w_N \in \mathcal{V}_N^{\text{pr}},\end{aligned} \tag{14}$$

where the time-dependent dual residual is given by

$$\overline{r}_h^{\mathrm{du}}(u; w, v; \mu, t) \equiv -m_h(w, \frac{\partial v}{\partial t}; \mu) + r_h'(u; w, v; \mu, t) - f_h'(u; w; \mu, t).$$

The dual problem (14) is solved backward in time starting with the terminal condition at time T. We refer to Appendix B.3 for fully discrete approximations of the dual problem using DIRK methods. The space-time form of the RB-RQ problem (9) and the associated dual problem can be written analogously by replacing $A(\cdot;\cdot)$ and $J(\cdot;\cdot)$ with their RQ approximations $\tilde{A}(\cdot;\cdot)$ and $\tilde{J}(\cdot;\cdot)$, respectively.

The following proposition summarizes OE1 due to the use of an RQ rule.

Proposition 7 (Output error due to the RQ approximation of $\overline{r}_h(\cdot,\cdot;\cdot,\cdot)$)**.** *Given $\mu \in \mathcal{D}$, let $u_N(\mu) \in \mathcal{V}_N^{\mathrm{pr}}$, $\tilde{u}_N(\mu) \in \mathcal{V}_N^{\mathrm{pr}}$, and $z_N^{\mathrm{pr}}(\mu) \in \mathcal{V}_N^{\mathrm{pr}}$ be the solutions to (7), (9), and (14), respectively, and $\hat{u}_N(\mu) \in \mathcal{V}_N^{\mathrm{pr}}$ be a surrogate state that may differ from both $u_N(\mu) \in \mathcal{V}_N^{\mathrm{pr}}$ and $\tilde{u}_N(\mu) \in \mathcal{V}_N^{\mathrm{pr}}$. Suppose $DA(\cdot;\mu) : \mathcal{V}_N^{\mathrm{pr}} \to \mathcal{L}(\mathcal{V}_N^{\mathrm{pr}\prime}, \mathcal{V}_N^{\mathrm{pr}\prime})$ and $D\tilde{A}(\cdot;\mu) : \mathcal{V}_N^{\mathrm{pr}} \to \mathcal{L}(\mathcal{V}_N^{\mathrm{pr}\prime}, \mathcal{V}_N^{\mathrm{pr}\prime})$ are continuously differentiable and bounded, and $DA^{-1}(\cdot;\mu) : \mathcal{V}_N^{\mathrm{pr}} \to \mathcal{L}(\mathcal{V}_N^{\mathrm{pr}\prime}, \mathcal{V}_N^{\mathrm{pr}\prime})$ and $D\tilde{A}^{-1}(\cdot;\mu) : \mathcal{V}_N^{\mathrm{pr}} \to \mathcal{L}(\mathcal{V}_N^{\mathrm{pr}\prime}, \mathcal{V}_N^{\mathrm{pr}\prime})$ are bounded. In addition, suppose*

$$|\int_I [\overline{r}_h(\hat{u}_N(\mu,t), z_N^{\mathrm{pr}}(\mu,t); \mu, t) - \tilde{\overline{r}}_h(\hat{u}_N(\mu,t), z_N^{\mathrm{pr}}(\mu,t); \mu, t))]dt| \le \delta^r, \tag{15}$$

$$\sup_{w \in \mathcal{V}_N^{\mathrm{pr}\prime}} \sup_{v \in \mathcal{V}_N'} \frac{\langle (I - DA(\hat{u}_N(\mu);\mu) D\tilde{A}(\hat{u}_N(\mu);\mu)^{-1}) w, v\rangle}{\|w\|_{\mathcal{V}_N^{\mathrm{pr}\prime}} \|v\|_{\mathcal{V}}} \le \delta^{DA} \tag{16}$$

for some $\delta^r \in \mathbb{R}_{>0}$ and $\delta^{DA} \in \mathbb{R}_{>0}$, where $\|w\|_{\mathcal{V}_N^{\mathrm{pr}\prime}} \equiv \sup_{v \in \mathcal{V}_N^{\mathrm{pr}}} \langle w, v\rangle / \|v\|_{\mathcal{V}}$ is the dual norm with respect to $\mathcal{V}_N^{\mathrm{pr}}$. Then

$$|J_h(u_N;\mu) - J_h(\tilde{u}_N;\mu)| \le \delta^r + O(\tilde{\delta}^2) + O((\delta^{DA})^2) + O(\hat{\delta}^2),$$

where $\tilde{\delta} \equiv \|u_N(\mu) - \tilde{u}_N(\mu)\|_{\mathcal{V}}$ and $\hat{\delta} \equiv \|u_N(\mu) - \hat{u}_N(\mu)\|_{\mathcal{V}}$.

Proof. See Appendix A. □

Proposition 7 shows that the error in the output $J(u_N(\mu);\mu) - J(\tilde{u}_N(\mu);\mu)$ due to the use of an RQ rule is bounded (i) primarily by the difference in the dual-weighted residual δ^r and (ii) secondarily by the error in the linearized operator δ^{DA}, the error in the norm of the solution $\tilde{\delta}$, and the error in the linearization point $\hat{\delta}$. We will now devise a systematic procedure to find an RQ rule informed by this output error estimate.

2.5. Empirical quadrature procedure

We recast the problem of finding the RQ rules $\{q^r(\tilde{q}), \tilde{\rho}_{\tilde{q}}^r\}_{\tilde{q}=1}^{\tilde{Q}^r}$ and $\{q^J(\tilde{q}), \tilde{\rho}_{\tilde{q}}^J\}_{\tilde{q}=1}^{\tilde{Q}^J}$ as constrained minimization problems. The general form of our approach to find $\{q^\bullet(\tilde{q}), \tilde{\rho}_{\tilde{q}}^\bullet\}_{\tilde{q}=1}^{\tilde{Q}^\bullet}$, where $\bullet$ may be "r" for the residual or "J" for the output functional, is the empirical quadrature procedure (EQP) [42].

Definition 8 (general form of the EQP to find $\{q^\bullet(\tilde{q}), \tilde{\rho}_{\tilde{q}}\}_{\tilde{q}=1}^{\tilde{Q}^\bullet}$)**.** Let $\mathrm{nnz}(\rho^\bullet)$ denote the number of nonzero elements in $\rho^\bullet \in \mathbb{R}^{Q_h}$. Given a training parameter set $\Xi_{\mathrm{eqp}} \subset \mathcal{D}$ of size N_{eqp}, find a sparse set of quadrature weights $\rho^{\bullet,\star} \in \mathbb{R}^{Q_h}$ such that

$$\rho^{\bullet,\star} = \arg\min_{\rho^\bullet \in \mathbb{R}^{Q_h}} \mathrm{nnz}(\rho^\bullet),$$

subject to Q_h non-negativity constraints

$$\rho_q^\bullet \ge 0, \quad q = 1, \dots, Q_h,$$

a constant-function accuracy constraint

$$|\int_\Omega dx - \sum_{q \in \{\text{vol. quad}\}} \rho_q^\bullet| < \delta^\bullet,$$

and $N_c \times N_{\mathrm{eqp}}$ manifold accuracy constraints

$$c(\rho^\bullet;\mu)_i \le \delta^\bullet, \quad i = 1, \dots, N_c, \quad \forall \mu \in \Xi_{\mathrm{eqp}},$$

where {vol. quad} is the set of indices of the FE quadrature points inside the domain Ω (as opposed to the boundary $\partial\Omega$), and $c : \mathbb{R}^{Q_h} \times \mathcal{D} \to \mathbb{R}^{N_c}$ are specific to the particular quadrature rule. Given a sparse set of weights $\{\rho_q^\bullet\}_{q=1}^{Q_h}$ that solves the optimization problem, the RQ mapping is $q^\bullet : \{1, \dots, \tilde{Q}^\bullet \equiv \mathrm{nnz}(\rho^\bullet)\} \to \{1, \dots, Q_h\}$ for the nonzero entries of $\rho^\bullet$, and the RQ weights are $\tilde{\rho}_{\tilde{q}}^\bullet \equiv \rho_{q^\bullet(\tilde{q})}^{\bullet,\star}$, $\tilde{q} = 1, \dots, \tilde{Q}^\bullet$. They collectively define an RQ rule $\{q^\bullet(\tilde{q}), \tilde{\rho}_{\tilde{q}}^\bullet\}_{\tilde{q}=1}^{\tilde{Q}^\bullet}$.

In other words, the EQP seeks a set of sparse quadrature weights $\{\rho_q^\bullet\}_{q=1}^{Q_h}$ that (i) contains the fewest number of nonzero points (i.e., $\min_\rho \cdot \text{nnz}(\rho^\bullet)$) that (ii) are non-negative, (iii) accurately integrate the constant function, and (iv) satisfy the manifold accuracy conditions that are specific to the RQ rule. The choice of the manifold accuracy constraints are crucial to control the error in the quantities of interest due to the RQ rule.

Remark 9. In practice, we do not solve the EQP problem (Definition 8) exactly, but approximate it as an ℓ^1 minimization problem, as done in [42, 59], or as a non-negative least squares (NNLS) problem, as in done in [22, 16] in the context of the ECSW method. In the case of NNLS, the standard ℓ^2 termination criterion is replaced by the EQP constraints (i.e., the solver iterates until the EQP constraints are satisfied). We have found that both methods reliably find sparse quadrature rules in practice. The parallel updatable NNLS method [16] is particularly well-suited for large-scale problems; the numerical example in Section 5.5 uses the algorithm on a problem that contains over 3×10^5 FE quadrature points and 10^3 constraints.

We now describe the manifold accuracy constraints for the residual RQ rule $\{q^r(\tilde{q}), \tilde{\rho}_{\tilde{q}}^r\}_{\tilde{q}=1}^{\tilde{Q}^r}$ that are designed to control OE1. Our constraints are based on the DWR condition (15) in Proposition 7, but we use a more conservative form of the constraints obtained by (i) splitting the time interval into K_{eqp} subintervals and (ii) modally decomposing the dual solution in the DWR into L pieces (see Remark 11). The resulting EQP statement is as follows:

Definition 10 (EQP for residual quadrature $\{q^r(\tilde{q}), \tilde{\rho}_{\tilde{q}}^r\}_{\tilde{q}=1}^{\tilde{Q}^r}$)**.** Given a training parameter set $\Xi_{\text{eqp}} \subset \mathcal{D}$ of size N_{eqp}, we introduce the associated training state set $\{\hat{u}_N(\mu) \in \mathcal{Y}_N^{\text{pr}}\}_{\mu\in\Xi_{\text{eqp}}}$ and training dual-state set $\{\hat{z}_N^{\text{pr}}(\mu) \in \mathcal{Y}_N^{\text{pr}}\}_{\mu\in\Xi_{\text{eqp}}}$. We in addition introduce $0 = t_0 < t_1 < \cdots < t_{K_{\text{eqp}}} = T$ and partition the time interval $I \equiv (0, T]$ into K_{eqp} subintervals $I_k = (t_{k-1}, t_k]$, $k = 1, \ldots, K_{\text{eqp}}$. Similarly, we introduce $1 = N_0 < N_1 < \cdots < N_L = N^{\text{pr}}$ and partition the RB function indices $\{1, \ldots, N^{\text{pr}}\}$ into L sets $S_l = \{N_{l-1}, \ldots, N_l\}$, $l = 1, \ldots, L$. We then apply the EQP (Definition 8) subject to manifold accuracy constraints

$$\left| \int_{I_k} [\overline{r}_h(\hat{u}_N(\mu, t), \Pi_{S_l}\hat{z}_N^{\text{pr}}(\mu, t); \mu, t) - \sum_{q=1}^{Q_h} \rho_q^r \overline{r}_q(\hat{u}_N(\mu, t), \Pi_{S_l}\hat{z}_N^{\text{pr}}(\mu, t); \mu, t)]dt \right| \le \frac{\delta^r}{L}\frac{|I_k|}{T} \tag{17}$$

for $k = 1, \ldots, K_{\text{eqp}}$, $l = 1, \ldots, L$, and $\mu \in \Xi_{\text{eqp}}$. Here, $|I_k| = t_k - t_{k-1}$, and $\Pi_{S_l} : \mathcal{V}_N^{\text{pr}} \to \text{span}\{\phi_i^{\text{pr}}\}_{i\in S_l}$ is the L^2-projection operator. The total number of manifold accuracy constraints is $K_{\text{eqp}} L N_{\text{eqp}}$.

Remark 11. In general, splitting the time interval I into smaller subintervals (i.e., larger K) results in a more conservative set of constraints since it prevents the positive and negative contributions to the DWR from cancelling with each other at different times. Similarly, splitting the RB functions into smaller sets results in more conservative constraints. The most conservative choice results from $L = N^{\text{pr}}$ so that the constraints are associated with $\overline{r}_h(\hat{u}_N(\mu, t), \mathbf{z}_i^{\text{pr}}\phi_i^{\text{pr}}; \mu, t)$ for $i = 1, \ldots, N^{\text{pr}}$, where $\mathbf{z}_i^{\text{pr}}$ is the i-th basis coefficient of $\hat{z}_N^{\text{pr}}(\mu, t)$; the least conservative choice results from $L = 1$ so that a single constraint is associated with $\overline{r}_h(\hat{u}_N(\mu, t), \hat{z}_N^{\text{pr}}(\mu, t) \equiv \sum_{i=1}^{N^{\text{pr}}} \mathbf{z}_i^{\text{pr}}\phi_i^{\text{pr}}; \mu, t)$. In theory, we could adaptively choose the size of each time interval $I_k = (t_{k-1}, t_k]$ and RB indices $S_l = \{N_{l-1}, \ldots, N_l\}$ so that, for instance, each constraint makes an equal contribution to the DWR statement. However, at least for the problems we have studied, we have found that the adaptive selection of the constraints does not make a significant difference in the resulting RQ rule. We hence use uniformly partitioned constraints in our numerical examples.

We observe that the EQP constraint (17) enforces the primary condition (15) in Proposition 7 that is required to control the output error (OE1) in the sense that

$$\begin{aligned} &\left| \int_I (\overline{r}_h(\hat{u}_N(\mu, t), z_N^{\text{pr}}(\mu, t); \mu, t) - \tilde{\overline{r}}_h(\hat{u}_N(\mu, t), z_N^{\text{pr}}(\mu, t); \mu, t)))dt \right| \\ &\le \sum_{k=1}^{K_{\text{eqp}}} \sum_{l=1}^{L} \left| \int_{I_k} [\overline{r}_h(\hat{u}_N(\mu, t), \Pi_{S_l}\hat{z}_N^{\text{pr}}(\mu, t); \mu, t) - \sum_{q=1}^{Q_h} \rho_q^r \overline{r}_q(\hat{u}_N(\mu, t), \Pi_{S_l}\hat{z}_N^{\text{pr}}(\mu, t); \mu, t)]dt \right| \le \sum_{k=1}^{K_{\text{eqp}}} \sum_{l=1}^{L} \frac{\delta^r}{L}\frac{|I_k|}{T} = \delta^r. \end{aligned} \tag{18}$$

On the other hand, we do not enforce the secondary condition (16) on the linearized operator $D\tilde{A}$ explicitly. We hope that enforcing the residual constraint (17) (and hence (15) more conservatively) indirectly controls the error in the linearized operator $D\tilde{A}$ and hence results in the satisfaction of (16). We also note that (16) is associated with the second-order error, as opposed to the first-order error. An analogous simplification was made in the context of steady-state problems in [58] and used successfully in practice. In Section 5, we will demonstrate using our numerical examples that the simplified constraints also work well for time-dependent problems.

Remark 12. Similarly to Remark 2 on the solution of the FE problem (2), in practice we apply the same time-marching scheme used for the FE problem to the semi-discrete residual $\bar{r}_h(\cdot,\cdot;\cdot,\cdot)$ in the EQP constraint (17) to obtain a fully discrete residual. We then create K_{eqp} subintervals that each contain multiple time steps of the time-marching scheme. We refer to Appendix B.6 for fully discrete approximations of the EQP constraints for DIRK methods.

We now describe the manifold accuracy constraints for the output functional RQ rule $\{q^J(\tilde{q}),\tilde{\rho}_{\tilde{q}}^J\}_{\tilde{q}=1}^{\tilde{Q}^J}$ that are designed to control OE2. The EQP statement is as follows:

Definition 13 (EQP for output functional quadrature $\{q^J(\tilde{q}),\tilde{\rho}_{\tilde{q}}^J\}_{\tilde{q}=1}^{\tilde{Q}^J}$)**.** Given a training parameter set $\Xi_{\text{eqp}} \subset \mathcal{D}$ of size N_{eqp}, we introduce a training parameter set Ξ_{eqp} and the associated RB-RQ solutions $\{\tilde{u}_N(\mu) \in \mathcal{Y}_N^{\text{pr}}\}_{\mu\in\Xi_{\text{eqp}}}$ given by (9). We in addition introduce $0 = t_0 < t_1 < \cdots < t_{K_{\text{eqp}}} = T$ and partition the time interval $I \equiv (0,T]$ into K_{eqp} subintervals $I_k = (t_{k-1},t_k]$, $k = 1,\ldots,K_{\text{eqp}}$. We then apply the EQP (Definition 8) subject to the manifold accuracy constraints

$$\begin{aligned} &|\int_{I_k}[f_h(\tilde{u}_N(\mu,t);\mu,t) - \sum_{q=1}^{Q_h}\rho_q^J f_q(\tilde{u}_N(\mu,t);\mu,t)]dt| \le \delta^J\frac{|I_k|}{T}, \quad k = 1,\ldots,K_{\text{eqp}}-1, \\ &|\int_{I_{K_{\text{eqp}}}} f_h(\tilde{u}_N(\mu,t);\mu,t)dt + g_h(\tilde{u}_N(\mu,T);\mu) - \sum_{q=1}^{Q_h}\rho_q^J[\int_{I_{K_{\text{eqp}}}} f_q(\tilde{u}_N(\mu,t);\mu,t)dt + g_q(\tilde{u}_N(\mu,t);\mu)]| \le \delta^J\frac{|I_{K_{\text{eqp}}}|}{T}, \end{aligned} \tag{19}$$

for $\mu \in \Xi_{\text{eqp}}$. The total number of constraints is $K_{\text{eqp}}N_{\text{eqp}}$.

The first $K_{\text{eqp}} - 1$ constraints of (19) enforce the output functional accuracy constraint for all time intervals except the last one, and the last constraint enforces the output functional accuracy constraint over the last time interval $I_{K_{\text{eqp}}}$ and at the terminal time. We include the terminal-time condition in the final interval constraint so that their treatment is consistent with the dual problem constraints introduced in Section 3. We observe that the EQP constraints (19) control error (OE2) in the sense that

$$\begin{aligned} |J_h(\tilde{u}_N(\mu);\mu) - \tilde{J}_h(\tilde{u}_N(\mu);\mu)| &= |\sum_{k=1}^{K_{\text{eqp}}}\int_{I_k}(f_h(\tilde{u}_N(\mu,t);\mu,t) - \tilde{f}_h(\tilde{u}_N(\mu,t);\mu,t))dt + g_h(\tilde{u}_N(\mu,T);\mu) - \tilde{g}_h(\tilde{u}_N(\mu,T);\mu)| \\ &\le \sum_{k=1}^{K_{\text{eqp}}}\delta^J\frac{|I_k|}{T} = \delta^J. \end{aligned}$$

Hence we provide direct control of OE2.

It follows that the EQP for the residual (Definition 10) and the EQP for output functional (Definition 13) together control the output error due to the use of RQ rules in the sense that

$$|s_N(\mu) - \tilde{s}_N(\mu)| \le \text{(OE1)} + \text{(OE2)} \le \delta^r + \delta^J + \text{(higher-order terms)}, \quad \forall\mu \in \Xi_{\text{eqp}},$$

where the higher-order terms are identified in Proposition 7.

Remark 14. In principle, we can use a sufficiently rich training set $\Xi_{\text{eqp}} \subset \mathcal{D}$ to control the RQ errors everywhere in $\mathcal{D}$. This is the approach used for steady problems in [59, 58]. However, imposing the constraints over a rich training set Ξ_{eqp} can be prohibitively expensive for time-dependent problems. As such, we employ a relatively small Ξ_{eqp} in practice, as discussed in Section 4. Nevertheless, numerical results in Section 5 show that the EQP controls the output error well in practice.

Remark 15. In this work, we employ a (space-only) Galerkin formulation and standard multi-stage time integrators. The associated RB problem (7) is not guaranteed to be well-posed, even if the underlying FE problem (2) is well-posed; as noted in Remark 6, we *assume* our RB problem (7) is well-posed. If we used (a) a time-integrator that is amenable to variational interpretation with (b) a space-time Petrov-Galerkin formulation, then it would be possible to guarantee well-posedness of the RB formulation for well-posed linear PDEs [52] and for certain nonlinear PDEs [56], as the space-time inf-sup constant for the RB formulation would be bounded. However, in this work, we do not explore the space-time Petrov-Galerkin formulation.

3. Output *a posteriori* error estimation

In this section we present an approach to construct output *a posteriori* error estimates for the FE problem (2) and for the RB-RQ problem (9). Our approach is based on the DWR method [9].

3.1. Estimation of FE error

We first describe our approach to estimate the FE output error $|s(\mu) - s_h(\mu)|$ associated with (2). Our approach is based on the DWR method [9], and in particular its application to time-dependent problems [7, 49, 26]. To this end, we first introduce an "enriched" FE approximation space $\mathcal{V}_{\hat{h}} \supset \mathcal{V}_h$. In our numerical examples in Section 5, we enrich the space by increasing the polynomial degree of the FE approximation space by one. We then introduce the FE dual problem in the enriched space: given $\mu \in \mathcal{D}$ and the associated FE solution $u_h(\mu) \in \mathcal{V}_N^{\mathrm{pr}}$, find $z_{\hat{h}}(\mu, t) \in \mathcal{V}_{\hat{h}}$ such that

$$\begin{aligned} \overline{r}_{\hat{h}}^{\mathrm{du}}(u_h(\mu,t); w_{\hat{h}}, z_{\hat{h}}(\mu,t); \mu, t) = 0 \quad \forall w_{\hat{h}} \in \mathcal{V}_{\hat{h}},\ \forall t \in [0,T), \\ m(w_{\hat{h}}, z_{\hat{h}}(\mu,T); \mu) = g'_{\hat{h}}(u_h(\mu,T); w_{\hat{h}}; \mu) \quad \forall w_{\hat{h}} \in \mathcal{V}_{\hat{h}}, \end{aligned} \tag{20}$$

where we recall that the time-dependent dual residual is given by

$$\overline{r}_{\hat{h}}^{\mathrm{du}}(u; w, z; \mu, t) \equiv -m_{\hat{h}}(w, \frac{dz}{dt}; \mu) + r'_{\hat{h}}(u; w, z; \mu, t) - f'_{\hat{h}}(u; w; \mu, t). \tag{21}$$

We then evaluate the DWR error estimate

$$|s(\mu) - \tilde{s}_h(\mu)| \approx \eta_h^{\mathrm{fe}}(\mu) \equiv |\int_I \overline{r}_h(\tilde{u}_h(\mu), z_{\hat{h}}(\mu); \mu, t) dt|. \tag{22}$$

The accuracy of the approximation ($\approx$) depends on the effect of neglecting the mean-value linearization, and on the ability of the enriched FE approximation space $\mathcal{V}_{\hat{h}}$ to represent the true dual solution; see [9] for a detailed analysis of the DWR error estimate. The dual problem (20) is solved backward in time starting from the terminal condition at time T. Expression (22) follows from the DWR error representation formula [9]. The DWR formulation provides (merely) an error estimate and is not guaranteed to be reliable. However, in practice, the estimate has been used successfully in many complex nonlinear problems, including those in aerodynamics [25]. We will continue to present our formulation in the semi-discrete form throughout this section, as discussed in Remark 2. We refer to Appendix B for fully discrete approximations of the dual problem and the DWR error estimate using DIRK methods.

3.2. Reduced-basis and reduced-quadrature approximation of DWR

We now describe our approach to construct an online-efficient *a posteriori* error estimate for the error $|s_h(\mu) - \tilde{s}_N(\mu)|$ based on the DWR method [9] and its RB-RQ approximation. To begin, we first introduce the FE dual problem (in the non-enriched space): given $\mu \in \mathcal{D}$ and the associated RB-RQ solution $\tilde{u}_N(\mu) \in \mathcal{V}_N^{\mathrm{pr}}$, find $z_h(\mu, t) \in \mathcal{V}_h$ for all $t \in [0, T)$ such that

$$\begin{aligned} \overline{r}_h^{\mathrm{du}}(\tilde{u}_N(\mu,t); w_h, z_h(\mu,t); \mu, t) = 0 \quad \forall w_h \in \mathcal{V}_h,\ \forall t \in [0,T), \\ m(w_h, z_h(\mu,T); \mu) = g'_h(\tilde{u}_N(\mu,T); w_h; \mu) \quad \forall w_h \in \mathcal{V}_h, \end{aligned} \tag{23}$$

where $\overline{r}_h^{\mathrm{du}}(\cdot;\cdot,\cdot;\cdot,\cdot)$ is given by (21), and then evaluate the DWR error estimate

$$|s_h(\mu) - \tilde{s}_N(\mu)| \approx \eta_h(\mu) \equiv |\int_I \overline{r}_h(\tilde{u}_N(\mu), z_h(\mu); \mu, t) dt|. \tag{24}$$

The accuracy of the approximation ($\approx$) depends on the effect of neglecting the mean-value linearization; see [9]. Note that $\eta_h^{\mathrm{fe}}(\mu)$ given by (22) approximates the error $|s(\mu) - s_h(\mu)|$ between the true output and the FE approximation to the output, while $\eta_h(\mu)$ given by (24) approximates the error $|s(\mu) - \tilde{s}_N(\mu)|$ between the true output and the RB-RQ approximation to the output. The dual residual (21) inherits quadrature-point-wise decompositions of $m_h(\cdot,\cdot;\cdot)$ and $r_h(\cdot,\cdot;\cdot,\cdot)$ in (5) and $f_h(\cdot;\cdot,\cdot)$ in (6) so that

$$\overline{r}_h^{\mathrm{du}}(u; w, z; \mu, t) = \sum_{q=1}^{Q_h} \rho_q \overline{r}_q^{\mathrm{du}}(u; w, z; \mu, t).$$

From the triangle inequality, we have

$$|s(\mu) - \tilde{s}_N(\mu)| \le |s(\mu) - s_h(\mu)| + |s_h(\mu) - \tilde{s}_N(\mu)| \approx \eta_h^{\mathrm{fe}}(\mu) + \eta_h(\mu).$$

In practice, we set the FE error tolerance to be an order of magnitude lower than the target error, so that the RB-RQ error $|s_h(\mu) - \tilde{s}_N(\mu)|$ dominates the overall error (see Remark 25). In addition, as the DWR error estimate has been

used successfully in many complex nonlinear problems [25], we *assume* that the accuracy of $\eta_h(\mu)$ (the DWR error estimate based on the FE dual solution) is satisfactory. Hence our goal is to enable an online-efficient evaluation of $\eta_h(\mu) \approx |s_h(\mu) - \tilde{s}_N(\mu)| \approx |s(\mu) - \tilde{s}_N(\mu)|$.

To enable an online-efficient approximation of the DWR error estimate, we employ RB and RQ approximation. To this end, we first introduce an N^{du}-dimensional RB approximation space for the dual problem $\mathcal{V}_N^{\mathrm{du}} \in \mathcal{V}_h$ spanned by a reduced basis $\{\phi_i^{\mathrm{du}}\}_{i=1}^{N^{\mathrm{du}}}$ for $N^{\mathrm{du}} \ll N_h$. We defer the discussion of the procedure to construct $\{\phi_i^{\mathrm{du}}\}_{i=1}^{N^{\mathrm{du}}}$ to Section 4.1 and here assume the basis is given. Our RB approximation of the DWR problem is the following: given $\mu \in \mathcal{D}$ and $\tilde{u}_N(\mu) \in \mathcal{V}_N^{\mathrm{pr}}$, find $z_N^{\mathrm{du}}(\mu, t) \in \mathcal{V}_N^{\mathrm{du}}$ such that

$$\begin{aligned} \overline{r}_h^{\mathrm{du}}(u_N(\mu, t); w_N, z_N^{\mathrm{du}}(\mu, t); \mu, t) = 0 \quad \forall w_N \in \mathcal{V}_N^{\mathrm{du}},\ \forall t \in [0, T), \\ m(w_N, z_N^{\mathrm{du}}(\mu, T); \mu) = g_h'(u_N(\mu, T); w_N; \mu) \quad \forall w_N \in \mathcal{V}_N^{\mathrm{du}}. \end{aligned} \tag{25}$$

and then evaluate the error estimate

$$|s_h(\mu) - s_N(\mu)| \approx \eta_N(\mu) \equiv |\int_I \overline{r}_h(\tilde{u}_N(\mu, t); z_N^{\mathrm{du}}(\mu, t); \mu, t) dt|. \tag{26}$$

While the RB dual-solution $z_N^{\mathrm{du}}(\mu, t) \in \mathcal{V}_N^{\mathrm{du}}$ is approximated in a space of dimension $N^{\mathrm{du}} \ll N_h$, the evaluation of the DWR estimate requires $O(Q_h) = O(N_h)$ operations because the solution of the dual problem (25) and the evaluation of the DWR expression (26) require the evaluation of the FE residual (21) and the associated Jacobian. Before we introduce the RQ rules needed for hyperreduction, we make a remark on the accuracy of the DWR error estimate:

Remark 16. The RB-approximated DWR error estimate (26) is different from the true error $|s_h(\mu) - s_N(\mu)|$ for two reasons: (i) the adjoint problem (25) is linearized about $\tilde{u}_N$ instead of the mean-value-linearized adjoint problem (see [9]), and (ii) we solve for $z_N^{\mathrm{du}}(\mu, t)$ in $\mathcal{V}_N^{\mathrm{du}}$ instead of $\mathcal{V}_h$. The error due to (i) scales with $\|u_h - \tilde{u}_N\|_{\mathcal{Y}}^2$, so it is second-order in the primal error. This error may be non-negligible for a coarse RB approximation, but if we assume that the primal error decreases rapidly with increasing N, then so does the error due to (i). The error due to (ii) can be significant for certain problems, depending on the choice of $\mathcal{V}_N^{\mathrm{du}}$. For instance, if $\mathcal{V}_N^{\mathrm{du}} \subset \mathcal{V}_N^{\mathrm{pr}}$ so that $z_N^{\mathrm{du}}(\mu, t) \in \mathcal{V}_N^{\mathrm{pr}}$, then the error estimate would be zero by Galerkin orthogonality. We must carefully choose $\mathcal{V}_N^{\mathrm{du}}$ to ensure that the error estimate is effective. Specifically, the space must be chosen so that $z(\mu, t) - \Pi_{\mathcal{V}_N^{\mathrm{pr}}} z(\mu, t)$ is well approximated in $\mathcal{V}_N^{\mathrm{du}}$. In Section 4.1, we present a method to adaptively construct the dual space $\mathcal{V}_N^{\mathrm{du}}$ to meet the desired error estimate effectivity (for the training parameter values).

We introduce RQ rules to enable rapid evaluation of the FE residual. To this end, we first introduce an RQ rule $\{q^{\eta,1}(\tilde{q}), \tilde{\rho}_{\tilde{q}}^{\eta,1}\}_{\tilde{q}=1}^{\tilde{Q}^{\eta,1}}$ for the dual residual that governs the dual problem (25) (i.e., to solve the dual problem) so that $\overline{r}_h^{\mathrm{du}}(u; w, v; \mu, t)$ is approximated by

$$\tilde{\overline{r}}_h^{\mathrm{du},\eta,1}(u; w, v; \mu, t) \equiv \sum_{\tilde{q}=1}^{\tilde{Q}^{\eta,1}} \tilde{\rho}_{\tilde{q}}^{\eta,1} \overline{r}_{q^{\eta,1}(\tilde{q})}^{\mathrm{du}}(u; w, v; \mu, t).$$

Similarly, we introduce an RQ rule $\{q^{\eta,2}(\tilde{q}), \tilde{\rho}_{\tilde{q}}^{\eta,2}\}_{\tilde{q}=1}^{\tilde{Q}^{\eta,2}}$ for the primal residual for the DWR (26) (i.e., to evaluate the DWR error estimate) so that $\overline{r}_h(w, v; \mu, t)$ is approximated by

$$\tilde{\overline{r}}_h^{\eta,2}(w, v; \mu, t) \equiv \sum_{\tilde{q}=1}^{\tilde{Q}^{\eta,2}} \tilde{\rho}_{\tilde{q}}^{\eta,2} \overline{r}_{q^{\eta,2}(\tilde{q})}(w, v; \mu, t).$$

Note that this RQ approximation of the residual, which is used to evaluate the RB-RQ DWR error estimate, is different from the approximate residual (8), which is used to evaluate the primal RB-RQ solution; we use EQP constraints that are tailored to the specific purpose of each RQ rule to construct two different RQ rules, which results in $\{\tilde{\rho}_{\tilde{q}}^{\eta,2}\}_{\tilde{q}=1}^{\tilde{Q}^{\eta,2}} \neq \{\tilde{\rho}_{\tilde{q}}^{r}\}_{\tilde{q}=1}^{\tilde{Q}^{r}}$. Given these approximations of the residual, our RB-RQ approximation of the DWR problem is as follows: given $\mu \in \mathcal{D}$ and $\tilde{u}_N(\mu) \in \mathcal{V}_N$, find $\tilde{z}_N^{\mathrm{du}}(\mu, t) \in \mathcal{V}_N^{\mathrm{du}}$ such that

$$\begin{aligned} \tilde{\overline{r}}_h^{\mathrm{du},\eta,1}(u_N(\mu, t); w_N, \tilde{z}_N^{\mathrm{du}}(\mu, t); \mu, t) = 0 \quad \forall w_N \in \mathcal{V}_N^{\mathrm{du}},\ \forall t \in [0, T), \\ m(w_N, \tilde{z}_N^{\mathrm{du}}(\mu, T); \mu) = g_h'(u_N(\mu, T); w_N; \mu) \quad \forall w_N \in \mathcal{V}_N^{\mathrm{du}}, \end{aligned} \tag{27}$$

and then evaluate the DWR error estimate

$$|s_h(\mu) - \tilde{s}_N(\mu)| \approx \tilde{\eta}_N(\mu) \equiv |\int_I \tilde{\bar{r}}_h^{\eta,2}(\tilde{u}_N(\mu,t), \tilde{z}_N^{\mathrm{du}}(\mu,t); \mu, t)dt|. \tag{28}$$

We refer to Appendix B for fully discrete approximations of the dual problem and the DWR error estimate using DIRK methods. As before, assuming the number of RQ points $\tilde{Q}^{\eta,1}$ and $\tilde{Q}^{\eta,2}$ are $O(N^{\mathrm{du}}) \ll N_h$, the RB-RQ approximation permits the evaluation of the DWR error estimate in $O((N^{\mathrm{du}})^\bullet)$ cost, which is independent of N_h and Q_h.

3.3. Error in the RB approximation of DWR due to RQ

We now analyze the error in the DWR error estimate due to the use of the RQ rules, $|\eta_N(\mu) - \tilde{\eta}_N(\mu)|$. We appeal again to the space-time framework to facilitate this analysis. (We use the framework only to enable formal manipulations; see Remark 6.) Using the space-time framework, we can compactly express the DWR error estimates (26) and (28) as

$$\eta_N(\mu) = \langle A(\tilde{u}_N(\mu);\mu), z_N^{\mathrm{du}}(\mu)\rangle \quad \text{and} \quad \tilde{\eta}_N(\mu) = \langle \tilde{A}(\tilde{u}_N(\mu);\mu), \tilde{z}_N^{\mathrm{du}}(\mu)\rangle,$$

where the space-time operator A is given by (11). We can decompose the error in the DWR error estimate due to the use of the RQ as

$$\begin{aligned} |\eta_N(\mu) - \tilde{\eta}_N(\mu)| = &|\underbrace{\langle A(\tilde{u}_N(\mu);\mu), z_N^{\mathrm{du}}(\mu) - \tilde{z}_N^{\mathrm{du}}(\mu)\rangle}_{\text{EE1}}| + |\underbrace{\langle A(\tilde{u}_N(\mu);\mu) - \tilde{A}(\tilde{u}_N(\mu);\mu), z_N^{\mathrm{du}}(\mu)\rangle}_{\text{EE2}}| \\ &+ |\underbrace{\langle A(\tilde{u}_N(\mu);\mu) - \tilde{A}(\tilde{u}_N(\mu);\mu), z_N^{\mathrm{du}}(\mu) - \tilde{z}_N^{\mathrm{du}}(\mu)\rangle}_{\text{EE3}}|. \end{aligned} \tag{29}$$

We have three sources of error: "EE1" associated with the error in dual RB-RQ solution $\tilde{z}_N^{\mathrm{du}}(\mu,t)$ given by (27) relative to the dual RB solution $z_N^{\mathrm{du}}(\mu,t)$ given by (25); "EE2" associated with the RQ approximation of the space-time operator $A(\cdot,\cdot)$ by $\tilde{A}(\cdot,\cdot)$; and "EE3" associated with the product of these two errors. Here "EE" stands for the error in the error estimate. We will introduce EQP constraints to directly control EE1 and EE2. We hope that by directly controlling EE1 and EE2, we will indirectly control the product term EE3.

To analyze EE1, we apply the DWR error analysis to the DWR error estimate given by (26), which results from solving the dual problem (25). We first recall that the dual problem is concisely expressed in the space-time framework as follows: given $\mu \in \mathcal{D}$ and $\tilde{u}_N(\mu) \in \mathcal{V}_N^{\mathrm{pr}}$, find $z_N^{\mathrm{du}}(\mu) \in \mathcal{V}_N^{\mathrm{du}}$ such that

$$\langle DA^*(\tilde{u}_N(\mu);\mu) z_N^{\mathrm{du}}(\mu), v_N\rangle = \langle DJ(\tilde{u}_N(\mu);\mu), v_N\rangle \quad \forall v_N \in \mathcal{V}_N^{\mathrm{du}},$$

where the space-time operators DA and DJ are given by (12) and (13), respectively, and $DA^*(\tilde{u}_N(\mu);\mu)$ is the formal adjoint of $DA(\tilde{u}_N(\mu);\mu)$; i.e., $\langle DA^*(\tilde{u}_N(\mu);\mu)v, w\rangle = \langle DA(\tilde{u}_N(\mu);\mu)w, v\rangle \ \forall w, v \in \mathcal{V}_N^{\mathrm{du}}$. In this context, the "output" or "quantity of interest" is the DWR error estimate

$$\eta_N(\mu) = \langle A(\tilde{u}_N), z_N^{\mathrm{du}}\rangle.$$

To analyze the error $\eta_N(\mu) - \tilde{\eta}_N(\mu)$ due to the use of RQ rules, we introduce the dual-of-the-dual problem, where we use the DWR framework to control and estimate the error in the DWR error estimate: given $\mu \in \mathcal{D}$ and $\tilde{u}_N(\mu) \in \mathcal{V}_N^{\mathrm{pr}}$, find $\psi_N(\mu) \in \mathcal{V}_N^{\mathrm{du}}$ such that

$$\langle DA(\tilde{u}_N(\mu);\mu)\psi_N(\mu), v_N\rangle = \langle A(\tilde{u}(\mu)), v_N\rangle \quad \forall v_N \in \mathcal{V}_N^{\mathrm{du}}.$$

The dual-of-the-dual problem is associated with the tangent operator DA, so we refer to the dual-of-the-dual problem as the *tangent problem* for brevity. We also refer to its solution as the *tangent solution*, or simply the *tangent*. We may appeal to the expressions for the space-time operators (11), (12), and (13) to obtain a more explicit form of the tangent problem: given $\mu \in \mathcal{D}$ and $\tilde{u}_N(\mu) \in \mathcal{V}_N^{\mathrm{pr}}$, find $\psi_N(\mu,t) \in \mathcal{V}_N^{\mathrm{du}}$ such that

$$\begin{aligned} m_h(\frac{\partial \psi_N}{\partial t}(\mu,t), v_N;\mu) + r_h'(\tilde{u}_N(\mu,t);\psi_N(\mu,t), v_N;\mu,t) = \bar{r}_h(\tilde{u}_N(\mu,t), v_N;\mu,t) \quad \forall v_N \in \mathcal{V}_N^{\mathrm{du}},\ t \in (0,T], \\ \psi_N(\mu, t=0) = 0. \end{aligned} \tag{30}$$

The tangent problem, like the dual problem, is linear; however, the tangent problem, like the primal problem and unlike the dual problem, is solved forward in time. We refer to Appendix B.5 for fully discrete approximations of the tangent problem using DIRK methods.

The following two propositions summarize EE1 and EE2 due to the use of RQ rules.

Proposition 17 (Error in DWR error estimate due to the RQ approximation of $\overline{r}_h^{\mathrm{du}}(\cdot;\cdot,\cdot;\cdot,\cdot)$)**.** *Let $\tilde{u}(\mu) \in \mathcal{V}_N$, $z_N^{\mathrm{du}}(\mu) \in \mathcal{V}_N^{\mathrm{du}}$, $\tilde{z}_N^{\mathrm{du}}(\mu) \in \mathcal{V}_N^{\mathrm{du}}$, and $\psi_N(\mu) \in \mathcal{V}_N^{\mathrm{du}}$ be the solutions to (9), (25), (27), and (30), respectively. Suppose $DA^*(\tilde{u}_N(\mu);\mu)$, $D\tilde{A}^*(\tilde{u}_N(\mu);\mu)$, $DA^{-*}(\tilde{u}_N(\mu);\mu)$, and $D\tilde{A}^{-*}(\tilde{u}_N(\mu);\mu)$ are bounded. In addition, suppose*

$$\left|\int_I [\overline{r}_h^{\mathrm{du}}(\tilde{u}_N(\mu,t);\psi_N(\mu,t),z_N^{\mathrm{du}}(\mu,t);\mu,t) - \tilde{\overline{r}}_h^{\mathrm{du}}(\tilde{u}_N(\mu,t);\psi_N(\mu,t),z_N^{\mathrm{du}}(\mu,t);\mu,t)]dt\right|$$
$$+ |m_h(\psi_N(\mu,T),z_N^{\mathrm{du}}(\mu,T)) - \tilde{m}_h(\psi_N(\mu,T),z_N^{\mathrm{du}}(\mu,T)) - g_h'(\tilde{u}_N(\mu,T);\psi_N(\mu,T)) + \tilde{g}_h'(\tilde{u}_N(\mu,T);\psi_N(\mu,T))| \le \frac{\delta^{\eta,1}}{2}, \tag{31}$$

$$\sup_{w\in\mathcal{V}_N^{\mathrm{du}},}\ \sup_{v\in\mathcal{V}_N^{\mathrm{du}}} \frac{\langle (I - DA^*(\tilde{u}_N(\mu);\mu)D\tilde{A}^{-*}(\tilde{u}_N(\mu);\mu))w, v\rangle}{\|w\|_{\mathcal{V}_N^{\mathrm{du}},}\|v\|_{\mathcal{V}}} \le \delta^{DA}, \tag{32}$$

for some $\delta^{\eta,1} \in \mathbb{R}_{>0}$ and $\delta^{DA} \in \mathbb{R}_{>0}$, where the dual norm is given by $\|w\|_{\mathcal{V}_N^{\mathrm{du}},} \equiv \sup_{v\in\mathcal{V}_N^{\mathrm{du}}}\langle w, v\rangle/\|v\|_{\mathcal{V}}$. Then

$$|\langle A(\tilde{u}_N(\mu)), z_N^{\mathrm{du}}(\mu) - \tilde{z}_N^{\mathrm{du}}(\mu)\rangle| \le \frac{\delta^{\eta,1}}{2} + O(\tilde{\delta}^2) + O((\delta^{DA})^2), \tag{33}$$

where $\tilde{\delta} \equiv \|z_N^{\mathrm{du}}(\mu) - \tilde{z}_N^{\mathrm{du}}(\mu)\|_{\mathcal{V}}$.

Proof. See Appendix A. □

Proposition 18 (Error in DWR error estimate due to the RQ approximation of $\overline{r}_h(\cdot,\cdot;\cdot,\cdot)$)**.** *Let $\tilde{u}_N(\mu) \in \mathcal{V}_N$ and $z_N^{\mathrm{du}}(\mu) \in \mathcal{V}_N^{\mathrm{du}}$ be the solutions to (9) and (25), respectively. Suppose*

$$\left|\int_I [\overline{r}(\tilde{u}_N(\mu,t),z_N^{\mathrm{du}}(\mu,t);\mu,t) - \tilde{\overline{r}}(\tilde{u}_N(\mu,t),z_N^{\mathrm{du}}(\mu,t);\mu,t)]dt\right| \le \frac{\delta^{\eta,2}}{2} \tag{34}$$

for $\delta^{\eta,2} \in \mathbb{R}_{>0}$. Then,

$$|\langle A(\tilde{u}_N) - \tilde{A}(\tilde{u}_N), z_N^{\mathrm{du}}\rangle| \le \frac{\delta^{\eta,2}}{2}. \tag{35}$$

Proof. Proof follows from the definition of the space-time operators A and $\tilde{A}$ and forms $\overline{r}(\cdot,\cdot;\cdot,\cdot)$ and $\tilde{\overline{r}}(\cdot,\cdot;\cdot,\cdot)$. □

Propositions 17 and 18 show that EE1 and EE2 in the DWR error estimate due to the use of RQ rules are bounded (i) primarily by the difference in the tangent-weighted dual residual (31) and the error in the dual-weighted residual (34), and (ii) secondarily by the error in the adjoint operators (32) and the norm of the adjoint error. In particular, note that the tangent solution is required to develop (31), which bounds EE1 as a difference between the original residual and the RQ-approximated residual; this is needed to develop appropriate EQP constraints. While the error in the error estimate is rarely explicitly controlled in typical model reduction settings, this explicit control is required for our hyperreduction framework, since the error estimate based on an arbitrary quadrature rule can have an arbitrarily poor effectivity. We now devise a systematic procedure to find RQ rules informed by EE1 and EE2.

3.4. Empirical quadrature procedure

We now describe the manifold accuracy constraints for the dual RQ rule $\{q^{\eta,1}(\tilde{q}),\tilde{\rho}_{\tilde{q}}^{\eta,1}\}_{\tilde{q}=1}^{\tilde{Q}^{\eta,1}}$, which are designed to control EE1. Our constraints are based on the tangent-weighted dual residual (31), but we obtain more conservative constraints by (i) splitting the time interval into K_{eqp} subintervals and (ii) modally decomposing the tangent solution in the tangent-weighted dual residual into L pieces. The resulting EQP statement is as follows:

Definition 19 (EQP for the dual residual RQ rule $\{q^{\eta,1}(\tilde{q}),\tilde{\rho}_{\tilde{q}}^{\eta,1}\}_{\tilde{q}=1}^{\tilde{Q}^{\eta,1}}$)**.** Given a training parameter set $\Xi_{\mathrm{eqp}} \subset \mathcal{D}$ of size N_{eqp}, we introduce the associated RB-RQ solutions $\{\tilde{u}_N(\mu) \in \mathcal{V}_N^{\mathrm{pr}}\}_{\mu\in\Xi_{\mathrm{eqp}}}$, the dual training set $\{\hat{z}_N^{\mathrm{du}}(\mu) \in \mathcal{V}_N^{\mathrm{du}}\}_{\mu\in\Xi_{\mathrm{eqp}}}$, and the tangent training set $\{\psi_N(\mu) \in \mathcal{V}_N^{\mathrm{du}}\}_{\mu\in\Xi_{\mathrm{eqp}}}$. We in addition introduce $0 = t_0 < t_1 < \cdots < t_{K_{\mathrm{eqp}}} = T$ and partition the time interval $I \equiv [0,T)$ into K_{eqp} subintervals $I_k = [t_{k-1},t_k)$, $k = 1,\ldots,K_{\mathrm{eqp}}$. Similarly, we introduce $1 = N_0 < N_1 < \cdots < N_L = N^{\mathrm{du}}$ and partition $\{1,\ldots,N^{\mathrm{du}}\}$ into L sets $S_l = \{N_{l-1},\ldots,N_l\}$, $l = 1,\ldots,L$. We then apply the EQP (Definition 8) subject to the manifold accuracy constraints

$$\left|\int_{I_k} [\overline{r}_h^{\mathrm{du}}(\tilde{u}_N(\mu,t);\Pi_{S_l}\psi_N(\mu,t),\hat{z}_N^{\mathrm{du}}(\mu,t);\mu,t) - \sum_{\tilde{q}=1}^{\tilde{Q}^{\eta,1}} \rho_{\tilde{q}}^{\eta,1}\overline{r}_q^{\mathrm{du}}(\tilde{u}_N(\mu,t);\Pi_{S_l}\psi_N(\mu,t),\hat{z}_N^{\mathrm{du}}(\mu,t);\mu,t)]dt\right| \le \frac{\delta^{\eta_1}}{2L}\frac{|I_k|}{T}, \tag{36}$$

for $k = 1, \ldots, K_{\text{eqp}} - 1$, $l = 1, \ldots, L$, and $\mu \in \Xi_{\text{eqp}}$, and

$$\begin{aligned} &\Big| \int_{I_{K_{\text{eqp}}}} \overline{r}_h^{\text{du}}(\tilde{u}_N(\mu,t); \Pi_{S_l}\psi_N(\mu,t), \hat{z}_N^{\text{du}}(\mu,t); \mu, t)dt + m_h(\Pi_{S_l}\psi_N(\mu,T), \hat{z}_N^{\text{du}}(\mu,T)) - g_h'(\tilde{u}_N(\mu,t); \Pi_{S_l}\psi_N(\mu,t); \mu) \\ &- \sum_{\tilde{q}=1}^{\tilde{Q}^{\eta,1}} \rho_{\tilde{q}}^{\eta,1} \Big[\int_{I_{K_{\text{eqp}}}} \overline{r}_q^{\text{du}}(\tilde{u}_N(\mu,t); \Pi_{S_l}\psi_N(\mu,t), \hat{z}_N^{\text{du}}(\mu,t); \mu, t)dt + m_q(\Pi_{S_l}\psi_N(\mu,T), \hat{z}_N^{\text{du}}(\mu,T)) \\ &\qquad - g'_{q^{\eta,1}(\tilde{q})}(\tilde{u}_N(\mu,t); \Pi_{S_l}\psi_N(\mu,t); \mu)\Big]\Big| \le \frac{\delta^{\eta_1}}{2L}\frac{|I_{K_{\text{eqp}}}|}{T}, \end{aligned} \tag{37}$$

for $l = 1, \ldots, L$, and $\mu \in \Xi_{\text{eqp}}$. Here, $|I_k| \equiv t_k - t_{k-1}$, and $\Pi_{S_l} : \mathcal{V}_N^{\text{du}} \to \text{span}\{\phi_i^{\text{du}}\}_{i \in S_l}$ is the L^2-projection operator. The total number of constraints is $K_{\text{eqp}} L N_{\text{eqp}}$.

We refer to Remark 11 for a discussion on the selection of K_{eqp} and L. The constraints (36) are associated with the time-dependent dual residual for all time subintervals except the final subinterval; the constraint (37) associated with the final subinterval incorporates the terminal condition. The EQP constraints (36) and (37) enforce the primary condition (31) of Proposition 17 in the sense that

$$\Big| \int_I [\overline{r}_h^{\text{du}}(\tilde{u}_N(\mu,t); \psi_N^{\text{du}}(\mu,t), z_N^{\text{du}}(\mu,t); \mu, t) - \tilde{\overline{r}}_h^{\text{du}}(\tilde{u}_N(\mu,t); \psi_N^{\text{du}}(\mu,t), z_N^{\text{du}}(\mu,t); \mu, t)]dt\Big| \le \sum_{k=1}^{K_{\text{eqp}}} \sum_{i=1}^{N} \frac{\delta^{\eta,1}}{2L}\frac{|I_k|}{T} \le \frac{\delta^{\eta,1}}{2}.$$

(As the argument is identical to the bound (18) for the primal error control, we here omit the presentation for brevity.) Hence the EQP given by Definition 19 controls the primary condition (31) of Proposition 17 and, in turn, EE1. On the other hand, similarly to the EQP constraint for the primal problem, we do not enforce the secondary condition (32). We again hope that conservatively enforcing the primary condition will indirectly control the error in the tangent operator $D\tilde{A}^*$ and hence results in the satisfaction of (32). We will demonstrate in Section 5 that the set of constraints works well in practice.

We now describe the manifold accuracy constraints for the dual RQ rule $\{q^{\eta,2}(\tilde{q}), \tilde{\rho}_{\tilde{q}}^{\eta,2}\}_{\tilde{q}=1}^{\tilde{Q}^{\eta,2}}$, which are designed to control EE2. Our EQP based on (34) is the following:

Definition 20 (EQP for the dual-weighted residual quadrature $\{q^{\eta,2}(\tilde{q}), \tilde{\rho}_{\tilde{q}}^{\eta,2}\}_{\tilde{q}=1}^{\tilde{Q}^{\eta,2}}$)**.** Given a training parameter set $\Xi_{\text{eqp}} \subset \mathcal{D}$ of size N_{eqp}, we introduce the associated RB-RQ solutions $\{\tilde{u}_N(\mu) \in \mathcal{V}_N^{\text{pr}}\}_{\mu \in \Xi_{\text{eqp}}}$ and the dual training set $\{\hat{z}_N^{\text{du}}(\mu) \in \mathcal{V}_N^{\text{du}}\}_{\mu \in \Xi_{\text{eqp}}}$. We in addition introduce $0 = t_0 < t_1 < \cdots < t_{K_{\text{eqp}}} = T$ and partition the time interval $I \equiv (0, T]$ into K_{eqp} subintervals $I_k = (t_{k-1}, t_k)$, $k = 1, \ldots, K_{\text{eqp}}$. Similarly, we introduce $1 = N_0 < N_1 < \cdots < N_L = N^{\text{du}}$ and partition $\{1, \ldots, N^{\text{du}}\}$ into L sets $S_l = \{N_{l-1}, \ldots, N_l\}$, $l = 1, \ldots, L$. We then apply the EQP (Definition 8) subject to the manifold accuracy constraints

$$\Big| \int_{I_k} [\overline{r}_h(\tilde{u}_N(\mu,t), \Pi_{S_l}\hat{z}_N^{\text{du}}(\mu,t); \mu, t) - \sum_{\tilde{q}=1}^{\tilde{Q}^{\eta,2}} \rho_{\tilde{q}}^{\eta,2} r_q(\tilde{u}_N(\mu,t), \Pi_{S_l}\hat{z}_N^{\text{du}}(\mu,t); \mu, t)]dt\Big| \le \frac{\delta^{\eta,2}}{2L}\frac{|I_k|}{T}, \tag{38}$$

for $k = 1, \ldots, K_{\text{eqp}}$, $l = 1, \ldots, L$, and $\mu \in \Xi_{\text{eqp}}$. Here, $|I_k| \equiv t_k - t_{k-1}$, and $\Pi_{S_l} : \mathcal{V}_N^{\text{du}} \to \text{span}\{\phi_i^{\text{du}}\}_{i \in S_l}$ is the L^2-projection operator. The total number of constraints is $K_{\text{eqp}} L N_{\text{eqp}}$.

We readily observe that (38) enforces (34) in Proposition 18 in the sense that

$$\begin{aligned} &\Big| \int_I [\overline{r}(\tilde{u}_N(\mu,t), \hat{z}_N^{\text{du}}(\mu,t); \mu, t) - \tilde{\overline{r}}(\tilde{u}_N(\mu,t), \hat{z}_N^{\text{du}}(\mu,t); \mu, t)]dt\Big| \\ &= \Big| \sum_{k=1}^{K_{\text{eqp}}} \sum_{i=1}^{N} \int_{I_k} [\overline{r}_h(\tilde{u}_N(\mu,t), \Pi_{S_l}\hat{z}_{N,i}^{\text{du}}(\mu,t); \mu, t) - \sum_{\tilde{q}=1}^{\tilde{Q}^{\eta,2}} \rho_{\tilde{q}}^{\eta,2} r_q(\tilde{u}_N(\mu,t), \Pi_{S_l}\hat{z}_{N,i}^{\text{du}}(\mu,t); \mu, t)]dt\Big| \le \sum_{k=1}^{K_{\text{eqp}}} \sum_{i=1}^{N} \frac{\delta^{\eta,2}}{2L}\frac{|I_k|}{T} \le \frac{\delta^{\eta,2}}{2}. \end{aligned}$$

Hence the EQP given by Definition 20 enforces the condition (18) in Proposition 18 and, in turn, EE2. We conclude this section with a few remarks about the EQPs for the DWR error estimate.

Remark 21. In the above, we invoke two separate EQPs (Definitions 19 and 20) to find two separate RQ rules $\{q^{\eta,1}(\tilde{q}), \tilde{\rho}_{\tilde{q}}^{\eta,1}\}_{\tilde{q}=1}^{\tilde{Q}^{\eta,1}}$ and $\{q^{\eta,2}(\tilde{q}), \tilde{\rho}_{\tilde{q}}^{\eta,2}\}_{\tilde{q}=1}^{\tilde{Q}^{\eta,2}}$ for (27) and (28), respectively. We may instead invoke a single EQP to find a single common RQ rule $\{q^{\eta}(\tilde{q}), \tilde{\rho}_{\tilde{q}}^{\eta}\}_{\tilde{q}=1}^{\tilde{Q}^{\eta}}$ for both (27) and (28). This EQP would include constraints (36), (37), and (38), and the total number of constraints would be $2K_{\text{eqp}} L N_{\text{eqp}}$. This is the procedure we use in the numerical examples in Section 5.

Remark 22. In this section we seek to identify EQP accuracy constraints to control the hyperreduction error in the DWR error estimate. In the original *steady* goal-oriented EQP formulation [58], the authors seek to identify these constraints by analyzing an algebraic equation. However, we cannot use a similar algebraic framework for *unsteady* PDEs because we must solve a differential (rather than algebraic) equation in time. To overcome this challenge, we have appealed to the space-time framework and identified the differential equation (30) associated with the dual-of-the-dual problem. We use the dual-of-the-dual problem to develop tractable EQP constraints (36) and (37) for unsteady DWR. This approach readily extends to steady PDEs. In fact, the algebraic constraints identified in [58] implicitly contain the algebraic solution to the dual-of-the-dual problem.

Remark 23. While EQPs in Definitions 19 and 20 are designed to control EE1 given by $\langle A(\tilde{u}_N(\mu);\mu), z_N^{\mathrm{du}}(\mu) - \tilde{z}_N^{\mathrm{du}}(\mu)\rangle$ and EE2 given by $\langle A(\tilde{u}_N(\mu);\mu) - \tilde{A}(\tilde{u}_N(\mu);\mu), z_N^{\mathrm{du}}(\mu)\rangle$, respectively, we do not control the product term EE3 given by $\langle A(\tilde{u}_N(\mu);\mu) - \tilde{A}(\tilde{u}_N(\mu);\mu), z_N^{\mathrm{du}}(\mu) - \tilde{z}_N^{\mathrm{du}}(\mu)\rangle$. The steady goal-oriented EQP framework presented in [58] modified constraints for EE1 and EE2 so that EE3 is indirectly controlled. However, we have found an analogous modification is difficult, if not impossible, for time-dependent problems for the reasons explained in Remark 22. We could in principle introduce additional manifold accuracy constraints to directly control the global dual error $\|z_N^{\mathrm{du}} - \tilde{z}_N^{\mathrm{du}}\|_{\mathcal{Y}}$ and the error in the residual evaluation $\|A(\tilde{u}_N) - \tilde{A}(\tilde{u}_N)\|_{\mathcal{Y}'}$; however, this approach would significantly increase the number of EQP constraints and hence the training cost. In this work, we will instead *assume* that the existing constraints control EE3 reasonably well. If EE3 is not well controlled, then we could employ the iterative procedure described in [51] (see Remark 24).

Remark 24. In the thesis [51] on which this paper is based, we consider more conservative EQP constraints than those defined in Definitions 19 and 20 by splitting the constraints into portions associated with the derivative of the output and portions associated with the derivative of the residual. In addition, we explore an iterative procedure to modify $\delta^{\eta,1}$ and $\delta^{\eta,2}$ so that we achieve the desired error tolerance on $|\tilde{\eta}_N(\mu) - \eta_N(\mu)|$. This iterative procedure also accounts for the effect of EE3, which is it not explicitly controlled by the EQP. The EQP for the dual problem (Definition 19) requires the solution of the dual problem and is more computationally expensive than the EQP for the DWR (Definition 20). Thus, if we wish to achieve $|\tilde{\eta}_N(\mu) - \eta_N(\mu)| \le \delta^{\eta}$, then we choose $\delta^{\eta,1} = \delta^{\eta}$ and iteratively decrease $\delta^{\eta,2} < \delta^{\eta}$ until we satisfy the desired error tolerance; i.e., we decrease $\delta^{\eta,2}$ and re-solve the EQP for the DWR (Definition 20) until the error tolerance is satisfied. With this iterative procedure, we can control all three error sources of the DWR error estimate in (29). We refer to [51] for details.

4. Offline-online computational decomposition

4.1. Offline training

We now present an offline training procedure to simultaneously construct FE spaces, RBs, and RQ rules. Our approach is based on the POD-greedy algorithm [31, 30, 27]. Like other greedy algorithms, we leverage the online-efficient error estimate to rapidly explore the parameter space $\mathcal{D}$ and to minimize the number of FE solves by choosing an "optimal" sequence of points. Given that our problems are nonlinear, we simultaneously train RBs and RQs, as explored for RB-EIM in [20, 19] and for steady RB-EQP in [59, 58]. In addition, we simultaneously adapt the FE mesh, as explored for steady problems in [2, 57, 58].

Algorithm 1 summarizes our training procedure. The algorithm is designed to control the output error in the RB-RQ solution (9) with respect to the PDE (1). We first provide a high-level description of the algorithm in the context of output error control:

$$|s(\mu) - \tilde{s}_N(\mu)| \le \overbrace{\underbrace{|s(\mu) - s_h(\mu)|}_{\substack{\text{FE error: } \delta^{\mathrm{fe}} \\ \text{(line 3)}}}}^{\text{estimated by } \eta_h^{\mathrm{fe}}(\mu)} + \overbrace{\underbrace{|s_h(\mu) - s_N(\mu)|}_{\substack{\text{RB error: } \delta^{\mathrm{pr}} \\ \text{(line 5)}}} + \underbrace{|s_N(\mu) - \tilde{s}_N(\mu)|}_{\substack{\text{RQ errors: } \delta^{r} \text{ and } \delta^{J} \\ \text{(line 7)}}}}^{\text{estimated by } \tilde{\eta}_N(\mu) \text{ for all } \mu \in \Xi_{\mathrm{train}}}. \tag{39}$$

The algorithm controls the FE, RB, and RQ errors to their respective tolerances δ^{fe}, δ^{rb}, and $\{\delta^r, \delta^J\}$ at $\mu \in \Xi_{\mathrm{greedy}}$ in the lines indicated. (We discuss the selection of the tolerances in Remark 25.) The algorithm terminates when the error estimate $\tilde{\eta}_N(\mu)$ for the RB-RQ error is less than δ^{greedy} for all $\mu \in \Xi_{\mathrm{train}}$. Similarly, the error in the error estimate can be decomposed as

$$\big|\underbrace{|s_h(\mu) - \tilde{s}_N(\mu)|}_{\approx \text{true error}} - \tilde{\eta}_N(\mu)\big| \approx |\eta_h(\mu) - \tilde{\eta}_N(\mu)| \le \underbrace{|\eta_h(\mu) - \eta_N(\mu)|}_{\substack{\text{RB error: } \delta^{\mathrm{du,rel}} \\ \text{(line 6)}}} + \underbrace{|\eta_N(\mu) - \tilde{\eta}_N(\mu)|}_{\substack{\text{RQ errors: } \delta^{\eta,1} \text{ and } \delta^{\eta,2} \\ \text{(line 8)}}} \tag{40}$$

The algorithm again controls the RB and RQ errors in the DWR error estimate to their respective tolerances $\delta^{\text{du,rel}}$ and $\{\delta^{\eta,1}, \delta^{\eta,2}\}$ at $\mu \in \Xi_{\text{greedy}}$. We discuss each step of the algorithm and the associated computational cost in more detail in the following paragraphs.

Algorithm 1: Simultaneous RB-RQ POD-greedy training for time-dependent problems

input : training set $\Xi_{\text{train}} \subset \mathcal{D}$
Overall greedy tolerance: δ^{greedy}
FE tolerance: δ^{fe}
RB tolerances: δ^{pr}, $\delta^{\text{du,rel}}$
EQP tolerances: δ^{r}, δ^{J}, $\delta^{\eta,1}$, $\delta^{\eta,2}$
output: RB spaces: $\mathcal{V}_N^{\text{pr}}$ and $\mathcal{V}_N^{\text{du}}$
RQ rules: $\{q^r(\tilde{q}), \tilde{\rho}_{\tilde{q}}^r\}_{\tilde{q}=1}^{\tilde{Q}^r}$, $\{q^J(\tilde{q}), \tilde{\rho}_{\tilde{q}}^J\}_{\tilde{q}=1}^{\tilde{Q}^J}$, $\{q^{\eta,1}(\tilde{q}), \tilde{\rho}_{\tilde{q}}^{\eta,1}\}_{\tilde{q}=1}^{\tilde{Q}^{\eta,1}}$, and $\{q^{\eta,2}(\tilde{q}), \tilde{\rho}_{\tilde{q}}^{\eta,2}\}_{\tilde{q}=1}^{\tilde{Q}^{\eta,2}}$

1 Set $\Xi_{\text{greedy}} = \{\mu^{(1)} \equiv \text{centroid}(\Xi_{\text{train}})\}$ and $N_{\text{greedy}} = 1$.
2 **while** $N_{\text{greedy}} < N_{\text{train}}$ **do**
3 *Solve primal FE problem:* solve (2) for $u_h(\mu^{(N_{\text{greedy}})}) \in \mathcal{Y}_h$; adapt the mesh as necessary to meet δ^{fe}.
4 *Solve dual FE problem:* solve (23) for $z_h(\mu^{(N_{\text{greedy}})}) \in \mathcal{Y}_h$.
5 *Construct primal RB:* set $\mathcal{V}_N^{\text{pr}} \leftarrow \mathcal{V}_N^{\text{pr}} \oplus \text{POD}_{N_{\text{add}}^{\text{pr}}}(\{\Pi^{\perp}_{\mathcal{V}_N^{\text{pr}}} u_h(\mu^{(N_{\text{greedy}})}, t)\}_{t \in I})$, where $N_{\text{add}}^{\text{pr}}$ is chosen so that $|s_h(\mu^{(N_{\text{greedy}})}) - s_N(\mu^{(N_{\text{greedy}})})| \leq \delta^{\text{pr}}$.
6 *Construct dual RB:* set $\mathcal{V}_N^{\text{du}} \leftarrow \mathcal{V}_N^{\text{du}} \oplus \text{POD}_{N_{\text{add}}^{\text{du}}}(\{\Pi^{\perp}_{\mathcal{V}_N^{\text{du}}} z_h(\mu^{(N_{\text{greedy}})}, t)\}_{t \in I})$, where $N_{\text{add}}^{\text{du}}$ is chosen so that the relative error with respect to a reference error estimate is less than $\delta^{\text{du,rel}}$.
7 *Construct primal RQs:* find $\{q^r(\tilde{q}), \rho_{\tilde{q}}^r\}_{\tilde{q}=1}^{\tilde{Q}^r}$ and $\{q^J(\tilde{q}), \rho_{\tilde{q}}^J\}_{\tilde{q}=1}^{\tilde{Q}^J}$ using EQPs for residual and output functional (Definitions 10 and 13) for the training set $\Xi_{\text{eqp}} = \Xi_{\text{greedy}}$.
8 *Construct dual and DWR RQs:* find $\{q^{\eta,1}(\tilde{q}), \rho_{\tilde{q}}^{\eta,1}\}_{\tilde{q}=1}^{\tilde{Q}^{\eta,1}}$ and $\{q^{\eta,2}(\tilde{q}), \rho_{\tilde{q}}^{\eta,2}\}_{\tilde{q}=1}^{\tilde{Q}^{\eta,2}}$ using EQPs for dual problem and DWR (Definitions 19 and 20) for $\Xi_{\text{eqp}} = \Xi_{\text{greedy}}$.
9 *Find parameter with maximum error:* for all $\mu \in \Xi_{\text{train}}$, solve (9) for $\tilde{u}_{N_{\text{pr}}}(\mu)$, solve (27) for $\tilde{z}_N^{\text{du}}(\mu)$, and evaluate (28) for $\tilde{\eta}_N(\mu)$. Then set $\mu^{\star} = \arg\sup_{\mu \in \Xi_{\text{train}}} |\tilde{\eta}_N(\mu)|$.
10 *Check convergence:* if $\tilde{\eta}_N(\mu^{\star}) \leq \delta^{\text{greedy}}$, terminate.
11 *Update* Ξ_{greedy}*:* set $\Xi_{\text{greedy}} \leftarrow \{\Xi_{\text{greedy}}, \mu^{(N_{\text{greedy}}+1)} \equiv \mu^{\star}\}$ and $N_{\text{greedy}} \leftarrow N_{\text{greedy}} + 1$.
12 **end**

Solve the (primal) FE problem (line 3). In this step, we compute the "truth" FE solution used to construct the primal RB $\{\phi_i^{\text{pr}}\}_{i=1}^{N^{\text{pr}}}$. To ensure the snapshots are accurate, we use the standard adaptation strategy based on solve, evaluate, mark, refine steps (see, e.g., [9]). Specifically, we construct a single adapted mesh for the entire time interval I instead of time-dependent adapted meshes. This ensures that all solution snapshots lie in the same FE space. In principle we could employ time-dependent mesh refinement and perform any operations that involve functions in two different spaces in a common enriched space; however, for simplicity, we do not consider this approach. Our adaptive mesh refinement algorithm is as follows:

1. Solve the (primal) FE problem (2) for $\{u_h(\mu^{(N_{\text{greedy}})}, t) \in \mathcal{V}_h\}_{t \in I}$.
2. Solve the enriched FE dual problem (20) for $\{z_{\hat{h}}(\mu^{(N_{\text{greedy}})}, t) \in \mathcal{V}_{\hat{h}}\}_{t \in I}$.
3. Evaluate the DWR error estimate $\eta_h^{\text{fe}}(\mu^{(N_{\text{greedy}})})$ given by (22).
4. If $\eta_h^{\text{fe}}(\mu^{(N_{\text{greedy}})}) < \delta^{\text{fe}}$, terminate.
5. Compute the time-integrated element-wise localized error indicator associated with $\eta_h^{\text{fe}}(\mu^{(N_{\text{greedy}})})$ given by

$$\eta_h^{\text{fe}}(\mu^{(N_{\text{greedy}})}) = \left| \int_I \overline{r}_h(\tilde{u}_h(\mu); z_{\hat{h}}(\mu)|_{\kappa}; \mu, t) dt \right|,$$

   where κ is an element in the mesh. Each element-wise error indicator is integrated over the entire time interval I.
6. Mark 10% of elements with largest error indicator for refinement.
7. Refine the marked elements, and go to Step 1.

Whenever the mesh is refined for a given parameter value, we invoke the adaptive FE solver for all previous parameter values using the new refined mesh as the starting mesh. This approach incurs a higher computational cost than simply re-representing the previously computed reduced basis on the refined mesh. However, because

the FE error (or the FE DWR error estimate) is not guaranteed to be nonincreasing under mesh refinement, the strategy is necessary to ensure that the FE error tolerance is met at all previous parameter values.

The overall computational cost is dominated by Steps 1 and 2. Each solution of the primal problem in Step 1 requires $O(Q_h) + O(N_h^\bullet)$ computation for the residual (and Jacobian) evaluation and the linear system solve. Similarly, each solution of the dual problem in Step 2 requires $O(Q_{\hat{h}}) + O(N_{\hat{h}}^\bullet)$ computation. The dual problem is linear and requires only one linear solve per time step. However, the cost of the primal and dual solves are comparable in practice because the time-dependent primal problem requires relatively few Newton iterations to converge, while the dual problem is solved in an enriched space. The total cost is equal to the product of these costs with the number of adaptation iterations.

Solve dual FE problem (line 4). We solve the dual FE problem (23) for $\{z_h(\mu^{(N_{\text{greedy}})}, t) \in \mathcal{V}_h\}_{t\in I}$. The computational cost of the FE solve is $O(Q_h) + O(N_h^\bullet)$ for the dual residual and Jacobian evaluation and linear system solve. Instead of solving (23), we could approximate $\{z_h(\mu^{(N_{\text{greedy}})}, t) \in \mathcal{V}_h\}_{t\in I}$ by projecting the enriched dual solution $\{z_{\hat{h}}(\mu^{(N_{\text{greedy}})}, t) \in \mathcal{V}_{\hat{h}}\}_{t\in I}$ onto $\mathcal{V}_h$; however, as the solution of the (non-enriched) dual problem (23) is relatively inexpensive, we do not use this approach. We note that we solve the dual FE problem only once per parameter to reduce the computational cost of the offline training stage (see remark 27). Similarly to the primal solution, the dual solution is recomputed for all previous parameter values whenever the mesh is refined.

Construct primal RB: (line 5). In this step, we enrich the primal RB space $\mathcal{V}_N^{\text{pr}}$ so that the solution to the RB problem (7) meets the user-prescribed output error tolerance δ^{pr} for the current parameter value $\mu^{(N_{\text{greedy}})}$. We find the appropriate level of enrichment iteratively using the following algorithm:

1. Project $\{u_h(\mu^{(N_{\text{greedy}})}, t)\}_{t\in I}$ onto the space orthogonal to $\mathcal{V}_N^{\text{pr}}$ to identify components of the solution that lie outside of the current RB space $\mathcal{V}_N^{\text{pr}}$; we denote the resulting functions by $\{y(t)\}_{t\in I} \equiv \{\Pi^\perp_{\mathcal{V}_N^{\text{pr}}} u_h(\mu^{(N_{\text{greedy}})}, t)\}_{t\in I}$.
2. Apply POD to $\{y(t)\}_{t\in I}$ to obtain a set of dominant modes $\{\zeta_i\} \equiv \text{POD}(\{y(t)\}_{t\in I})$.
3. Augment $\mathcal{V}_N^{\text{pr}}$ with the $N_{\text{add}}^{\text{pr}}$ most dominant modes $\{\zeta_i\}_{i=1}^{N_{\text{add}}^{\text{pr}}}$; i.e., $\mathcal{V}_N^{\text{pr}} \leftarrow \mathcal{V}_N^{\text{pr}} \oplus \text{span}\{\zeta_i\}_{i=1}^{N_{\text{add}}^{\text{pr}}}$.
4. Solve the RB problem (7) for $u_N(\mu^{(N_{\text{greedy}})}) \in \mathcal{V}_N^{\text{pr}}$ using the updated $\mathcal{V}_N^{\text{pr}}$ and evaluate $s_N(\mu^{(N_{\text{greedy}})})$.
5. Check if the output error satisfies $|s_h(\mu^{(N_{\text{greedy}})}) - s_N(\mu^{(N_{\text{greedy}})})| \le \delta^{\text{pr}}$.
6. If the tolerance is met, then terminate; otherwise, increase $N_{\text{add}}^{\text{pr}}$ and go back to Step 3.

In words, in Steps 1 and 2 we identify the most dominant modes that are missing from the existing $\mathcal{V}_N^{\text{pr}}$. In Steps 3–6, we iteratively enrich $\mathcal{V}_N^{\text{pr}}$ until the RB output $s_N(\mu^{(N_{\text{greedy}})})$ meets the user-prescribed error tolerance. Our implementation of the POD-greedy algorithm differs from its typical implementation (in e.g. [31, 27, 30]), where only the single most dominant POD mode is added to the RB per greedy iteration; we use this alternative approach that adds multiple POD modes to reduce the number of greedy iterations and hence the number of FE solves. Step 4, which requires the solution of the (non-hyperreduced) RB problem (7) dominates the computational cost of this algorithm. The overall cost is $O(N_{\text{iter}}(Q_h + (N^{\text{pr}})^3))$, where N_{iter} is the number of iterations required to find a suitable $N_{\text{add}}^{\text{pr}}$.

We make one cautionary remark. As the output error does not decrease monotonically for all $\mu \in \Xi_{\text{greedy}}$ under basis enrichment, the addition of $N_{\text{add}}^{\text{pr}}$ basis functions may result in the violation of the output error condition for $\mu \in \Xi_{\text{greedy}} \setminus \mu^{(N_{\text{greedy}})}$. As a result, the greedy algorithm may need to revisit a parameter that is already in Ξ_{greedy} later in the training procedure.

Construct dual RB: (line 6). In this step, we enrich the dual RB space $\mathcal{V}_N^{\text{du}}$ so that the DWR error estimate (26) meets the user-prescribed (relative) error estimate accuracy $\delta^{\text{du,rel}}$ for the current parameter value $\mu^{(N_{\text{greedy}})}$. We find the appropriate level of enrichment iteratively using the following algorithm:

1. Project $\{z_h(\mu^{(N_{\text{greedy}})}, t)\}_{t\in I}$ onto the space orthogonal to $\mathcal{V}_N^{\text{du}}$ to identify components of the solution that lie outside of the current RB space $\mathcal{V}_N^{\text{du}}$; we denote the resulting functions by $\{v(t)\}_{t\in I} \equiv \{\Pi^\perp_{\mathcal{V}_N^{\text{du}}} z_h(\mu^{(N_{\text{greedy}})}, t)\}_{t\in I}$.
2. Apply POD to $\{v(t)\}_{t\in I}$ to obtain a set of dominant modes $\{\chi_i\} \equiv \text{POD}(\{v(t)\}_{t\in I})$.
3. Construct an enriched reference dual space $\mathcal{V}_N^{\text{du,ref}} = \mathcal{V}_N^{\text{du}} \oplus \text{span}\{\chi_i\}_{i=1}^{N_{\text{add}}^{\text{du,ref}}}$ for $N^{\text{du,ref}}$ sufficiently large.
4. Using the reference space $\mathcal{V}_N^{\text{du,ref}}$, solve the RB dual problem (25) for $z_N^{\text{du,ref}}(\mu^{(N_{\text{greedy}})}) \in \mathcal{Y}_N^{\text{du,ref}}$, and evaluate the error estimate $\eta_N^{\text{ref}}(\mu^{(N_{\text{greedy}})})$ using (26).
5. Augment $\mathcal{V}_N^{\text{du}}$ with the $N_{\text{add}}^{\text{du}}$ most dominant modes $\{\chi_i\}_{i=1}^{N_{\text{add}}^{\text{du}}}$; i.e., $\mathcal{V}_N^{\text{du}} \leftarrow \mathcal{V}_N^{\text{du}} \oplus \text{span}\{\chi_i\}_{i=1}^{N_{\text{add}}^{\text{du}}}$.
6. Using the updated $\mathcal{V}_N^{\text{du}}$, solve the RB dual problem (25) for $z_N^{\text{du}}(\mu^{(N_{\text{greedy}})}) \in \mathcal{Y}_N^{\text{du}}$, and evaluate the error estimate $\eta_N(\mu^{(N_{\text{greedy}})})$ using (26).
7. Check if the error estimate is saturated in the sense that $|\eta_N^{\text{ref}}(\mu^{(N_{\text{greedy}})}) - \eta_N(\mu^{(N_{\text{greedy}})})|/|\eta_N^{\text{ref}}(\mu^{(N_{\text{greedy}})})| < \delta^{\text{du,rel}}$.
8. If the tolerance is met, then terminate; otherwise, increase $N_{\text{add}}^{\text{du}}$ and go back to Step 5.

In words, in Steps 1 and 2 we identify most dominant modes that are missing from the original $\mathcal{V}_N^{\mathrm{du}}$. In Steps 3 and 4, we compute the "reference" dual solution $z_N^{\mathrm{du,ref}}(\mu^{(N_{\mathrm{greedy}})}) \in \mathcal{Y}_{N^{\mathrm{ref}}}^{\mathrm{du}}$ and the DWR error estimate $\eta_{N^{\mathrm{ref}}}(\mu^{(N_{\mathrm{greedy}})})$ using an enriched dual space $\mathcal{Y}_{N^{\mathrm{ref}}}^{\mathrm{du}}$. In Steps 5–8, we iteratively enrich $\mathcal{V}_N^{\mathrm{du}}$ until the relative error in the DWR error estimate (measured with respect to the reference DWR error estimate $\eta_{N^{\mathrm{ref}}}(\mu^{(N_{\mathrm{greedy}})})$) is smaller than the user-prescribed relative error tolerance $\delta^{\mathrm{du,rel}}$.

Steps 4 and 6, which require the solution of the (non-hyperreduced) RB dual problem (25) and the evaluation of the (non-hyperreduced) DWR error estimate (26), dominate the computational cost. The overall cost is $(O(N_{\mathrm{iter}} + 1)(Q_h + (N^{\mathrm{du}})^3))$, where N_{iter} is the number of iterations required to find a suitable $N_{\mathrm{add}}^{\mathrm{du}}$. The solution of the "reference" dual solution $z_N^{\mathrm{du,ref}}(\mu^{(N_{\mathrm{greedy}})})$ is only slightly more expensive than the dual solution $z_N^{\mathrm{du}}(\mu^{(N_{\mathrm{greedy}})})$ because the cost of the non-hyperreduced RB problems are dominated by the residual and Jacobian evaluations, which do not depend strongly on the dimension of the RB space. However, it is still important to minimize N^{du}, because the number of RQ points increases with higher N^{du}.

Construct residual and output functional RQs (line 7). We invoke the residual EQP (Definition 10) and output functional EQP (Definition 13) to find $\{q^r(\tilde{q}), \rho_{\tilde{q}}^r\}_{\tilde{q}=1}^{\tilde{Q}^r}$ and $\{q^J(\tilde{q}), \rho_{\tilde{q}}^J\}_{\tilde{q}=1}^{\tilde{Q}^J}$, respectively. The EQP training parameter set Ξ_{eqp} is the current greedy parameter set $\Xi_{\mathrm{greedy}} \equiv \{\mu^{(i)}\}_{i=1}^{N_{\mathrm{greedy}}}$. This choice allows us to use the EQP training state set $\{\hat{u}_N(\mu) \equiv u_N(\mu)\}_{\mu\in\Xi_{\mathrm{greedy}}}$, which has been computed in line 5. We use the FE dual solution projected onto $\mathcal{V}_N^{\mathrm{pr}}$ as the dual state set $\{\hat{z}_N^{\mathrm{pr}}(\mu) \equiv \Pi_{\mathcal{V}_N^{\mathrm{pr}}} z_h(\mu)\}_{\mu\in\Xi_{\mathrm{greedy}}}$ to avoid the additional (non-hyperreduced) RB dual solve. The cost of this step is divided into two parts: the cost to compute the EQP constraints, which scales with $Q_h N_{\mathrm{eqp}}$; the cost to solve for the sparse weights, which scales with the number of constraints LKN_{eqp} and the number of FE quadrature Q_h.

Construct dual and DWR RQs (line 8). We invoke the dual residual EQP (Definition 19) and DWR EQP (Definition 20) to find $\{q^{\eta,1}(\tilde{q}), \rho_{\tilde{q}}^{\eta,1}\}_{\tilde{q}=1}^{\tilde{Q}^{\eta,1}}$ and $\{q^{\eta,2}(\tilde{q}), \rho_{\tilde{q}}^{\eta,2}\}_{\tilde{q}=1}^{\tilde{Q}^{\eta,2}}$. We choose the current greedy parameter set $\Xi_{\mathrm{greedy}} \equiv \{\mu^{(i)}\}_{i=1}^{N_{\mathrm{greedy}}}$ as the EQP training parameter set Ξ_{eqp}. This choice allows us to use the EQP training dual set $\{\hat{z}_N^{\mathrm{du}}(\mu) = z_N^{\mathrm{du}}(\mu)\}_{\mu\in\Xi_{\mathrm{greedy}}}$, which has been computed in line 6. We solve the (non-hyperreduced) tangent problem (30) to obtain $\{\psi_N(\mu)\}_{\mu\in\Xi_{\mathrm{greedy}}}$. The cost of this step is again divided into two parts: the cost to compute the EQP constraints, which scales with $Q_h N_{\mathrm{eqp}}$; the cost to solve for the sparse weights, which scales with LKN_{eqp} and Q_h.

Find parameter with maximum error (line 9). We solve the RB-RQ primal problem (9), solve the RB-RQ dual problem (27), and evaluate the RB-RQ DWR error estimate (28) for all $\mu \in \Xi_{\mathrm{train}}$. We then identify the parameter value $\mu^\star$ with the largest error estimate, and choose it as the next greedy parameter. For each $\mu \in \Xi_{\mathrm{train}}$, the solution of the primal and dual problems dominate the computation cost, which scales as $O((N^{\mathrm{pr}})^3 + (N^{\mathrm{pr}})^2\tilde{Q}^r)$ and $O((N^{\mathrm{du}})^3 + (N^{\mathrm{du}})^2\tilde{Q}^{\eta,1})$, respectively. Assuming $\tilde{Q}^{\mathrm{pr}} = O(N^r)$ and $\tilde{Q}^{\eta,1} = O(N^{\mathrm{du}})$, the overall computational cost is hence $O(((N^{\mathrm{pr}})^3 + (N^{\mathrm{du}})^3)N_{\mathrm{train}})$.

Having described Algorithm 1, we conclude the section with three remarks on the choice of the tolerances and overall computational cost.

Remark 25. The tolerance that controls the overall output accuracy of the RB-RQ model is δ^{greedy}. This parameter is set by the user based on the engineering need; e.g., in an aerodynamics simulation, the engineer might target 1% error in drag. Given δ^{greedy}, we then (somewhat conservatively) budget the errors. Typically, we use the FE tolerance of $\delta^{\mathrm{fe}} = \delta^{\mathrm{greedy}}/10$, the primal RB tolerance of $\delta^{\mathrm{pr}} = \delta^{\mathrm{greedy}}/10$, the dual relative RB tolerance of $\delta^{\mathrm{du,rel}} = 1/10$, and the EQP tolerances of $\delta^r = \delta^{\eta,1} = \delta^{\eta,2} = \delta^{\mathrm{greedy}}/10$ and $\delta^J = \delta^{\mathrm{greedy}}/100$. We refer to the error decompositions (39) and (40) for the errors controlled by each tolerance.

All tolerances are naturally specified once the user selects the overall output tolerance δ^{greedy} based on the engineering need. This is only possible because all tolerances are associated with a single common quantity of interest that is engineering-relevant. This would not be possible if, for example, the tolerance parameter for hyperreduction were associated with interpolation error of the residual operator. In this sense, the greedy algorithm presented, which builds on the goal-oriented hyperreduction framework, automates the RB-RQ training.

Remark 26. As we will see in Section 5, in a typical problem of interest, $\approx 50\%$ of the overall offline training time is spent on the (adaptive) FE solves (lines 3 and 4). The next two dominant costs are the identification of the appropriate primal and dual RB sizes using the recursive algorithm (lines 5 and 6) and the EQPs (lines 7 and 8), which each take $\approx 25\%$ of the offline training time. The error sampling time is a small fraction $\lesssim 5\%$ of the training time even for relatively dense training set thanks to the online efficient *a posteriori* error estimate.

Remark 27. We use the dual FE solution computed in line 4 to construct the dual RB (line 6), to find primal RQs (line 7), and to find DWR RQs (line 8). In principle, the dual FE problems involved in each of these steps could be linearized about the appropriate primal solution (i.e., the solution for which we wish to compute the error

estimate) to improve the quality of our error estimate and error control. However, solving multiple FE problem is computationally expensive. The choice of linearization point does not significantly impact the quality of the error estimate, so we choose to compute only one dual FE solution per greedy iteration, and to use this solution in all of the above steps.

4.2. Online dataset and evaluation

The training procedure described in Section 4.1 yields an `OnlineDataset` that comprises the following:

- RQ rules: RQ indices $\{q^\bullet(\tilde{q})\}_{\tilde{q}=1}^{\tilde{Q}^\bullet}$ and the associated RQ points $\{\tilde{x}_{\tilde{q}}^\bullet \equiv x_{q^\bullet(\tilde{q})}\}_{\tilde{q}=1}^{\tilde{Q}^\bullet}$ and RQ weights $\{\tilde{\rho}_{\tilde{q}}^\bullet \equiv \rho_{q^\bullet(\tilde{q})}^\bullet\}_{\tilde{q}=1}^{\tilde{Q}^\bullet}$ for $\bullet \in \{$ "r", "J", "$\eta, 1$", "$\eta, 2$"$\}$.
- Primal RB evaluated at RQ points: $\{\{\phi_i^{\text{pr}}(\tilde{x}_{\tilde{q}}^\bullet)\}_{i=1}^{N^{\text{pr}}}\}_{\tilde{q}=1}^{\tilde{Q}^\bullet}$ and $\{\{\nabla\phi_i^{\text{pr}}(\tilde{x}_{\tilde{q}}^\bullet)\}_{i=1}^{N^{\text{pr}}}\}_{\tilde{q}=1}^{\tilde{Q}^\bullet}$ for $\bullet \in \{$"r", "J", "$\eta, 1$", "$\eta, 2$"$\}$.
- Dual RB evaluated at RQ points: $\{\{\phi_i^{\text{du}}(\tilde{x}_{\tilde{q}}^\bullet)\}_{i=1}^{N^{\text{du}}}\}_{\tilde{q}=1}^{\tilde{Q}^\bullet}$ and $\{\{\nabla\phi_i^{\text{du}}(\tilde{x}_{\tilde{q}}^\bullet)\}_{i=1}^{N^{\text{du}}}\}_{\tilde{q}=1}^{\tilde{Q}^\bullet}$ for $\bullet \in \{$ "$\eta, 1$", "$\eta, 2$" $\}$.

Given the `OnlineDataset`, we invoke the RB-RQ model to compute the output and the associated *a posteriori* error estimate as follows:

1. Given $\mu \in \mathcal{D}$, solve the RB-RQ problem (9) for $\tilde{u}_N(\mu) \in \mathcal{V}_N^{\text{pr}}$.
2. Evaluate RB-RQ output functional (10) to obtain the associated output $\tilde{J}_N(\mu)$.
3. Solve the associated RB-RQ dual problem (27) for $\tilde{z}_N(\mu) \in \mathcal{V}_N^{\text{du}}$.
4. Evaluate the RB-RQ DWR error estimate (28) to obtain $\tilde{\eta}_N(\mu)$.

We can readily evaluate all of the RB-RQ forms that appear in these four steps using the `OnlineDataset` thanks to the quadrature-point-wise decomposition of the forms; see Remark 1. We may choose to perform only Steps 1 and 2 if the *a posteriori* error estimate is not needed. The online computational cost scales as $O((N^{\text{pr}})^3 + \tilde{Q}^r)$ for the output prediction and $O((N^{\text{du}})^3 + \tilde{Q}^{\eta,1} + \tilde{Q}^{\eta,2})$ for the error estimate.

5. Numerical results

5.1. Preliminary

In this section, we apply our model reduction framework to two families of two-dimensional problems: flow in a cavity driven by an oscillatory lid and flow past a NACA0012 airfoil at a high angle of attack. For both problems, the flow is governed by the compressible Navier-Stokes equations in entropy variables [6]: we seek the entropy variable

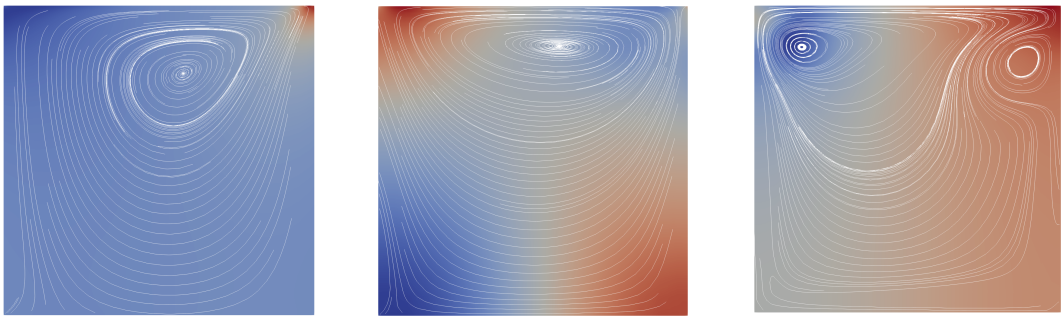
$$u \equiv \left(-\frac{s}{\gamma-1} + \frac{\gamma+1}{\gamma-1} - \frac{\rho e}{p}, \quad \frac{\rho v}{p}, \quad -\frac{\rho}{p}\right)$$

that satisfies the Navier-Stokes equations

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) &= 0, \\ \frac{\partial \rho v}{\partial t} + \nabla \cdot (\rho v \otimes v + pI - \tau) &= 0, \\ \frac{\partial \rho e}{\partial t} + \nabla \cdot ((\rho e + p)v - \tau v - \kappa \nabla T) &= 0, \end{aligned}$$

where $\rho : \Omega \times I \to \mathbb{R}$ is the density, $v : \Omega \times I \to \mathbb{R}^2$ is the velocity, $e : \Omega \times I \to \mathbb{R}$ is the specific internal energy, $s = \log(p/\rho^\gamma)$ is the thermodynamic entropy, $p \equiv (\gamma-1)(\rho e - \rho\|v\|_2^2/2)$ is the pressure, $T = p/(\rho R)$ is the temperature, $\tau = \mu_v(\nabla v + \nabla v^T) + \lambda_v(\nabla \cdot v)I$ is the stress tensor, $\gamma = 1.4$ is the ratio of specific heats, μ_v is the dynamic viscosity, and $\lambda_v = -2/3\mu_v$ is the bulk viscosity coefficient. The entropy (as opposed to conservative) variables symmetrize the flux Jacobian and the viscous tensor, which yields an entropy stable discontinuous Galerkin (DG) method assuming exact integration [34, 6].

We use an adaptive $\mathbb{P}^2$ DG method [18, 4] for the FE approximation. The inviscid numerical flux is computed using Roe's approximate Riemann solver [44], and the viscous terms are discretized using the interior penalty method [4]. The time integration is performed using a three-stage diagonally implicit Runge-Kutta (DIRK) method [1]. The nonlinear system arising in each stage is solved using the Newton's method. The linear system is solved using GMRES [48] that uses an additive Schwarz preconditioner globally and block-ILU(0) with minimum discarded fill ordering locally [43].



(a) Re = 10, $\omega = \pi$, $t = 7/8$ (b) Re = 200, $\omega = \pi$, $t = 1$ (c) Re = 1000, $\omega = 2\pi$, $t = 4$

Figure 1: Instantaneous pressure fields and streamlines for the lid-driven cavity flow.

5.2. *Lid-driven cavity flow: lower Reynolds number*

We first consider a lid-driven cavity flow in a lower Reynolds number regime. Our spatial domain is a unit square $\Omega \equiv (0,1)^2 \subset \mathbb{R}^2$ with a lid on the upper boundary $\Gamma_{\text{lid}} \equiv (0,1)\times\{x_2 = 1\}$. Our time interval is $I \equiv (0, T \equiv 1]$. The flow is driven by an oscillatory lid so that the tangential velocity on Γ_{lid} is

$$v_1(x,t;\mu) = \exp\left(-\frac{0.05}{0.25-(x_1-0.5)^2}\right)\sin(\omega t), \quad x \in \Gamma_{\text{lid}},\ t \in I;$$

we enforce zero normal velocity $v_2 = 0$ on Γ_{lid} and no-slip conditions on all other boundaries $\Omega \setminus \Gamma_{\text{lid}}$. All boundaries are adiabatic. The flow is stationary at the initial time; the initial density and pressure are chosen so that the Mach number associated with the peak lid velocity is $M_{\text{lid}} = 0.7$. Our first parameter is the Reynolds number based on the domain width of 1 and the maximum lid velocity of 1: $\text{Re} \equiv 1/\nu \in [10, 200]$. Our second parameter is the oscillation frequency $\omega \in [\pi/(4T), 2\pi/T]$. We hence have $\mu \equiv (\text{Re}, \omega) \in [10, 200] \times [\pi/(4T), 2\pi/T] \equiv \mathcal{D}_{\text{low}}$. (The subscript "low" indicates that this is the lower Reynolds number case.) The quantity of interest is the time-averaged kinetic energy

$$s(\mu) = J(u(\mu),\mu) = \frac{1}{2T}\int_I \int_\Omega \rho v(\mu,t)\cdot v(\mu,t) dx dt.$$

Figure 1 shows snapshots of solutions for a few different values of Reynolds numbers.

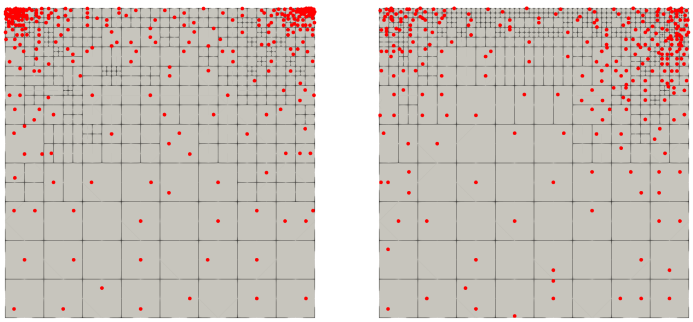
We use the following greedy algorithm settings. We use the adaptive $\mathbb{P}^2$ DG method and the three-stage DIRK method with $\Delta t = 1/32$ for the FE approximation; we refer to Section 5.1 for details. The initial FE space comprises an 8×8 square of elements and $N_h = 1536$ degrees of freedom. The training set Ξ_{train} comprises a 10×10 grid of parameter points that are uniformly distributed over $\mathcal{D}$. Our goal is to achieve the error level of $\delta^{\text{greedy}} = 5 \times 10^{-5}$, which corresponds to $\approx 0.5\%$ error level. Following Remark 25, we set the FE tolerance to $\delta^{\text{fe}} = 5 \times 10^{-6}$, the primal RB tolerance to $\delta^{\text{pr}} = 5 \times 10^{-6}$, the dual RB relative tolerance to $\delta^{\text{du,rel}} = 0.1$, and the EQP tolerances to $\delta^r = \delta^\eta = 5 \times 10^{-6}$ and $\delta^J = 5 \times 10^{-7}$. We solve a single EQP problem for the dual problem and DWR; see Remark 21.

We invoke the greedy algorithm (Algorithm 1) to construct an RB-RQ model. Table 1 summarizes the behavior of the greedy algorithm. We observe that the algorithm converges in five iterations. The dimensions of the primal and dual spaces increase to $N^{\text{pr}} = 22$ and $N^{\text{du}} = 28$, respectively. The number of RQ points increase monotonically with the sizes of the reduced bases. The final RB-RQ approximation comprises $\tilde{Q}^r = 411$ (primal) residual RQ points, $\tilde{Q}^J = 42$ output RQ points, and $\tilde{Q}^\eta = 340$ dual and DWR RQ points. The FE approximation uses $N_h = 9720$ degrees of freedom and $Q_h = 14765$ quadrature points, and hence the RB-RQ approximation is significantly reduced in terms of both the number of degrees of freedom and the number of quadrature points. Figure 2a shows the residual RQ points obtained using the EQP in the final greedy iteration; we observe that the RQ points are clustered towards the upper corners, where the solution exhibits the largest variations in both time and between parameters.

We next discuss the accuracy of the RB-RQ approximations reported in Table 1. In the seventh column, we observe that the maximum error estimate over Ξ_{train}, $\max_{\mu\in\Xi_{\text{train}}} \tilde{\eta}_N(\mu)$, rapidly decays with the greedy iteration. Figure 3b also depicts this rapid convergence behavior. In columns 8–11, we assess the accuracy of the RB-RQ approximation in the "predictive setting" by testing the approximation against the solution computed on $\Xi_{\text{test}} \subset \mathcal{D}$

Table 1: Behavior of the greedy algorithm for the lid-driven cavity flow problem for $\mathcal{D}_{\text{low}}$. The FE problem on the final adapted mesh has $N_h = 9720$ degrees of freedom and $Q_h = 14765$ quadrature points.

						max over Ξ_{train}	max over Ξ_{test}			
iter	N^{pr}	N^{du}	$\tilde{Q}^r$	$\tilde{Q}^J$	$\tilde{Q}^\eta$	$\tilde{\eta}_N$	$\lvert s_h - \tilde{s}_N \rvert$	$\tilde{\eta}_N$	$\lvert s_N - \tilde{s}_N \rvert$	$\lvert \eta_N - \tilde{\eta}_N \rvert$
1	6	4	54	11	50	2.50×10^{-3}	3.21×10^{-3}	2.33×10^{-3}	3.39×10^{-4}	2.10×10^{-5}
2	10	16	136	20	143	5.52×10^{-4}	1.17×10^{-4}	2.08×10^{-4}	1.15×10^{-4}	2.11×10^{-5}
3	14	26	227	31	238	3.98×10^{-4}	2.05×10^{-4}	2.74×10^{-4}	1.79×10^{-5}	6.94×10^{-6}
4	20	26	373	39	327	7.27×10^{-5}	5.48×10^{-5}	5.60×10^{-5}	5.78×10^{-6}	5.87×10^{-6}
5	22	28	411	42	340	3.56×10^{-5}	2.32×10^{-5}	2.80×10^{-5}	3.67×10^{-6}	6.40×10^{-6}



(a) $\mathcal{D}_{\text{low}}$ (b) $\mathcal{D}_{\text{high}}$

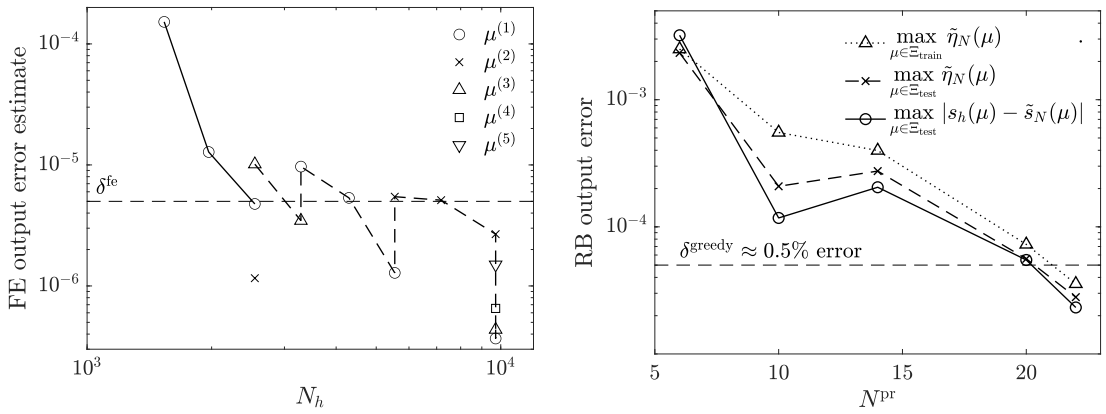
Figure 2: The residual RQ points for the lid-driven cavity flow problems.

comprises $N_{\text{test}} = 20$ randomly chosen test parameter values. We observe that maximum error over the test set $\max_{\mu \in \Xi_{\text{test}}} |s_h - \tilde{s}_N|$ decays rapidly. The error estimate $\max_{\mu \in \Xi_{\text{test}}} \tilde{\eta}_N(\mu)$ also effectively estimates the output error in this predictive setting of $\Xi_{\text{test}} \neq \Xi_{\text{train}}$, as also depicted in Figure 3b. The maximum error in the output due to the use of the residual and output functional RQs is well controlled by the EQP for $N_{\text{greedy}} \geq 4$; i.e., $\max_{\mu \in \Xi_{\text{test}}} |s_N(\mu) - \tilde{s}_N(\mu)| \approx \delta^r = 5 \times 10^{-6}$. In this predictive setting, the RQ error control is not as tight for $N_{\text{greedy}} \leq 3$ due to the relatively poor coverage of the parameter space provided by the EQP parameter set $\Xi_{\text{eqp}} = \Xi_{\text{greedy}}$. Nevertheless, the RQ error is never the dominant source of output error, as shown in column 10 of Table 1. The maximum error in the DWR error estimate due to the use of the RQ rule also behaves similarly. In column 11 of Table 1, we see that it is well controlled by the EQP for $N_{\text{greedy}} \geq 3$; i.e., $\max_{\mu \in \Xi_{\text{test}}} |\eta_N(\mu) - \tilde{\eta}_N(\mu)| \approx \delta^\eta = 5 \times 10^{-6}$.

Having discussed the behavior of the RB-RQ approximation, we next comment on the behavior of the adaptive FE refinement. Figure 3a shows that adaptive mesh refinement is performed for the first two parameter values to meet the specified FE error tolerance. The dimension of the FE space increases from $N_h = 1536$ to 9720 through the adaptive mesh refinement; the number of quadrature points on the final adapted mesh is $Q_h = 14765$. The final adapted mesh is shown in Figure 2a. Much of the refinement is in the vicinity of the upper boundary and the upper corners.

We next discuss the offline training time. Table 2 summarizes the wall-clock time observed on an eight-core computer. All computation times are normalized by the time for a single FE solve on the final adapted mesh, without the solution of the FE dual problem or DWR. We first note that $\approx 50\%$ of the computational time is spent on the adaptive FEM. All adaptive mesh refinements are performed in the first and third greedy iterations; while the other greedy iterations do not require adaptive mesh refinement, we nevertheless must solve the (enriched) FE dual problem and compute the DWR error estimate to ensure the snapshots are sufficiently accurate. The next dominant cost is associated with the EQPs for the residual, output functional, and dual+DWR; these steps take up to $\approx 25\%$ of the computational time in the final greedy iteration, and the cost increases as the size of the EQP training set $\Xi_{\text{eqp}} = \Xi_{\text{greedy}}$ increases. The next dominant cost is the identification of the appropriate primal and dual RB sizes using the recursive algorithm (Algorithm 1, lines 5 and 6); while this cost is non-negligible, it nevertheless is more efficient than using a POD-greedy algorithm that adds only one RB at a time and requires more AFEM solves. The error sampling is a relatively small fraction of the overall cost; the efficient RB-RQ error estimate $\tilde{\eta}_N(\mu)$ enables rapid evaluation of the error at all $|\Xi_{\text{train}}| = 100$ parameter values.

We finally discuss the online evaluation time. Table 3 summarizes the wall-clock speed up observed relative to a single FE solve on the final adapted mesh, without the solution of the FE dual problem or DWR, on the eight-



(a) FE error

(b) RB-RQ error

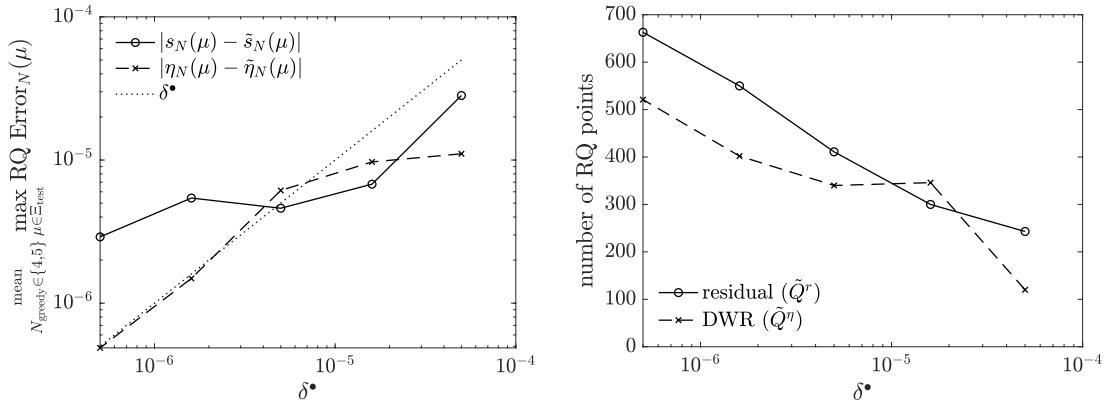
Figure 3: FE and RB-RQ error convergence for the lid-driven cavity flow problem for $\mathcal{D}_{\text{low}}$. In (a), the symbols indicate the parameter value for which the FE problem is solved; the solid and dashed lines correspond to the first and third greedy iterations, in which the mesh is adapted.

			EQP				
iter	AFEM	RBs	residual	output	dual+DWR	sampling	total
1	2.01	0.31	0.04	0.00	0.10	0.50	3.00
2	1.03	0.67	0.11	0.01	0.29	1.09	3.33
3	20.16	3.27	0.58	0.03	2.54	1.71	29.14
4	3.50	2.49	1.19	0.04	3.56	2.40	13.99
5	3.36	2.01	1.59	0.05	4.60	2.70	15.24
	48.5%	14.1%	5.6%	0.2%	17.9%	13.6%	100.0%

Table 2: Wall-clock training time for the lid-driven cavity flow problem for $\mathcal{D}_{\text{low}}$. The times are normalized by the time for a single FE solve (without the DWR error estimate) on the final adapted mesh.

iter	output (only)	output + error estimate
1	325	200
2	166	91
3	113	58
4	74	41
5	62	37

Table 3: Online wall-clock speed up for the lid-driven cavity flow problem for $\mathcal{D}_{\text{low}}$. The times are normalized by the time for a single FE solve (without the DWR error estimate) on the final adapted mesh.



(a) RQ error

(b) RQ points

Figure 4: The RQ errors and the number of RQ points as a function of EQP tolerance $\delta^\bullet$ for the lid-driven cavity flow problem for $\mathcal{D}_{\text{low}}$.

core computer. The RB-RQ approximation, which meets the output error tolerance of ≈0.5%, achieves a speed up of 62 for the (primal) solution and output evaluation; if we wish to also compute the error estimate, which involves the solution of the dual problem and the evaluation of the DWR, the speed up is 37. The cost to evaluate the error estimate is comparable to solving the primal problem for this case as the dual and primal problems are comparable in size ($N^{\text{pr}} = 22$ vs $N^{\text{du}} = 28$ and $\tilde{Q}^r = 411$ vs $\tilde{Q}^\eta = 340$).

5.3. RQ error control via EQP

We will now assess the ability of the EQPs to control the error in the RB-RQ approximations for the lid-driven cavity flow. The greedy algorithm setting is identical to those used in Section 5.2, except that we vary the EQP tolerances from a loose setting of $\delta^r = \delta^\eta = 10\delta^J = 5 \times 10^{-5}$ to a tight setting of $\delta^r = \delta^\eta = 10\delta^J = 5 \times 10^{-7}$.

Figure 4a summarizes the maximum RQ error over a random test set Ξ_{test} that comprises $N_{\text{test}} = 20$ parameter values. We report the average error for the last two iterations, $N_{\text{greedy}} = 4$ and 5. We observe that the RQ error in the output $|s_N(\mu) - \tilde{s}_N(\mu)|$ is well controlled for $\delta^r \geq 5 \times 10^{-6}$. We do not attain a tighter control for $\delta^r < 5 \times 10^{-6}$ because the EQP training parameter set Ξ_{eqp} with just $N_{\text{eqp}} = 4$ and 5 parameter points provide limited coverage of the parameter space. On the other hand, the RQ error in the DWR error estimate $|\eta_N(\mu) - \tilde{\eta}_N(\mu)|$ is well controlled for all EQP tolerances δ^η.

Figure 4b shows the number of residual and DWR RQ points as a function of the EQP tolerances. As expected, the number of RQ points $\tilde{Q}^\bullet$ increase as the tolerance $\delta^\bullet$ is tightened, which results in a commensurate increase in the online computing time. In fact, our standard choice of $\delta^r = \delta^\eta = \delta^{\text{greedy}}/10 = 5\times10^{-6}$, as used in Section 5.2, is somewhat conservative. A looser EQP tolerance would yield a larger online computational speed up for a minimal change in the accuracy.

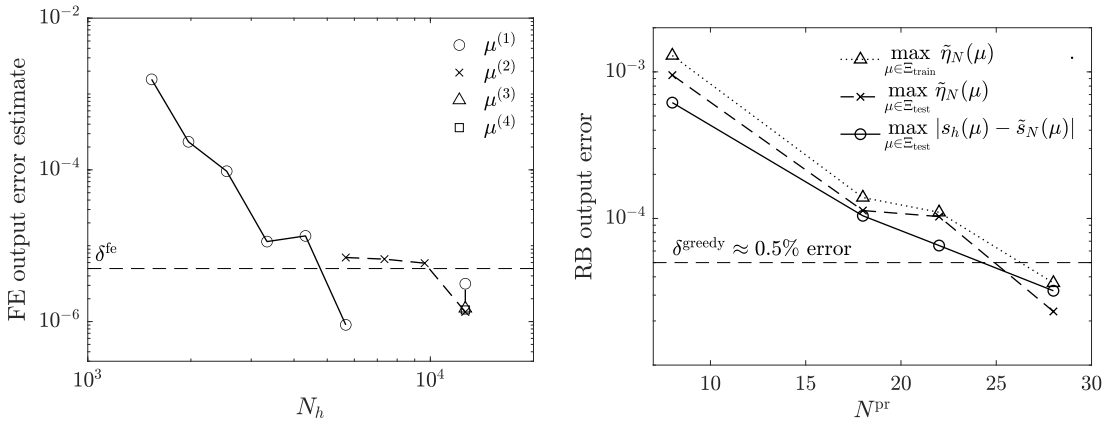
5.4. Lid-driven cavity flow: higher Reynolds number

We now consider the same lid-driven cavity flow problem that we considered in Section 5.2 but in a higher Reynolds number regime. Our new time interval is $I \equiv (0, T \equiv 4]$. Our parameters lie in $\mu \equiv (\text{Re}, \omega) \in [500, 1000] \times [\pi/(4T), 2\pi/T] \equiv \mathcal{D}_{\text{high}}$. Figure 1c shows a snapshot of the solution for the highest Reynolds number case. The solution significantly differs from the lower Reynolds number case.

We invoke the greedy algorithm (Algorithm 1) to construct an RB-RQ model. The greedy algorithm setting is the same as the lower Reynolds number case considered in Section 5.2. For brevity, we focus on observations

Table 4: Behavior of the greedy algorithm for the lid-driven cavity flow problem for $\mathcal{D}_{\text{high}}$. The FE problem on the final adapted mesh has $N_h = 12648$ degrees of freedom and $Q_h = 19165$ quadrature points.

						max over Ξ_{train}	max over Ξ_{test}			
iter	N^{pr}	N^{du}	$\tilde{Q}^r$	$\tilde{Q}^J$	$\tilde{Q}^\eta$	$\tilde{\eta}_N$	$\lvert s_h - \tilde{s}_N \rvert$	$\tilde{\eta}_N$	$\lvert s_N - \tilde{s}_N \rvert$	$\lvert \eta_N - \tilde{\eta}_N \rvert$
1	8	4	83	17	64	1.29×10^{-3}	6.17×10^{-4}	9.50×10^{-4}	2.59×10^{-4}	1.28×10^{-4}
2	18	12	253	40	175	1.39×10^{-4}	1.04×10^{-4}	1.13×10^{-4}	1.41×10^{-5}	1.16×10^{-5}
3	22	20	321	45	246	1.10×10^{-4}	6.54×10^{-5}	1.03×10^{-4}	1.59×10^{-5}	1.31×10^{-5}
4	28	26	436	60	385	3.63×10^{-5}	3.21×10^{-5}	2.32×10^{-5}	3.92×10^{-6}	3.16×10^{-6}



(a) FE error

(b) RB-RQ error

Figure 5: FE and RB-RQ error convergence for the lid-driven cavity flow problem for $\mathcal{D}_{\text{high}}$. In (a), the symbols indicate the parameter value for which the FE problem is solved; the solid and dashed lines correspond to the first and second greedy iterations, in which the mesh is adapted.

that differ from the lower-Reynolds-number lid-driven cavity flow case; we refer to Section 5.2 for more detailed discussions of the general behavior of the algorithm.

Table 4 summarizes the behavior of the greedy algorithms. The greedy algorithm progressively enriches both the primal and dual spaces, and the number of RQ points increases accordingly. We observe rapid decrease in the maximum error over the 100 training parameter values in Ξ_{train}. The maximum error over the 20 test parameter values in $\Xi_{\text{test}} \neq \Xi_{\text{train}}$ also decays rapidly as the number of greedy iterations increases; we also note that the error estimate $\tilde{\eta}_N$ is also effective. Figure 5b depicts these convergence behaviors over Ξ_{train} and Ξ_{test}. We observe that the RQ errors over Ξ_{test} for $N_{\text{greedy}} = 4$ is $\max_{\mu \in \Xi_{\text{test}}} |s_N(\mu) - \tilde{s}_N(\mu)| \approx \delta^r$ for the output and $\max_{\mu \in \Xi_{\text{test}}} |\eta_N(\mu) - \tilde{\eta}_N(\mu)| \approx \delta^\eta$ for the error estimate. We emphasize that this is for the predictive setting of $\Xi_{\text{test}} \neq \Xi_{\text{eqp}}$, and Ξ_{eqp} contains just four parameter values; as the cost of the FE solve dominates the overall computational cost, the ability to achieve tight error control using a small training set is crucial. The residual RQ points for the final RB-RQ model is depicted in Figure 2b.

Figure 5a summarizes the behavior of the FE error with adaptive mesh refinement. For this higher Reynolds number case, adaptive mesh refinement is performed for the first and second parameter values. The final adapted mesh is shown in Figure 2b; we observe refinement in the vicinity of the lid and the upper corners.

Table 5 summarizes offline computation time. We report the wall-clock time observed on an eight-core computer. All times are normalized by the time for a single FE solve on the final adapted mesh, without the solution of the FE dual problem or DWR. The timing breakdown is similar to the lower Reynolds number case considered in Section 5.2. The most dominant cost ($\approx 45\%$) of the offline training are the adaptive FE solves. The cost to identify appropriate primal and dual RB sizes using the recursive algorithm (Algorithm 1, lines 5 and 6) and the EQP costs are similar at $\approx 20\%$ each. The error sampling cost over $|\Xi_{\text{train}}| = 100$ training point is relatively small thanks to the online-efficient error estimate.

Table 6 summarizes the online wall-clock computational time using the eight-core computer. The final model, which achieves an output error of less than 0.5%, achieves online speed up of 68 in the wall-clock time for output evaluation. If we wish to also evaluate the output error estimate, the speed up is 40.

iter	AFEM	RBs	EQP residual	EQP output	EQP dual+DWR	EQP sampling	total
1	4.58	0.70	0.05	0.00	0.17	0.49	6.02
2	10.94	3.41	0.38	0.02	1.01	1.28	17.40
3	3.21	2.96	0.66	0.03	2.03	1.76	11.33
4	3.21	3.84	1.21	0.05	3.78	2.49	15.50
	45.4%	22.6%	4.8%	0.2%	14.5%	12.5%	100.0%

Table 5: Wall-clock training time for the lid-driven cavity flow problem for $\mathcal{D}_{\text{high}}$. The times are normalized by the time for a single FE solve (without the DWR error estimate) on the final adapted mesh.

iter	output (only)	output + error estimate
1	329	202
2	123	77
3	93	56
4	68	40

Table 6: Online wall-clock speed up for the lid-driven cavity flow problem for $\mathcal{D}_{\text{high}}$. The times are normalized by the time for a single FE solve (without the DWR error estimate) on the final adapted mesh.

5.5. *Flow past NACA0012*

We now consider a two-dimensional compressible flow past a NACA0012 airfoil. Our parametrized flow configuration is as follows. The freestream Mach number is fixed at $M_\infty = 0.2$, and the angle of attack is fixed at $\alpha = 20°$. Our parameter is the (chord-based) Reynolds number $\mu \equiv \text{Re} \in [300, 600] \equiv \mathcal{D}$. The initial condition is a snapshot of a (nearly) periodic solution for Re = 450. The quantity of interest is the time-averaged drag over a window of 20 non-dimensionalized time units, which corresponds to approximately two periods of the (nearly) periodic solution. The drag output functional, unlike the kinetic-energy functional for the lid-driven cavity flow, is localized to the airfoil and hence serves as a good test of the goal-oriented error estimation and control framework. Figure 6 shows snapshots of the primal and dual solutions.

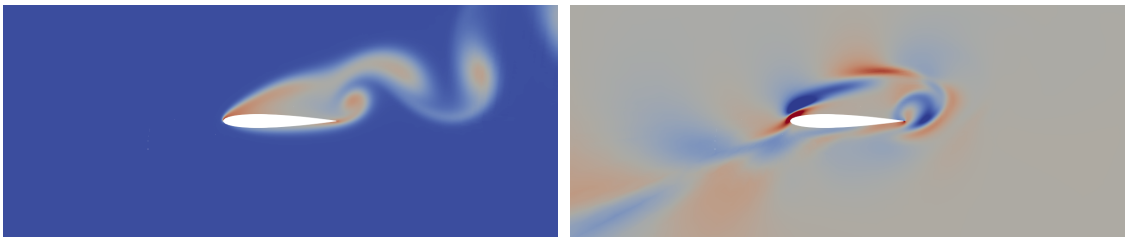
We again use an adaptive $\mathbb{P}^2$ DG method and the three-stage DIRK method with $\Delta t = 1/4$ for the FE approximation, as described in Section 5.1. The initial FE space comprises 1296 elements and $N_h = 31104$ degrees of freedom. The greedy training set comprises 30 parameter points that are uniformly distributed over $\mathcal{D}$. Our goal is to achieve the error level $\delta^{\text{greedy}} = 2 \times 10^{-3}$, which corresponds to ≈0.5% drag error. Following Remark 25, we set the FE tolerance to $\delta^{\text{fe}} = 2 \times 10^{-4}$, the primal RB tolerance to $\delta^{\text{rb}} = 1 \times 10^{-3}$, the dual RB relative tolerance to $\delta^{\text{du,tol}} = 1/10$, and the EQP tolerances to $\delta^r = \delta^\eta = 2 \times 10^{-4}$ and $\delta^J = 2 \times 10^{-5}$.

We invoke the greedy algorithm (Algorithm 1) to construct an RB-RQ model. (For brevity, we again refer to Section 5.2 for general observations and focus on observations that are different from the lower-Reynolds-number lid-driven cavity flow case.) Table 7 summarizes the behavior of the greedy algorithm. The NACA0012 flow exhibits more complicated dynamics than the lid-driven cavity flow. Accordingly, the number of RBs and RQs is substantially higher than the lid-driven cavity flow, requiring $N^{\text{pr}} = 42$, $N^{\text{du}} = 24$, $\tilde{Q}^r = 1267$, $\tilde{Q}^J = 35$, and $\tilde{Q}^\eta = 799$. However, as the FE approximation requires $N_h = 155904$ degrees of freedom and $Q_h = 318036$ quadrature points, the relative reduction is greater for this case. The residual RQ points obtained in the final greedy iteration are shown in Figure 7; the majority of the RQ points are in the wake.

We next comment on the accuracy of the RB-RQ approximation. We observe that the maximum error over the training set $\max_{\mu \in \Xi_{\text{train}}} \tilde{\eta}_N(\mu)$ decays rapidly, as also shown in Figure 8b. The error in the "predictive" setting over the test set Ξ_{test} comprises randomly selected parameter values also decays rapidly. We also confirm that the RB-RQ error estimate $\tilde{\eta}_N(\mu)$ is effective for $N_{\text{greedy}} \geq 2$. We finally comment on the error due to the use of the RQ rules. While the RQ error in the error estimate is well controlled (i.e., $\max_{\mu \in \Xi_{\text{test}}} |\eta_N(\mu) - \tilde{\eta}_N(\mu)| < \delta^\eta$),

Table 7: Convergence of the greedy algorithm for the NACA0012 flow problem. The FE problem on the final adapted mesh has $N_h = 155904$ degrees of freedom and $Q_h = 318036$ quadrature points.

iter	N^{pr}	N^{du}	$\tilde{Q}^r$	$\tilde{Q}^J$	$\tilde{Q}^\eta$	max over Ξ_{train} $\tilde{\eta}_N$	max over Ξ_{test} $\lvert s_h - \tilde{s}_N \rvert$	max over Ξ_{test} $\tilde{\eta}$	max over Ξ_{test} $\lvert s_N - \tilde{s}_N \rvert$	max over Ξ_{test} $\lvert \eta_N - \tilde{\eta}_N \rvert$
1	12	12	226	19	351	5.28×10^{-2}	1.56×10^{-2}	3.32×10^{-2}	3.73×10^{-4}	2.04×10^{-2}
2	24	24	654	31	785	1.78×10^{-2}	7.13×10^{-3}	7.26×10^{-3}	2.83×10^{-3}	1.79×10^{-4}
3	42	24	1267	35	799	6.05×10^{-4}	1.04×10^{-3}	6.00×10^{-4}	6.50×10^{-4}	8.56×10^{-5}



(a) primal solution (entropy), $t = 20$ (b) adjoint (continuity equation), $t = 0$

Figure 6: Snapshots of the primal and adjoint solutions for the NACA0012 flow problem for Re = 450.

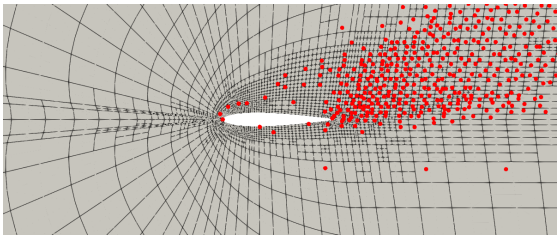


Figure 7: The residual RQ points for the NACA0012 flow problem obtained using EQP in the third greedy iteration. Note that the figure is zoomed into the vicinity of the airfoil.

the RQ error in the output is not as tightly controlled as the lower-Reynolds-number lid-driven cavity flow case (i.e., $\max_{\mu\in\Xi_{\text{test}}} |s_N(\mu) - \tilde{s}_N(\mu)| < 3.5\delta^r$). We again emphasize that this is for the predictive setting of $\Xi_{\text{test}} \neq \Xi_{\text{eqp}}$, and Ξ_{eqp} contains just three parameter points. If a tighter RQ error control over the entire $\mathcal{D}$ is desired, we could increase the size of Ξ_{eqp} and/or use a tighter RQ tolerance at the expense of a higher computational cost. The residual RQ points for the final RB-RQ model are depicted in Figure 7; most RQ points are in the wake. The overall effectiveness of the RB-RQ approximation and error estimates is illustrated in Figure 8b.

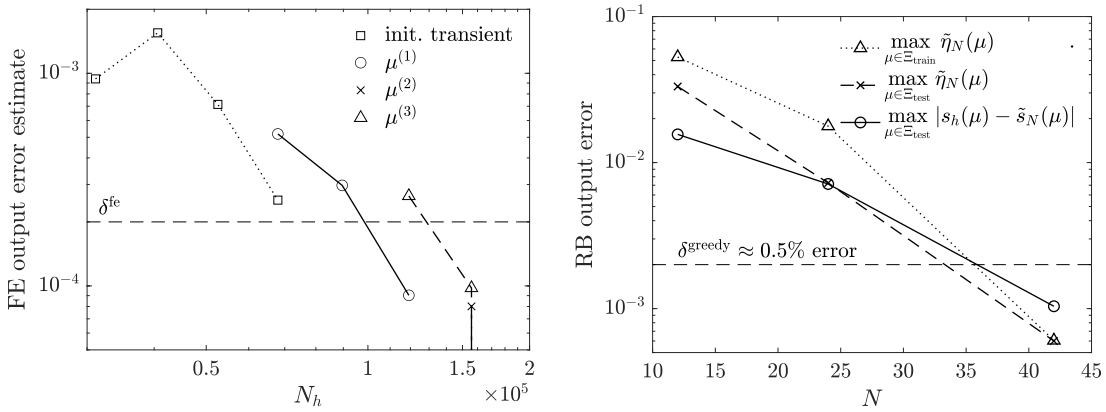
Figure 8a summarizes the convergence behavior of the FE error with adaptive mesh refinement. Figure 7 shows the final adapted mesh. As expected, most of the refinement is in the trailing edge and the wake.

Table 8 summarizes the wall-clock time observed on a 40-core computer. All computational times are normalized by the time for a single FE solve on the final adapted mesh, without the solution of the FE dual problem or DWR. The computation of the adaptive FE solutions and the FE DWR error estimates dominates the overall computational cost at ≈54%. The identification of the appropriate primal and dual RB size using the recursive algorithm (Algorithm 1, lines 5 and 6) also requires ≈27% of the overall computational time. The EQPs comprise a smaller fraction of the computational time than for the lid-driven cavity flows. The error sampling time is negligible, thanks to the efficient RB-RQ error estimate and the relatively small size of the training set $|\Xi_{\text{train}}| = 30$.

Table 9 summarizes the wall-clock speed up observed on a 40-core computer relative to a single FE solve on the final adapted mesh. The RB-RQ approximation, which yields ≈0.5% output error in the predictive setting, achieves a speed up of 328 for the (primal) solution and output evaluation, and a speed up of 240 if we wish to also evaluate the output error estimate. We note that the wall-clock speed up is "conservative" as the relatively small size of the RB-RQ model affects the parallel efficiency on the 40-core computer. Nevertheless, we achieve a wall-clock speed up of two-orders of magnitude in this "conservative" setting.

			EQP				
iter	AFEM	RBs	residual	output	dual+DWR	sampling	total
1	8.02	2.08	0.15	0.01	0.56	0.04	10.87
2	4.54	4.81	0.80	0.03	2.39	0.08	12.97
3	24.21	11.32	3.52	0.09	5.27	0.13	45.56
	54.0%	26.8%	6.6%	0.2%	12.1%	0.4%	100.0%

Table 8: Wall-clock training time for the NACA0012 flow problem. The times are normalized by the time for a single FE solve (without the DWR error estimate) on the final adapted mesh.



(a) FE error

(b) RB-RQ error

Figure 8: FE and RB-RQ error convergence for the NACA0012 flow problem. In (a), the symbols indicate the parameter value for which the FE problem is solved, including the initial transient solve; the solid and dashed lines correspond to the first and third greedy iterations, in which the mesh is adapted.

iter	output (only)	output + error estimate
1	1306	716
2	678	374
3	328	240

Table 9: Online wall-clock speed up for the NACA0012 flow problem. The times are normalized by the time for a single FE solve (without the DWR error estimate) on the final adapted mesh.

6. Summary and perspectives

The present work extends the previous work on model reduction for parametrized time-dependent (nonlinear) PDEs (e.g., [29, 31, 20, 27]) in three areas. Our first contribution is the development of EQPs for time-dependent problems that provide quantitative control of the output error due to hyperreduction. The EQP constraints are informed by an error analysis that appeals to the space-time DWR error representation formula, and the constraints can be used to hyperreduce all quadrature-based spatial discretization and time-marching schemes. Our second contribution is the development of an online-efficient output *a posteriori* error estimate for time-dependent PDEs with general nonlinearities. We approximate the DWR error estimate using a dual RB combined with RQ rules, where the RQ rules are identified using EQP with appropriate accuracy constraints. Our third contribution is the development of a POD-greedy algorithm that simultaneously constructs the finite element mesh, primal and dual RBs, and RQ rules; given a user-prescribed output error tolerance, the training is completely automated. We have demonstrated the efficacy of the framework for two families of unsteady compressible Navier-Stokes flows in a lid-driven cavity and over a NACA0012 airfoil. We believe that systematic construction of ROMs with quantitative output error control, as well as online-efficient *a posteriori* error estimates, are important ingredients needed to apply model reduction techniques to increasingly challenging problems, such as those in aerodynamics.

We also acknowledge that this work focuses on the above contributions and does not explore several topics that are important for model reduction of time-dependent nonlinear PDEs. First, while we employ the discontinuous Galerkin (DG) method and entropy variables for the compressible Navier-Stokes equations to stabilize the "truth" problem, we do not consider an explicit stabilization of the ROM using, e.g., a closure model [54, 35], minimum-residual projection [12], or entropy-stable formulation [15]. Second, we use spatial RBs based on POD and do not consider space-time RB and do not consider space-time formulations based on global snapshots [52] or tensor decomposition [17]. These techniques could further improve the robustness and efficiency of ROMs.

Appendix A. Proof of error estimates

Proof of Proposition 7. We first apply the standard DWR error estimate (e.g., [9, Section 2.2]) to the space-time form of the problem to obtain

$$J(u_N) - J(\tilde{u}_N) = -\langle A(\tilde{u}_N), z_N^{\text{pr}} \rangle + O(\tilde{\delta}^2).$$

We next note that

$$\begin{aligned} J(u_N) - J(\tilde{u}_N) =& -\langle A(\tilde{u}_N), z_N^{\mathrm{pr}}\rangle + \langle A(u_N), z_N^{\mathrm{pr}}\rangle + O(\tilde{\delta}^2) \\ =& -\langle A(\hat{u}_N), z_N^{\mathrm{pr}}\rangle - \langle DA(\hat{u}_N)(\tilde{u}_N - \hat{u}_N), z_N^{\mathrm{pr}}\rangle + O(\|\tilde{u}_N - \hat{u}_N\|_{\mathscr{Y}}^2) \\ &+ \langle A(\hat{u}_N), z_N^{\mathrm{pr}}\rangle + \langle DA(\hat{u}_N)(u_N - \hat{u}_N), z_N^{\mathrm{pr}}\rangle + O(\|u_N - \hat{u}_N\|_{\mathscr{Y}}^2) \\ =& \langle DA(\hat{u}_N)(u_N - \hat{u}_N), z_N^{\mathrm{pr}}\rangle - \langle DA(\hat{u}_N)(\tilde{u}_N - \hat{u}_N), z_N^{\mathrm{pr}}\rangle + O(\|\tilde{u}_N - \hat{u}_N\|_{\mathscr{Y}}^2) + O(\|u_N - \hat{u}_N\|_{\mathscr{Y}}^2) \\ =& \langle DA(\hat{u}_N)\delta\tilde{u}_N, z_N^{\mathrm{pr}}\rangle + O(\tilde{\delta}^2) + O(\hat{\delta}^2), \end{aligned} \tag{A.1}$$

where $\tilde{\delta} \equiv \|\delta\tilde{u}_N\|_{\mathscr{Y}}$ and $\hat{\delta} \equiv \|\delta\hat{u}_N\|_{\mathscr{Y}}$ for $\delta\tilde{u}_N \equiv u_N - \tilde{u}_N$ and $\delta\hat{u}_N \equiv u_N - \hat{u}_N$. Here, the first equality follows from $\langle A(u_N), z_N^{\mathrm{pr}}\rangle = 0$, the second equality follows from the Taylor series expansions of A about $\hat{u}_N$, and the last equality follows from the triangle inequality $\|\tilde{u}_N - \hat{u}_N\|_{\mathscr{Y}}^2 = \|(u_N - \tilde{u}_N) - (u_N - \hat{u}_N)\|_{\mathscr{Y}}^2 \le O(\tilde{\delta}^2) + O(\hat{\delta}^2)$. On the other hand, we note that, $\forall v \in \mathscr{V}$,

$$\begin{aligned} \langle A(\hat{u}_N) - \tilde{A}(\hat{u}_N), v\rangle =& \langle A(u_N), v\rangle + \langle DA(u_N)(\hat{u}_N - u_N), v\rangle + O(\|\hat{u}_N - u_N\|_{\mathscr{Y}}^2) \\ &- \langle \tilde{A}(\tilde{u}_N), v\rangle - \langle D\tilde{A}(\tilde{u}_N)(\hat{u}_N - \tilde{u}_N), v\rangle + O(\|\hat{u}_N - \tilde{u}_N\|_{\mathscr{Y}}^2) \\ =& -\langle DA(\hat{u}_N)\delta\hat{u}_N, v\rangle - \langle D\tilde{A}(\hat{u}_N)(\delta\tilde{u}_N - \delta\hat{u}_N), v\rangle + O(\hat{\delta}^2) + O(\tilde{\delta}^2) \\ =& -\langle D\tilde{A}(\hat{u}_N)\delta\tilde{u}_N, v\rangle - \langle (DA(\hat{u}_N) - D\tilde{A}(\hat{u}_N))\delta\hat{u}_N, v\rangle + O(\hat{\delta}^2) + O(\tilde{\delta}^2), \end{aligned} \tag{A.2}$$

where the first equality follows from the Taylor series expansion about u_N and $\tilde{u}_N$ for the first and second terms, respectively, and the second equality follows from $\langle A(u_N), z_N^{\mathrm{pr}}\rangle = 0$, $\langle \tilde{A}(\tilde{u}_N), z_N^{\mathrm{pr}}\rangle = 0$, the Taylor series expansion of DA and $D\tilde{A}$ about $\hat{u}_N$, and grouping second-order terms. We rearrange the expression to obtain, $\forall v \in \mathscr{V}$,

$$\langle D\tilde{A}(\hat{u}_N)\delta\tilde{u}_N, v\rangle = -\langle A(\hat{u}_N) - \tilde{A}(\hat{u}_N), v\rangle - \langle (DA(\hat{u}_N) - D\tilde{A}(\hat{u}_N))\delta\hat{u}_N, v\rangle + O(\hat{\delta}^2) + O(\tilde{\delta}^2).$$

We now take $v = D\tilde{A}^{-*}(\hat{u}_N)DA^*(\hat{u}_N)z_N^{\mathrm{pr}}$, where $(\cdot)^*$ denotes the formal adjoint to obtain

$$\begin{aligned} &\langle DA(\hat{u}_N)\delta\tilde{u}_N, z_N^{\mathrm{pr}}\rangle \\ &= -\langle DA(\hat{u}_N)D\tilde{A}(\hat{u}_N)^{-1}(A(\hat{u}_N) - \tilde{A}(\hat{u}_N)), z_N^{\mathrm{pr}}\rangle - \langle DA(\hat{u}_N)D\tilde{A}(\hat{u}_N)^{-1}(DA(\hat{u}_N) - D\tilde{A}(\hat{u}_N))\delta\hat{u}_N, z_N^{\mathrm{pr}}\rangle + O(\hat{\delta}^2) + O(\tilde{\delta}^2) \\ &= -\langle A(\hat{u}_N) - \tilde{A}(\hat{u}_N), z_N^{\mathrm{pr}}\rangle + \langle (I - DA(\hat{u}_N)D\tilde{A}(\hat{u}_N)^{-1})(A(\hat{u}_N) - \tilde{A}(\hat{u}_N)), z_N^{\mathrm{pr}}\rangle \\ &\quad + \langle (I - DA(\hat{u}_N)D\tilde{A}(\hat{u}_N)^{-1})DA(\hat{u}_N)\delta\hat{u}_N, z_N^{\mathrm{pr}}\rangle + O(\hat{\delta}^2) + O(\tilde{\delta}^2) \\ &= -\langle A(\hat{u}_N) - \tilde{A}(\hat{u}_N), z_N^{\mathrm{pr}}\rangle + \langle B(A(\hat{u}_N) - \tilde{A}(\hat{u}_N)), z_N^{\mathrm{pr}}\rangle + \langle B(DA(\hat{u}_N)\delta\hat{u}_N), z_N^{\mathrm{pr}}\rangle + O(\hat{\delta}^2) + O(\tilde{\delta}^2), \end{aligned} \tag{A.3}$$

where we have defined $B \equiv I - DA(\hat{u}_N)D\tilde{A}(\hat{u}_N)^{-1}$. We combine (A.1) and (A.3) to obtain

$$|J(u_N) - J(\tilde{u}_N)| \le |\langle A(\hat{u}_N) - \tilde{A}(\hat{u}_N), z_N^{\mathrm{pr}}\rangle| + |\langle B(A(\hat{u}_N) - \tilde{A}(\hat{u}_N)), z_N^{\mathrm{pr}}\rangle| + |\langle B(DA(\hat{u}_N)\delta\hat{u}_N), z_N^{\mathrm{pr}}\rangle| + O(\hat{\delta}^2) + O(\tilde{\delta}^2) \tag{A.4}$$

$$\equiv \text{(I)} + \text{(II)} + \text{(III)} + O(\hat{\delta}^2) + O(\tilde{\delta}^2). \tag{A.5}$$

We now bound the three terms. We appeal to the residual constraint (15) to bound the first term

$$\text{(I)} = |\langle A(\hat{u}_N) - \tilde{A}(\hat{u}_N), z_N^{\mathrm{pr}}\rangle| = \left|\int_I (r(\hat{u}_N, z_N^{\mathrm{pr}}) - \tilde{r}(\hat{u}_N, z_N^{\mathrm{pr}}))dt\right| \le \delta_r. \tag{A.6}$$

To bound the second term, we first appeal to (A.2) to obtain

$$\begin{aligned} \text{(II)} &= |\langle B(A(\hat{u}_N) - \tilde{A}(\hat{u}_N)), z_N^{\mathrm{pr}}\rangle| \\ &= |-\langle BD\tilde{A}(\hat{u}_N)\delta\tilde{u}_N, z_N^{\mathrm{pr}}\rangle - \langle B(DA(\hat{u}_N) - D\tilde{A}(\hat{u}_N))\delta\hat{u}_N, z_N^{\mathrm{pr}}\rangle + O(\hat{\delta}^2) + O(\tilde{\delta}^2)| \\ &= |-\langle BD\tilde{A}(\hat{u}_N)\delta\tilde{u}_N, z_N^{\mathrm{pr}}\rangle - \langle B(DA(\hat{u}_N)D\tilde{A}(\hat{u}_N)^{-1} - I)D\tilde{A}(\hat{u}_N)\delta\hat{u}_N, z_N^{\mathrm{pr}}\rangle + O(\hat{\delta}^2) + O(\tilde{\delta}^2)| \\ &= |-\langle BD\tilde{A}(\hat{u}_N)\delta\tilde{u}_N, z_N^{\mathrm{pr}}\rangle + \langle BBD\tilde{A}(\hat{u}_N)\delta\hat{u}_N, z_N^{\mathrm{pr}}\rangle + O(\hat{\delta}^2) + O(\tilde{\delta}^2)| \end{aligned}$$

We then appeal to the triangle inequality and the definition of operator norms

$$\begin{aligned} \text{(II)} &\le \|B\|_{\mathcal{L}(\mathscr{Y}_N', \mathscr{Y}_N')}\|D\tilde{A}(\hat{u}_N)\|_{\mathcal{L}(\mathscr{Y}_N, \mathscr{Y}_N')}\|\delta\tilde{u}_N\|_{\mathscr{Y}}\|z_N^{\mathrm{pr}}\|_{\mathscr{Y}} + \|B\|^2_{\mathcal{L}(\mathscr{Y}_N', \mathscr{Y}_N')}\|D\tilde{A}(\hat{u}_N)\|_{\mathcal{L}(\mathscr{Y}_N, \mathscr{Y}_N')}\|\delta\hat{u}_N\|_{\mathscr{Y}}\|z_N^{\mathrm{pr}}\|_{\mathscr{Y}} \\ &\le O(\delta_J^2) + O(\tilde{\delta}^2) + O(\hat{\delta}^2), \end{aligned} \tag{A.7}$$

where the operator norms are defined by

$$\|B\|_{\mathcal{L}(\mathcal{Y}'_N,\mathcal{Y}'_N)} = \sup_{w\in\mathcal{Y}'_N}\sup_{v\in\mathcal{Y}_N}\frac{\langle Bw,v\rangle}{\|w\|_{\mathcal{Y}'_N}\|v\|_{\mathcal{Y}_N}} \quad \text{and} \quad \|D\tilde{A}(\hat{u}_N)\|_{\mathcal{L}(\mathcal{Y}_N,\mathcal{Y}'_N)} = \sup_{w\in\mathcal{Y}_N}\sup_{v\in\mathcal{Y}_N}\frac{\langle DA(\hat{u}_N)w,v\rangle}{\|w\|_{\mathcal{Y}_N}\|v\|_{\mathcal{Y}_N}},$$

and the second inequality follows from $\|B\| \le \delta_J$ and the Young's inequality. Finally, we note that the third term is bounded by

$$\text{(III)} = |\langle B(DA(\hat{u}_N)\delta\hat{u}_N), z_N^{\text{pr}}\rangle| \le \|B\|_{\mathcal{L}(\mathcal{Y}'_N,\mathcal{Y}'_N)}\|DA(\hat{u}_N)\|_{\mathcal{L}(\mathcal{Y}_N,\mathcal{Y}'_N)}\|\delta\hat{u}_N\|_{\mathcal{Y}}\|\delta z_N^{\text{pr}}\|_{\mathcal{Y}} \le O(\delta_J^2) + O(\hat{\delta}^2), \tag{A.8}$$

where the first inequality follows from the definition of operator norms and the second inequality follows from the Young's inequality. We substitute (A.6), (A.7), and (A.8) into (A.5) to obtain

$$|J(u_N) - J(\hat{u}_N)| \le \delta_r + O(\delta_J^2) + O(\hat{\delta}^2) + O(\tilde{\delta}^2),$$

which is the desired inequality. □

Proof of Proposition 17. Throughout this proof, all DA, $D\tilde{A}$, DJ, and $D\tilde{J}$ are evaluated about $(\tilde{u}(\mu);\mu)$ and hence we omit the argument for notational convenience. We first note that

$$\begin{aligned} z_N^{\text{du}} - \tilde{z}_N^{\text{du}} &= z_N^{\text{du}} + D\tilde{A}^{-*}D\tilde{J} = DA^{-*}DA^*D\tilde{A}^{-*}(D\tilde{A}^*z_N^{\text{du}} + D\tilde{J}) = DA^{-*}(I - (I - DA^*D\tilde{A}^{-*}))(D\tilde{A}^*z_N^{\text{du}} + D\tilde{J}) \\ &= DA^{-*}(I-B)(D\tilde{A}^*z_N^{\text{du}} + D\tilde{J}), \end{aligned} \tag{A.9}$$

where the first equality follows from $\tilde{z}_N^{\text{du}} = -D\tilde{A}^{-*}D\tilde{J}$, and we have introduced $B \equiv I - DA^*D\tilde{A}^{-*}$. It follows that

$$\begin{aligned} \langle A(\tilde{u}(\mu)), z_N^{\text{du}} - \tilde{z}_N^{\text{du}}\rangle &= \langle A(\tilde{u}(\mu)), DA^{-*}(I-B)(D\tilde{A}^*z_N^{\text{du}} + D\tilde{J})\rangle = \langle (I-B)(D\tilde{A}^*z_N^{\text{du}} + D\tilde{J}), DA^{-1}A(\tilde{u}(\mu))\rangle \\ &= \langle (I-B)(D\tilde{A}^*z_N^{\text{du}} + D\tilde{J}), \psi_N\rangle = \langle D\tilde{A}^*z_N^{\text{du}} + D\tilde{J}, \psi_N\rangle + \langle B(D\tilde{A}^*z_N^{\text{du}} + D\tilde{J}), \psi_N\rangle \equiv \text{(I)} + \text{(II)}, \end{aligned}$$

where the third equality follows from $\psi_N = DA^{-1}A(\tilde{u}(\mu))$. The term (I) is bounded by

$$\begin{aligned} \text{(I)} &= \langle (DA^*z_N^{\text{du}} + DJ) - (D\tilde{A}^*z_N^{\text{du}} + D\tilde{J}), \psi_N\rangle \\ &= \int_I [\overline{r}_h^{\text{du}}(\tilde{u}_N(\mu,t);\psi_N^{\text{du}}(\mu,t), z_N^{\text{du}}(\mu,t);\mu,t) - \tilde{\overline{r}}_h^{\text{du}}(\tilde{u}_N(\mu,t);\psi_N^{\text{du}}(\mu,t), z_N^{\text{du}}(\mu,t);\mu,t)]dt \\ &\quad + m_h(\psi_N^{\text{du}}(\mu,T), z_N^{\text{du}}(\mu,T)) - \tilde{m}_h(\psi_N^{\text{du}}(\mu,T), z_N^{\text{du}}(\mu,T)) - f'_h(\tilde{u}_N(\mu,T);\psi_N^{\text{du}}(\mu,T)) + \tilde{f}'_h(\tilde{u}_N(\mu,T);\psi_N^{\text{du}}(\mu,T)) \le \frac{\delta_{\eta,1}}{2}, \end{aligned}$$

where the first equality follows from "adding" $DA^*z_N^{\text{du}} + DJ = 0$, and the inequality follows from the constraint (31). The term (II) is bounded by

$$\text{(II)} = \langle B(D\tilde{A}^*(z_N^{\text{du}} - \tilde{z}_N^{\text{du}})), \psi_N\rangle \le \|B\|_{\mathcal{L}(\mathcal{Y}_N^{\text{du}\prime},\mathcal{Y}_N^{\text{du}\prime})}\|D\tilde{A}^*\|_{\mathcal{L}(\mathcal{Y}_N^{\text{du}},\mathcal{Y}_N^{\text{du}\prime})}\|z_N^{\text{du}} - \tilde{z}_N^{\text{du}}\|_{\mathcal{Y}}\|\psi_N\|_{\mathcal{Y}} \le O(\delta_J^2) + O(\tilde{\delta}^2),$$

where the first equality follows from $D\tilde{A}^*\tilde{z}_N^{\text{du}} = -D\tilde{J}$, the first equality follows from the constrains and the triangle inequality. □

Appendix B. Fully discrete approximations for DIRK methods

Appendix B.1. Problem description

We present details associated with a fully discrete approximation of the proposed method. We use the DIRK method to illustrate the construction, but the idea readily extends to other time-marching methods. To present the formulation for FE, RB, and RB-RQ approximations in a unified manner, we consider a general system of n equations for the time window $I \equiv (0,T]$. Namely, we find $u : I \to \mathbb{R}^n$ such that

$$\begin{aligned} \frac{d}{dt}M(u(t)) + R(u(t),t) = 0 \quad &\text{in } \mathbb{R}^n \times I, \\ u(t=0) = u^{\text{init}} \quad &\text{in } \mathbb{R}^n, \end{aligned} \tag{B.1}$$

and then evaluate the output

$$s = \int_0^T F(u(t),t)dt + G(u(T));$$

here $M : \mathbb{R}^n \to \mathbb{R}^n$ is the mass function, $R : \mathbb{R}^n \times I \to \mathbb{R}^n$ is the residual function, $F : \mathbb{R}^n \times I \to \mathbb{R}$ is the time-dependent output functional, and $G : \mathbb{R}^n \to \mathbb{R}$ is the terminal-time output functional. In the context of the FE approximation given by $\mathcal{V}_h = \text{span}\{\phi_i^{\text{fe}}\}_{i=1}^{N_h}$, we have $n = N_h$, $u(t) = \mathbf{u}_h(t)$ such that $u_h(t) = \sum_{j=1}^{N_h} \phi_j^{\text{fe}} \mathbf{u}_h(t)_j$, $M(u(t))_i = m(u_h(t), \phi_i^{\text{fe}}; \mu)$ for $i = 1, \ldots, n$, and $R(u(t))_i = r(u_h(t), \phi_i^{\text{fe}}; \mu, t)$ for $i = 1, \ldots, n$. Similarly, in the context of the RB approximation given by $\mathcal{V}_N^{\text{pr}} = \text{span}\{\phi_i^{\text{pr}}\}_{i=1}^{N^{\text{pr}}}$, we have $n = N^{\text{pr}}$, $u(t) = \mathbf{u}_N(t)$ such that $u_N(t) = \sum_{j=1}^{N^{\text{pr}}} \phi_j^{\text{pr}} \mathbf{u}_N(t)_j$, $M(u(t))_i = m(u_h(t), \phi_i^{\text{pr}}; \mu)$ for $i = 1, \ldots, n$, and $R(u(t))_i = r(u_h(t), \phi_i^{\text{pr}}; \mu, t)$ for $i = 1, \ldots, n$.

Appendix B.2. Primal problem

We now consider a fully discrete approximation of (B.1) using an s-stage DIRK method characterized by coefficients $\{a_{ij}\}_{i,j=1}^s$, $\{b_i\}_{i=1}^s$, and $\{c_j\}_{j=1}^s$; see, e.g., [1, 36]. Our time instances are given by $0 \equiv t^0 < t^1 < \cdots < t^K \equiv T$, and the time steps are $t^k \equiv t^k - t^{k-1}$. The application of the s-stage DIRK method to (B.1) yields the following problem: find $\{u^k, \{u_l^k\}_{l=1}^s\}_{k=1}^K$ such that

$$M(u_i^k) - M(u^{k-1}) + \Delta t^k \sum_{j=1}^{s} a_{ij} R(u_j^k, t_j^k) = 0, \quad i = 1, \ldots, s, \quad k = 1, \ldots, K, \tag{B.2}$$

$$M(u^k) - M(u^{k-1}) + \Delta t^k \sum_{j=1}^{s} b_j R(u_j^k, t_j^k) = 0, \quad k = 1, \ldots, K, \tag{B.3}$$

where u^k is the solution at time t^k, u_j^k is the j-th stage solution in the k-th time interval, and $t_j^k = t^{k-1} + c_j \Delta t^k$ for $\Delta t^k \equiv t^k - t^{k-1}$. The DIRK approximation of the output is given by

$$s = \sum_{k=1}^{K} \sum_{j=1}^{s} \Delta t^n b_j F(u_j^k, t_j^k) + G(u^K).$$

For DIRK methods, $a_{ij} = 0$ for $j > i$ by definition. Hence, in the k-th time interval, we successively solve (B.2) for stage states $u_1^k, u_2^k, \ldots, u_s^k$ using Newton's method, and then solve (B.3) for the updated state u^{k+1}. This procedure applies to the semi-discrete FE system (2), the RB system (7), and the RB-RQ system (9).

Appendix B.3. Dual problem

To obtain a fully discrete approximation of the dual problem using the s-stage DIRK method, we first introduce the Lagrangian

$$\begin{aligned} L(U, Z) \equiv & - \sum_{k=1}^{K} \sum_{j=1}^{s} \Delta t^k b_j F(u_j^k, t_j^k) - G(u^K) + z^0 (M(u^0) - M(u^{\text{init}})) \\ & + \sum_{k=1}^{K} \sum_{i=1}^{s} (z_i^k)^T \left(M(u_i^k) - M(u^{k-1}) + \Delta t^k \sum_{j=1}^{s} a_{ij} R(u_j^k, t_j^k) \right) \\ & + \sum_{k=1}^{K} (z^k)^T \left(M(u^k) - M(u^{k-1}) + \Delta t^k \sum_{j=1}^{s} b_j R(u_j^k, t_j^k) \right), \end{aligned}$$

where $U \equiv \{u^k, \{u_i^k\}_{i=1}^s\}_{k=1}^K$ and $Z \equiv \{z^k, \{z_i^k\}_{i=1}^s\}_{k=1}^K$ denote the collection of all state vectors. We differentiate the Lagrangian with respect to the first argument and regroup the terms to obtain

$$\begin{aligned}
L'(U;W,Z) = &-\sum_{k=1}^{K}\sum_{j=1}^{s}\Delta t^k b_j F'(u_j^k,t_j^k)^T w_j^k - G'(u^K)^T w^K + (z^0)^T M'(u^0)w^0 \\
&+\sum_{k=1}^{K}\sum_{i=1}^{s}(z_i^k)^T\left(M'(u_i^k)w_i^k - M'(u^{k-1})w^{k-1} + \Delta t^k\sum_{j=1}^{s}a_{ij}R'(u_j^k,t_j^k)w_j^k\right) \\
&+\sum_{k=1}^{K}(z^k)^T\left(M'(u^k)w^k - M'(u^{k-1})w^{k-1} + \Delta t^k\sum_{j=1}^{s}b_j R'(u_j^k,t_j^k)w_j^k\right) \\
= &((z^K)^T M'(u^K) - G'(u^K)^T)w^K + \sum_{k=0}^{K-1}\left(\left((z^k)^T - (z^{k+1})^T - \sum_{i=1}^{s}(z_i^{k+1})^T\right)M'(u^k)w^k\right) \\
&+\sum_{k=1}^{K}\sum_{j=1}^{s}\left(-\Delta t^k b_j F'(u_j^k,t_j^k)^T + (z_j^k)^T M'(u_j^k) + \Delta t^k\left(\sum_{i=1}^{s}a_{ij}(z_i^k)^T + b_j(z^k)^T\right)R'(u_j^k,t_j^k)\right)w_j^k,
\end{aligned}$$

where $F' : \mathbb{R}^n \times I \to \mathbb{R}^n$ and $G' : \mathbb{R}^n \times \mathbb{R}^n$ are the gradients of F and G, respectively, and $M' : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ and $R' : \mathbb{R}^n \times I \to \mathbb{R}^{n\times n}$ are the Jacobians of M and R, respectively. (We follow the convention where the gradient is a column, instead of a row, vector.) We recall that that the adjoint Z must satisfy $L'(U;W,Z) = 0$ for all W to obtain the following adjoint equations: find $\{z^k, \{z_l^k\}_{l=1}^s\}_{k=1}^K$ such that

$$M'(u_j^k)^T z_j^k + \Delta t^k R'(u_j^k,t_j^k)^T\left(\sum_{i=1}^{s}a_{ij}z_i^k + b_j z^k\right) = \Delta t^k b_j F'(u_j^k,t_j^k), \quad j = 1,\dots,s,\ k = 1,\dots,K, \tag{B.4}$$

$$M'(u^k)^T\left(z^k - z^{k+1} - \sum_{i=1}^{s}(z_i^{k+1})\right) = 0, \quad k = 0,\dots,K-1, \tag{B.5}$$

$$M'(u^K)^T z^K = G'(u^K). \tag{B.6}$$

We again recall that for DIRK methods $a_{ij} = 0$ for $j > i$. Hence, in the k-th time interval, we successively solve (B.4) for stage adjoints $z_s^k, z_{s-1}^k, \dots, z_1^k$ and then solve (B.5) for the updated adjoint z^{k-1}. We start at the final time interval $k = K$ and solve the equation backward in time until we reach $k = 0$. This approach to compute the fully discrete DIRK approximations of the dual problem applies to the semi-discrete FE system (23), RB system (25), RB-RQ system (27), and the enriched FE system (20).

Appendix B.4. DWR error estimate

Given the fully discrete DIRK residual and the associated primal and dual solutions, the *stage-wise* DWR error estimates are given by

$$\eta_i^k \equiv (z_s^k)^T(M(u_i^k) - M(u^{k-1}) + \Delta t^k\sum_{j=1}^{s}a_{ij}R(u_j^k,t_j^k)), \quad i = 1,\dots,s, \quad k = 1,\dots,K, \tag{B.7}$$

$$\eta^k \equiv (z^k)^T(M(u^k) - M(u^{k-1}) + \Delta t^k\sum_{j=1}^{s}b_j R(u_j^k,t_j^k)), \quad k = 1,\dots,K, \tag{B.8}$$

The (global) DWR error estimate is the sum of the stage-wise estimates:

$$\eta \equiv \sum_{k=1}^{K}(\eta^k + \sum_{i=1}^{s}\eta_i^k).$$

This approach to compute the fully discrete DWR error estimates applies to the FE DWR (24), the RB DWR (26), the RB-RQ DWR (28), and the FE DWR for adaptive mesh refinement (22).

Appendix B.5. Dual-of-the-dual (or tangent) problem

A fully discrete approximation of the tangent problem, and specifically (30), for the s-stage DIRK method is the following: find tangent states $\{\psi^k, \{\psi_l^k\}_{l=1}^s\}_{k=1}^K$ such that

$$M'(u_i^k)\psi_i^k - M'(u^{k-1})\psi^{k-1} + \Delta t^k \sum_{j=1}^s a_{ij} R'(u_j^k, t_j^k)\psi_j^k = M(u_i^k) - M(u^{k-1}) + \Delta t^k \sum_{j=1}^s a_{ij} R(u_j^k, t_j^k),$$
$$i = 1, \ldots, s, \quad k = 1, \ldots, K, \tag{B.9}$$

$$M'(u^k)\psi^k - M(u^{k-1})\psi^{k-1} + \Delta t^k \sum_{j=1}^s b_j R'(u_j^k, t_j^k)\psi_j^k = M(u^k) - M(u^{k-1}) + \Delta t^k \sum_{j=1}^s b_j R(u_j^k, t_j^k), \quad k = 1, \ldots, K, \tag{B.10}$$

where $\{u^k, \{u_i^k\}_{i=1}^s\}_{k=1}^K$ is the solution about which the tangent problem is linearized. We note that the right hand side corresponds to the (DIRK-discretized) time-dependent residual. For DIRK methods, $a_{ij} = 0$ for $j > i$ by definition. Hence, in the k-th time interval, we successively solve (B.9) for stage states $\psi_1^k, \psi_2^k, \ldots, \psi_s^k$, and then solve (B.10) for the updated state ψ^{k+1}. This approach to compute the fully discrete DIRK approximations of the tangent problem applies to RB system (30).

Appendix B.6. EQP constraints

The fully discrete approximation of the residual EQP constraint (17) in Definition 10 is given by

$$\Big| \int_{k'} [\overline{r}_h(\hat{u}_N(\mu,t), \Pi_{S_l}\hat{z}_N^{\mathrm{pr}}(\mu,t); \mu, t) - \sum_{q=1}^{Q_h} \rho_q^r \overline{r}_q(\hat{u}_N(\mu,t), \Pi_{S_l}\hat{z}_N^{\mathrm{pr}}(\mu,t); \mu, t) \approx \sum_{k \in \mathrm{ind}(I_{k'})} (\eta^{k,S_l} + \sum_{i=1}^s \eta_i^{k,S_l}) \leq \frac{\delta^r}{L}\frac{|I_k|}{T},$$

where the mapping "ind" identifies DIRK time indices $k \in [1, K]$ that belong to the EQP time interval $k' \in [1, K_{\mathrm{eqp}}]$, and the stage-wise DWR error estimates associated with the projected training dual states $\Pi_{S_l}\hat{z}_N^{\mathrm{pr}}(\mu, \cdot) \in \mathscr{V}_N^{\mathrm{pr}}$ are given by

$$\eta_i^{k,S_l} \equiv \sum_{m \in S_l} (z_s^k)_m (M(u_i^k) - M(u^{k-1}) + \Delta t^k \sum_{j=1}^s a_{ij} R(u_j^k, t_j^k))_m, \quad i = 1, \ldots, s, \quad k = 1, \ldots, K,$$

$$\eta^{k,S_l} \equiv \sum_{m \in S_l} (z^k)_m (M(u^k) - M(u^{k-1}) + \Delta t^k \sum_{j=1}^s b_j R(u_j^k, t_j^k))_m, \quad k = 1, \ldots, K,$$

for $l = 1, \ldots, L$. We use the same procedure to obtain the fully discrete approximations of the output functional EQP constraints (19) in Definition 13, the dual EQP constraints (36) and (37) in Definition 19, and the DWR EQP constraints (38) in Definition 20.

Acknowledgments

The financial support for this work was provided by the Natural Sciences and Engineering Research Council of Canada and the Ontario Graduate Scholarship. Some of the computations were performed on the Niagara supercomputer at the SciNet HPC Consortium. SciNet is funded by the Canada Foundation for Innovation; the Government of Ontario; Ontario Research Fund - Research Excellence; and the University of Toronto.

References

[1] R. Alexander. Diagonally implicit Runge-Kutta methods for stiff O.D.E.'s. *SIAM Journal on Numerical Analysis*, 14(6):1006–1021, 1977.

[2] M. Ali, K. Steih, and K. Urban. Reduced basis methods based upon adaptive snapshot computations. *Advances in Computational Mathematics*, 43:257–294, 2017.

[3] S. S. An, T. Kim, and D. L. James. Optimizing cubature for efficient integration of subspace deformations. *ACM Trans. Graph.*, 27(5):165:1–165:10, Dec. 2008.

[4] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptical problems. *SIAM Journal on Numerical Analysis*, 39(5):1749–1779, 2002.

[5] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera. An "empirical interpolation" method: application to efficient reduced-basis discretization of partial differential equations. *C. R. Acad. Sci. Paris, Ser. I*, 339:667—672, 2004.

[6] T. J. Barth. Numerical methods for gasdynamic systems on unstructured meshes. In D. Kröner, M. Ohlberger, and C. Rohde, editors, *An Introduction to Recent Developments in Theory and Numerics for Conservation Laws*, pages 195–282. Springer-Verlag, 1999.

[7] T. J. Barth. Space-time error representation and estimation in Navier-Stokes calculations. In *Lecture Notes in Computational Science and Engineering*, pages 29–48. Springer Berlin Heidelberg, 2007.

[8] C. E. Baumann and J. T. Oden. An adaptive-order discontinuous Galerkin method for the solution of the Euler equations of gas dynamics. *International Journal for Numerical Methods in Engineering*, 47(1-3):61–73, 2000.

[9] R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numerica*, 10:1–102, may 2001.

[10] P. Benner, S. Gugercin, and K. Willcox. A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM Review*, 57(4):483–531, 2015.

[11] K. Carlberg. Adaptive *h*-refinement for reduced-order models. *International Journal for Numerical Methods in Engineering*, 102(5):1192–1210, nov 2014.

[12] K. Carlberg, M. Barone, and H. Antil. Galerkin v. least-squares Petrov-Galerkin projection in nonlinear model reduction. *Journal of Computational Physics*, 330:693–734, feb 2017.

[13] K. Carlberg, C. Bou-Mosleh, and C. Farhat. Efficient non-linear model reduction via a least-squares Petrov-Galerkin projection and compressive tensor approximations. *International Journal for Numerical Methods in Engineering*, 86(2):155–181, 2011.

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

[15] J. Chan. Entropy stable reduced order modeling of nonlinear conservation laws. *Journal of Computational Physics*, 423:109789, dec 2020.

[16] T. Chapman, P. Avery, P. Collins, and C. Farhat. Accelerated mesh sampling for the hyper reduction of nonlinear computational models. *International Journal for Numerical Methods in Engineering*, 109(12):1623–1654, 2017.

[17] Y. Choi and K. Carlberg. Space-time least-squares Petrov-Galerkin projection for nonlinear model reduction. *SIAM Journal on Scientific Computing*, 41(1):A26–A58, jan 2019.

[18] B. Cockburn. Discontinuous Galerkin methods. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 83(11):731–754, 2003.

[19] C. Daversin-Catty and C. Prud'homme. Simultaneous empirical interpolation and reduced basis method for non-linear problems. *C. R. Acad. Sci. Paris, Ser. I*, 353:1105–1109, 2015.

[20] M. Drohmann, B. Haasdonk, and M. Ohlberger. Reduced basis approximation for nonlinear parametrized evolution equations based on empirical operator interpolation. *SIAM Journal on Scientific Computing*, 34(2):33, 2012.

[21] R. Everson and L. Sirovich. Karhunen-Loève procedure for gappy data. *Journal of the Optical Society of America A, Optics and Image Science*, 12(8):1657–1664, 1995.

[22] C. Farhat, P. Avery, T. Chapman, and J. Cortial. Dimensional reduction of nonlinear finite element dynamic models with finite rotations and energy-based mesh sampling and weighting for computational efficiency. *International Journal for Numerical Methods in Engineering*, 98(9):625–662, 2014.

[23] C. Farhat, T. Chapman, and P. Avery. Structure-preserving, stability, and accuracy properties of the energy-conserving sampling and weighting method for the hyper reduction of nonlinear finite element dynamic models. *International Journal for Numerical Methods in Engineering*, 102(5):1077–1110, 2015. nme.4820.

[24] L. Fick, Y. Maday, A. T. Patera, and T. Taddei. A stabilized POD model for turbulent flows over a range of Reynolds numbers: optimal parameter sampling and constrained projection. *Journal of Computational Physics*, 371:214–243, oct 2018.

[25] K. Fidkowski and D. Darmofal. Review of output-based error estimation and mesh adaptation in computational fluid dynamics. *AIAA Journal*, 49(4):673–694, 2011.

[26] K. J. Fidkowski and Y. Luo. Output-based space–time mesh adaptation for the compressible Navier–Stokes equations. *Journal of Computational Physics*, 230(14):5753 – 5773, 2011.

[27] M. A. Grepl. Certified reduced basis methods for nonaffine linear time-varying and nonlinear parabolic partial differential equations. *Mathematical Models and Methods in Applied Sciences*, 22(03):1150015, mar 2012.

[28] M. A. Grepl, Y. Maday, N. C. Nguyen, and A. T. Patera. Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations. *ESAIM: M2AN*, 41(3):575–605, 2007.

[29] M. A. Grepl and A. T. Patera. A posteriorierror bounds for reduced-basis approximations of parametrized parabolic partial differential equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 39(1):157–181, jan 2005.

[30] B. Haasdonk. Convergence rates of the POD-Greedy method. *ESAIM: Mathematical Modelling and Numerical Analysis*, 47(3):859–873, apr 2013.

[31] B. Haasdonk and M. Ohlberger. Reduced basis method for finite volume approximations of parametrized linear evolution equations. *Mathematical Modelling and Numerical Analysis*, 42(2):277–302, 2008.

[32] J. A. Hernández, M. A. Caicedo, and A. Ferrer. Dimensional hyper-reduction of nonlinear finite element models via empirical cubature. *Comput. Methods Appl. Mesh. Eng.*, 313:687–722, 2017.

[33] J. S. Hesthaven, G. Rozza, and B. Stamm. *Certified reduced basis methods for parametrized partial differential equations*. Springer, 2016.

[34] T. J. R. Hughes, L. P. Franca, and M. Mallet. A new finite element formulation for computational fluid dynamics: I Symmetric forms of the compressible Euler and Navier-Stokes equations and the second law of thermodynamics. *Comput. Methods Appl. Mech. Engrg.*, 54:223–234, 1986.

[35] T. Iliescu and Z. Wang. Variational multiscale proper orthogonal decomposition: Navier-stokes equations. *Numerical Methods for Partial Differential Equations*, 30(2):641–663, dec 2013.

[36] C. A. Kennedy and M. H. Carpenter. Diagonally implicit Runge-Kutta methods for ordinary differential equaitons. A review. Technical Report NASA/TM–2016–219173, NASA, 2016.

[37] D. J. Knezevic, N.-C. Nguyen, and A. T. Patera. Reduced basis approximation and a posteriori error estimation for the parametrized unsteady Boussinesq equations. *Mathematical Models and Methods in Applied Sciences*, 21(07):1415–1442, jul 2011.

[38] K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for parabolic problems. *Numerische Mathematik*, 90(1):117–148, nov 2001.

[39] K. Kunisch and S. Volkwein. Galerkin proper orthogonal dcomposition methods for a gneral euation in fluid dnamics. *SIAM Journal on Numerical Analysis*, 40(2):492–515, jan 2002.
[40] T. Lassila, A. Manzoni, A. Quarteroni, and G. Rozza. Model order reduction in fluid dynamics: challenges and perspectives. In *Reduced Order Methods for Modeling and Computational Reduction*, pages 235–273. Springer International Publishing, 2014.
[41] M. Meyer and H. G. Matthies. Efficient model reduction in non-linear dynamics using the Karhunen-Loève expansion and dual-weighted-residual methods. *Computational Mechanics*, 31(1-2):179–191, may 2003.
[42] A. T. Patera and M. Yano. An LP empirical quadrature procedure for parametrized functions. *Comptes Rendus Mathematique*, 355(11):1161–1167, nov 2017.
[43] P.-O. Persson and J. Peraire. Newton-GMRES preconditioning for discontinuous Galerkin discretizations of the Navier-Stokes equations. *SIAM Journal on Scientific Computing*, 30(6):2709–2733, jan 2008.
[44] P. L. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. *Journal of Computational Physics*, 43(2):357–372, 1981.
[45] D. V. Rovas, L. Machiels, and Y. Maday. Reduced-basis output bound methods for parabolic problems. *IMA Journal of Numerical Analysis*, 26(3):423–445, jul 2006.
[46] G. Rozza, D. B. P. Huynh, and A. T. Patera. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations — Application to transport and continuum mechanics. *Archives of Computational Methods in Engineering*, 15(3):229–275, 2008.
[47] D. Ryckelynck. A priori hyperreduction method: an adaptive approach. *Journal of Computational Physics*, 202(1):346–366, jan 2005.
[48] Y. Saad and M. H. Schultz. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on scientific and statistical computing*, 7(3):856–869, 1986.
[49] M. Schmich and B. Vexler. Adaptivity with dynamic meshes for space-time finite element discretizations of parabolic equations. *SIAM Journal on Scientific Computing*, 30(1):369–393, jan 2008.
[50] L. Sirovich. Turbulence and the dynamics of coherent structures. I. Coherent structures. *Quarterly of Applied Mathematics*, 45(3):561–571, oct 1987.
[51] M. Sleeman. Goal-oriented model reduction for time-dependent nonlinear parametrized partial differential equations. Master's thesis, University of Toronto, 2020.
[52] K. Urban and A. T. Patera. An improved error bound for reduced basis approximation of linear parabolic problems. *Mathematics of Computation*, 83(288):1599–1615, oct 2014.
[53] K. Veroy, C. Prud'homme, D. Rovas, and A. Patera. A posteriori error bounds for reduced-basis approximation of parametrized noncoercive and nonlinear elliptic partial differential equations. In *16th AIAA Computational Fluid Dynamics Conference*. American Institute of Aeronautics and Astronautics, jun 2003.
[54] Z. Wang, I. Akhtar, J. Borggaard, and T. Iliescu. Proper orthogonal decomposition closure models for turbulent flows: a numerical comparison. *Computer Methods in Applied Mechanics and Engineering*, 237-240:10–26, sep 2012.
[55] D. Wirtz, D. C. Sorensen, and B. Haasdonk. A psteriori eror etimation for DEIM rduced nnlinear dnamical sstems. *SIAM Journal on Scientific Computing*, 36(2):A311–A338, jan 2014.
[56] M. Yano. A space-time Petrov-Galerkin certified reduced basis method: application to the Boussinesq equations. *SIAM Journal on Scientific Computing*, 36(1):A232–A266, jan 2014.
[57] M. Yano. A minimum-residual mixed reduced basis method: exact residual certification and simultaneous finite-element reduced-basis refinement. *ESAIM: Mathematical Modelling and Numerical Analysis*, 50(1):163–185, Jan. 2016.
[58] M. Yano. Goal-oriented model reduction of parametrized nonlinear PDEs; application to aerodynamics. *International Journal for Numerical Methods in Engineering*, accepted, 2020.
[59] M. Yano and A. T. Patera. An LP empirical quadrature procedure for reduced basis treatment of parametrized nonlinear PDEs. *Computer Methods in Applied Mechanics and Engineering*, 344:1104–1123, feb 2019.
[60] M. Yano, A. T. Patera, and K. Urban. A space-time hp-interpolation-based certified reduced basis method for Burgers' equation. *Mathematical Models and Methods in Applied Sciences*, 24(09):1903–1935, aug 2014.