An LP empirical quadrature procedure for reduced basis treatment of parametrized nonlinear PDEs

Masayuki Yano[a], Anthony T Patera[b]

[a] *University of Toronto, 4925 Duffein Street, Toronto, ON, M3H 5T6, Canada*
[b] *Room 3-266, MIT, 77 Massachusetts Avenue, Cambridge, MA 02139, USA*

Abstract

We present a model reduction formulation for parametrized nonlinear partial differential equations (PDEs). Our approach builds on two ingredients: reduced basis (RB) spaces which provide rapidly convergent approximations to the parametric manifold; sparse empirical quadrature rules which provide rapid evaluation of the nonlinear residual and output forms associated with the RB spaces. We identify both the RB spaces and the sparse quadrature rules in the offline stage through a greedy training procedure over the parameter domain; the procedure requires the dual norm of the finite element (FE) residual at many training points in the parameter domain, but only very few FE solutions — the snapshots retained in the RB space. The quadrature rules are identified by a linear program (LP) empirical quadrature procedure (EQP) which (i) admits efficient solution by a simplex method, and (ii) directly controls the solution error induced by the approximate quadrature. We demonstrate the formulation for a parametrized neo-Hookean beam: the dimension of the approximation space and the number of quadrature points are both reduced by two orders of magnitude relative to FE treatment, with commensurate savings in computational cost.

Keywords: reduced basis method; parametrized nonlinear PDEs; hyperreduction; empirical quadrature; linear programming; neo-Hookean hyperelasticity

---

1. Introduction

In this work we consider rapid solution of parametrized nonlinear partial differential equations (PDEs) in continuum mechanics. In particular, we wish to evaluate an engineering quantity of interest (output) — such as displacement and local strain energy — for a given configuration parameter (input) — such as material properties and load conditions. Our interest is in many-query scenarios, which require the input-output evaluation for a large number of different parameter values; we will also typically demand real-time response. One approach in this context is model reduction based on offline-online computational decomposition: in the offline stage, we construct, once, a reduced model through a relatively expensive exploration of the parameter domain; in the online stage, we invoke, many times, the reduced model to evaluate the input-output map. This work focuses on model reduction for parametrized *nonlinear* PDEs.

To make the setting mathematically precise, we define an abstract parametrized second-order PDE. We introduce a parameter domain $\mathcal{D} \subset \mathbb{R}^P$, a spatial domain $\Omega \subset \mathbb{R}^d$, and a function (Hilbert)

Email addresses: `myano@utias.utoronto.ca` (Masayuki Yano), `patera@mit.edu` (Anthony T Patera)

space $\mathcal{V}$ such that $H_0^1(\Omega) \subset \mathcal{V} \subset H^1(\Omega)$; we associate to $\mathcal{V}$ an inner product $(\cdot,\cdot)_{\mathcal{V}}$ and induced norm $\|\cdot\|_{\mathcal{V}}$ (equivalent to the $H^1(\Omega)$ norm). We further define an $\mathcal{N}$-dimensional "truth" finite element (FE) approximation space $\mathcal{V}^h \subset \mathcal{V}$ for $\mathcal{N}$ relatively large, and an associated N-dimensional RB space $\mathcal{V}_N \subset \mathcal{V}^h \subset \mathcal{V}$ for $N \ll \mathcal{N}$. Our reduced basis approximation is given as follows: given $\mu \in \mathcal{D}$, find $u_N(\mu) \in \mathcal{V}_N$ such that

$$\int_\Omega \nabla v \cdot A(u_N(\mu);\mu)dx = \int_\Omega vB(u_N(\mu);\mu)dx \quad \forall v \in \mathcal{V}_N, \tag{1}$$

and then evaluate an output

$$s_N(\mu) \equiv \int_\Omega C(u_N(\mu);\mu)dx; \tag{2}$$

here $A : \mathcal{V} \times \mathcal{D} \to (L^2(\Omega))^d$, $B : \mathcal{V} \times \mathcal{D} \to L^2(\Omega)$, and $C : \mathcal{V} \times \mathcal{D} \to L^1(\Omega)$ are respectively the parametrized nonlinear flux, source, and output operators associated with our weak form. The objective of model reduction is to approximate the input-output map $\mu \mapsto s_N(\mu)$ for any $\mu \in \mathcal{D}$ in online complexity $\mathcal{O}(N^q) \ll \mathcal{O}(\mathcal{N})$ for q a small positive integer. (We will henceforth denote complexity I^α, for $0 \le \alpha \le 3$, simply as $\mathcal{O}(I^\cdot)$.) A naive approach to the problem is to directly evaluate the requisite integrals or, more precisely, to approximate these integrals using "truth" FE quadrature rules. The latter will comprise $\mathcal{O}(\mathcal{N})$ points and hence, even though our test and trial spaces are of dimension $N \ll \mathcal{N}$, the overall online complexity is $\mathcal{O}(\mathcal{N}) \gg \mathcal{O}(N^\cdot)$.

A special case arises if the operators are linear (or polynomially nonlinear) in the state and moreover admit a decomposition that is affine in functions of parameters:

$$A(w;\mu) = \sum_{q=1}^{Q^A} \Theta_q^A(\mu) A_q w$$

for some $\Theta_q^A \in L^\infty(\mathcal{D})$ and linear $A_q : \mathcal{V} \to (L^2(\Omega))^d$, $q = 1,\ldots,Q_A$, with similar representations for B and C. This is the setting typically considered by the reduced basis (RB) method; we refer to the review paper [18] for further details. The RB method exploits the affine decomposition to precompute all required integrals in the offline stage and thus reduce the online complexity to $\mathcal{O}(N^\cdot)$ — *rapid response*. In addition, the RB method provides an $\mathcal{O}(N^\cdot)$-complexity error indicator, often error bound, for any $\mu \in \mathcal{D}$ — *reliable response*. Hence, for linear PDEs which admit affine decomposition — often the case in linear elasticity, acoustics, and heat transfer — the RB method provides rapid and reliable solution.

However, many PDEs in continuum mechanics are nonlinear: A, B, and C in (1) and (2) are now nonlinear operators. Over the past decade, much work has been done to extend the efficiency and reliability of the RB method to nonlinear PDEs (see, e.g., [16]) in particular by hyperreduction approaches. A popular approach is interpolation-then-integration: we first approximate, for $\mu \in \mathcal{D}$, the nonlinear sets $A(u(\mu))$, $B(u(\mu))$, and $C(u(\mu))$, by construction of a suitable reduced basis and associated sparse interpolation system; we may then precompute all required integrals in the offline stage. We do not provide here a comprehensive survey of model reduction techniques based on interpolation-then-integration, but rather cite a few representative approaches. The *gappy POD method*, pioneered by Everson and Sirovich in the context of image reconstruction [11], has since been adapted to model reduction in the (Galerkin) gappy POD method [5], the missing point estimation (MPE) method [2], and the Gauss-Newton approximate tensor (GNAT) method [7]. The *empirical interpolation method* (EIM) [3, 13], and closely related methods such as the best point interpolation method (BPIM) [14, 15], also provide a foundation for treatment of nonlinear problems. Although these interpolation-then-integration approaches are quite effective in practice, the objectives of interpolation and integration are arguably quite different. Moreover, while error bounds for certain interpolatory procedures exist (e.g., [10] for EIM), to our knowledge most if not all of the interpolation-then-integration approaches do not quantify the effect of the interpolation error on the solution error (in norm $\|\cdot\|_{\mathcal{V}}$) — the quantity we wish to control; a judicious, or conservative, selection of the interpolation tolerance is often required to ensure that the solution error tolerance is satisfied.

An alternative approach to (1) and (2) is to forgo the interpolation of the nonlinear operators A, B, and C, and directly approximate the integrals with an empirical quadrature rule specialized for the relevant parametrized integrands. An ℓ_2 framework for empirical quadrature procedure of parametrized functions is developed by An et al. [1] and further extended by Farhat et al. [12]. In the ℓ_2 framework, sparsity must be introduced explicitly, either through a heuristic sequential point selection process (as in [1]) or through an approximate ℓ_0 optimization (as in [12]); in both cases, a somewhat challenging non-negative least-squares problem must be addressed. In [19], Ryu and Boyd propose an ℓ_1 framework to empirical quadrature construction; the stronger norm naturally yields quadrature rules that are sparse and furthermore the offline problem can be cast as a linear program (LP) efficiently treated by a simplex method. This LP quadrature framework is further developed in DeVore et al. [9] and extended to the parametric context in [17].

The current paper builds on [17] to provide an LP empirical quadrature procedure (EQP) approach for RB approximation of nonlinear PDEs. The new contribution is threefold. First, we introduce an LP EQP which invokes a preconditioned residual to provide, assuming a sufficiently dense training set, direct control of the RB solution error induced by the approximate quadrature. Second, we devise an efficient offline training procedure which exploits the dual norm of the residual to simultaneously identify an appropriate RB space $\mathcal{V}_N$ and the associated empirical quadrature rule; our approach is related to the simultaneous training procedure for RB and EIM introduced in [8]. Third, we demonstrate the efficiency and error control of our model reduction formulation for a parametrized neo-Hookean beam.

This paper is organized as follows. In Section 2, we review the LP EQP and error analysis first presented in [17]. In Section 3, we present the LP EQP RB method: we describe the formulation, present an error analysis, and then provide a greedy algorithm for offline training. In Section 4, we demonstrate the approach for a parametrized neo-Hookean beam.

2. Linear programming (LP) empirical quadrature procedure (EQP)

In this section, we review the LP EQP first presented in [17].

2.1. Formulation

We first introduce a compact parameter domain $\mathcal{D} \subset \mathbb{R}^P$ and a bounded integration domain $\Omega \subset \mathbb{R}^d$. We next introduce a parametrized vector-valued function $g : \mathcal{D} \times \Omega \to \mathbb{R}^M$ for M a finite positive integer. We now state our integration problem: given $\mu \in \mathcal{D}$, find $I(\mu) \in \mathbb{R}^M$ given by

$$I(\mu) = \int_\Omega g(\mu; x) dx. \tag{3}$$

We will approximate the integral (3) using a quadrature rule.

Towards this end, we introduce a "truth" FE quadrature rule defined by K^h points $\{x_k^h \in \Omega\}_{k=1}^{K^h}$ and the associated weights $\{\rho_k^h \in \mathbb{R}_{+0}\}_{k=1}^{K^h}$; we require the "truth" quadrature to integrate exactly the constant function,

$$\sum_{k=1}^{K^h} \rho_k^h = |\Omega|.$$

We then state the "truth"-quadrature integration problem: given $\mu \in \mathcal{D}$, find $I^h(\mu) \in \mathbb{R}^M$ such that

$$I^h(\mu) = \sum_{k=1}^{K^h} \rho_k^h g(\mu; x_k^h). \tag{4}$$

(Here $\mathbb{R}_+$ and $\mathbb{R}_{+0}$ refer to the positive and non-negative real numbers, respectively.) Following the custom in the reduced-basis community, we use the word "truth" to refer to sufficiently accurate (computable) FE approximations to our (in general, uncomputable) continuous problem (3); we characterize the "truth" FE approximations by a discretization parameter h. For instance, our "truth" quadrature may be a piecewise Gauss quadrature associated with a tessellation of Ω by elements of maximum diameter h.

We wish to approximate (4) for all $\mu \in \mathcal{D}$ by an empirical quadrature rule specifically constructed for the parametrized integrand. To begin, we define an empirical quadrature operator, $\hat{I}^h : \mathbb{R}^{K^h} \times \mathcal{D} \to \mathbb{R}^M$, such that

$$\hat{I}^h(\rho; \mu) \equiv \sum_{k=1}^{K^h} \rho_k g(\mu; x_k^h) \quad \forall \rho \in \mathbb{R}_{+0}^{K^h},\ \forall \mu \in \mathcal{D}.$$

We note that, for the "truth" quadrature weight $\rho^h \in \mathbb{R}^{K^h}$, $\hat{I}^h(\rho^h; \mu) = I^h(\mu)$, $\forall \mu \in \mathcal{D}$. Our goal is to find a quadrature rule $\{x_k^h, \rho_k^\star\}_{k=1}^{K^h}$ that is accurate — $\|I^h(\mu) - \hat{I}^h(\rho^\star; \mu)\|_\infty \leq \epsilon$, $\forall \mu \in \mathcal{D}$, for ϵ small — and sparse — $|\{k \mid \rho_k^\star > 0\}| \ll K^h$. (Here $\|\cdot\|_m$ refers to the usual ℓ_m norm.)

We now introduce an empirical quadrature procedure (EQP) to find $\rho^\star \in \mathbb{R}^{K^h}$. To this end, we specify an accuracy parameter $\delta \in \mathbb{R}_+$ and a parameter training set $\Xi_J^{\text{train}} \equiv \{\mu_j^{\text{train}}\}_{j=1}^J \subset \mathcal{D}$ of size J. We then define our hyperparameter $\nu \equiv \{h, \delta, \Xi_J^{\text{train}}\}$; the hyperparameter summarizes the dependence of our empirical quadrature rule on the underlying "truth" quadrature, as characterized by h, the accuracy parameter δ, and the parameter training set Ξ_J^{train}. We can now define our EQP linear program $\mathsf{LP}_{\text{EQP}}^\nu$: find a *basic feasible solution* $\rho^\star \in \mathbb{R}^{K^h}$ such that

$$\rho^\star = \arg\min_{\rho \in \mathbb{R}^{K^h}} \sum_{k=1}^{K^h} \rho_k$$

subject to K^h non-negativity constraints

$$\rho_k \geq 0, \quad \forall k = 1, \ldots, K^h, \tag{5}$$

an accuracy constraint associated with the constant function,

$$\left||\Omega| - \sum_{k=1}^{K^h} \rho_k\right| \leq \delta, \tag{6}$$

and MJ accuracy constraints associated with functions on the parameter manifold,

$$\|I^h(\mu) - \hat{I}^h(\rho;\mu)\|_\infty \le \delta \quad \forall \mu \in \Xi_J^{\rm train}. \tag{7}$$

We then extract the (strictly) positive quadrature weights $\{x_k^\nu, \rho_k^\nu\}_{k=1}^{K^\nu} \equiv \{\{x_i^h, \rho_i^\star\} \mid i \in \{k \mid \rho_k^\star > 0\}\}$. The resulting empirical quadrature rule is given by

$$I^\nu(\mu) \equiv \hat{I}^h(\rho^\star;\mu) \equiv \sum_{k=1}^{K^\nu} \rho_k^\nu g(\mu; x_k^\nu). \tag{8}$$

We make a few remarks. First, we observe that $\mathtt{LP}_{\rm EQP}^\nu$ is feasible: the "truth" quadrature weight $\rho^h \in \mathbb{R}^{K^h}$ satisfies (5), (6), and (7). Second, numerical evidence suggests that there exist basic feasible solutions of $\mathtt{LP}_{\rm EQP}^\nu$ which are sparse — $K^\nu \ll K^h$; indeed, given that our objective function is the ℓ_1 norm of ρ, the existence of a sparse basic feasible solution is anticipated. Third, numerical evidence suggests that the solution to $\mathtt{LP}_{\rm EQP}^\nu$ achieves rapid convergence — $\max_{\mu \in \mathcal{D}} \|I^h(\mu) - I^\nu(\mu)\|_\infty \to \delta$ rapidly as $J \to \infty$; we provide a proof of convergence, if not rapid convergence, in Section 2.2. Fourth, for a strictly positive (or negative) integral, we may control the relative error instead of the absolute error; we replace (6) and (7) by

$$\begin{aligned} ||\Omega| - \sum_{k=1}^{K^h} \rho_k| &\le \delta|\Omega|, \\ |I^h(\mu) - \hat{I}^h(\rho;\mu)|_i &\le \delta|I^h(\mu)|_i, \quad \forall i = 1,\dots,M, \ \forall \mu \in \Xi_J^{\rm train}, \end{aligned}$$

respectively, where the subscript i denotes the i-th entry of the M-vector.

We finally comment on the offline and online computational cost. The *offline* stage consists of two tasks: the formation of the LP constraint matrix; the solution of the LP. The LP has K^h unknowns, K^h positivity constraints, and $MJ+1$ absolute value bound constraints, one of which is associated with the constant function, (6), and MJ of which are associated with the manifold, (7). (In practice, the absolute value bounds are recast as $2(MJ+1)$ one-sided inequality constraints.) To populate entries of the LP constraint matrix associated with the manifold accuracy constraints, we must evaluate the integrand $g(\cdot;\cdot)$ for all J training parameter values in $\Xi_J^{\rm train}$ and all K^h "truth"-quadrature points $\{x_k^h\}_{k=1}^{K^h}$. We then proceed to find a basic feasible solution to the LP by application of the (dual) simplex method. In the *online* stage, given $\mu \in \mathcal{D}$, we evaluate the integrand $g(\cdot;\cdot)$ for the specified μ at the EQP points $\{x_k^\nu\}_{k=1}^{K^\nu}$, and then evaluate (8); hence the online reduction in computational cost (relative to FE truth quadrature) is $\approx K^\nu/K^h$.

2.2. A priori *error analysis*

We now provide convergence proofs. The results were first provided in [17], which we reproduce here for completeness. We first provide a general error bound.

Proposition 2.1. *For any $\mu \in \mathcal{D}$,*

$$\|I^h(\mu) - I^\nu(\mu)\|_\infty \le \max_{m \in \{1,\dots,M\}} \left[\inf_{\alpha \in \mathbb{R}^J} \delta \sum_{j=1}^J |\alpha_j| + 2|\Omega| |g_m(\mu;\cdot) - \sum_{j=1}^J \alpha_j g_m(\mu_j^{\rm train};\cdot)|_{L^\infty(\Omega)} \right] \tag{9}$$

Proof. See the proof of Lemma 2.1 in [17]. □

To interpret (9), we consider an interpolatory approximation of $g(\mu;\cdot)$ by our snapshots $\sum_{j=1}^{J}\alpha_j g(\mu_j^{\mathrm{train}};\cdot)$. The equation (9) quantifies (first term) the stability and (second term) the best-fit approximation error for any $\mu \in \mathcal{D}$.

We note that Proposition 2.1 is not actionable. However, under a suitable regularity assumption on $g(\cdot;\cdot)$, we arrive at the following error bound.

Proposition 2.2. *Suppose the integrand $g \in (L^\infty(\mathcal{D};L^\infty(\Omega)))^M$ satisfies a Lipschitz condition*

$$\sup_{(\mu',\mu'')\in\mathcal{D}^2} \|g(\mu';\cdot)-g(\mu'';\cdot)\|_{(L^\infty(\Omega))^M} \le L_g\|\mu'-\mu''\|_2$$

for L_g a finite constant and $\|\cdot\|_2$ the usual Euclidean norm in $\mathbb{R}^P$. We in addition let

$$\Delta^\nu \equiv \max_{\mu\in\mathcal{D}} \min_{\mu'\in\Xi_J^{\mathrm{train}}} \|\mu-\mu'\|_2.$$

Then, for a fixed J and any $\mu \in \mathcal{D}$,

$$\|I^h(\mu)-I^\nu(\mu)\|_\infty \le \delta + 2|\Omega|L_g\Delta^\nu.$$

Proof. See the proof of Lemma 2.2 in [17]. □

3. Reduced-basis empirical-quadrature-procedure (RB-EQP) method

The LP EQP introduced in Section 2 identifies a sparse quadrature rule for a family of parametrized integrals and admits an offline-online computational decomposition. We now apply the technique to RB approximations to provide online-efficient solution of parametrized nonlinear PDEs. The formulation is quite general, however, to present the method in a concrete manner, we consider a model problem: parametrized neo-Hookean hyperelasticity.

3.1. Model problem: neo-Hookean hyperelasticity

We first introduce a vector-valued function space $\mathcal{V} \equiv \{v \in (H^1(\Omega))^d \mid v|_{\Gamma_D} = 0\}$ over a Lipschitz domain $\Omega \subset \mathbb{R}^d$ with a non-empty Dirichlet boundary $\Gamma_D \subset \partial\Omega$. We associate displacement fields with $\mathcal{V}$. We then introduce the deformation gradient tensor field $F(w) = \nabla w + I \in (L^2(\Omega))^{d\times d}$ for a given displacement field $w \in \mathcal{V}$. We next recall the strain energy density function for a neo-Hookean solid:

$$\Psi(F(w);\mu) = \frac{\lambda_2(\mu)}{2}\mathrm{tr}(F^T(w)F(w)-I) - \lambda_2(\mu)\log(\det(F(w))) + \frac{\lambda_1(\mu)}{2}\log^2(\det(F(w))), \quad (10)$$

where the Lamé first and second constants, respectively $\lambda_1(\mu)$ and $\lambda_2(\mu)$, are in general functions of the parameter $\mu \in \mathcal{D}$. The differentiation of the potential with respect to the deformation gradient tensor $F(w)$ yields the first Piola-Kirchhoff stress:

$$P(F(w);\mu) = \lambda_2(\mu)(F(w)-F^{-T}(w)) + \lambda_1(\mu)\log(\det(F(w)))F^{-T}(w),$$

where ^{T} denotes algebraic transpose.

The parametrized residual form $r : \mathcal{V} \times \mathcal{V} \times \mathcal{D} \to \mathbb{R}$ associated with a neo-Hookean solid subject to a parametrized external volume load $f : \mathcal{D} \to L^2(\Omega)$ is then given by

$$r(w, v; \mu) \equiv \int_\Omega \eta(w, v; \mu; x) dx, \quad \forall w, v \in \mathcal{V}, \ \forall \mu \in \mathcal{D}, \tag{11}$$

where the integrand $\eta : \mathcal{V} \times \mathcal{V} \times \mathcal{D} \times \Omega \to \mathbb{R}$ is defined as

$$\eta(w, v; \mu; \cdot) \equiv \nabla v(\cdot) : P(\mu; F(w(\cdot))) - v(\cdot) \cdot f(\mu; \cdot), \quad \forall w, v \in \mathcal{V}, \ \forall \mu \in \mathcal{D}. \tag{12}$$

We can now state our model nonlinear problem: given $\mu \in \mathcal{D}$, find $u(\mu) \in \mathcal{V}$ such that

$$r(u(\mu), v; \mu) = 0 \quad \forall v \in \mathcal{V}. \tag{13}$$

We *assume* the problem (13) is well posed and that, for all $\mu \in \mathcal{D}$, we remain on a single solution branch (i.e., no bifurcation). For the example problem considered in Section 4, the FE convergence study indeed supports the assumption of well-posedness of the continuous (infinite-dimensional) problem.

3.2. FE "truth" approximation

We now consider a finite element approximation of (13). To this end, we first provide a "truth" finite element space

$$\mathcal{V}^h \equiv \{v \in \mathcal{V} \mid v|_\kappa \in (\mathbb{P}^p(\kappa))^d, \ \kappa \in \mathcal{T}^h\} \subset \mathcal{V};$$

here $\mathcal{T}^h$ is a tessellation of the domain Ω into non-overlapping elements $\{\kappa\}$, and $\mathbb{P}^p(\kappa)$ is the space of degree-p polynomials over κ. We next introduce a "truth" quadrature rule $\{x_k^h, \rho_k^h\}_{k=1}^{K^h}$ for the piecewise polynomial function space. As the neo-Hookean model exhibits non-polynomial nonlinearity, the exact integration of the residual (11) is not possible; in practice, we choose a rule that exactly integrates piecewise polynomials of degree up to $4p$. We then define a "truth" quadrature approximation of the residual form (11):

$$r^h(w, v; \mu) \equiv \sum_{k=1}^{K^h} \rho_k^h \eta(w, v; \mu; x_k^h), \quad \forall w, v \in \mathcal{V}, \ \forall \mu \in \mathcal{D}, \tag{14}$$

where η is the integrand (12). We can now state our finite element "truth" problem: given $\mu \in \mathcal{D}$, find $u^h(\mu) \in \mathcal{V}^h$ such that

$$r^h(u^h(\mu), v; \mu) = 0 \quad \forall v \in \mathcal{V}^h. \tag{15}$$

We again *assume* the problem is well posed and refer to the solution $u^h(\mu) \in \mathcal{V}^h$ as the "truth" solution.

In practice, we find the solution to (15) by application of a damped Newton's method. We here outline our nonlinear solution strategy for neo-Hookean hyperelasticity; however, we note that the choice of nonlinear solution strategy is highly dependent on the particular form of the governing equation. Given a fixed $\mu \in \mathcal{D}$, we first initialize the solution to $u^{h,i=0} = 0 \in \mathcal{V}^h$. We now describe the steps by which we proceed from iterate i to iterate $i+1$. We first solve for the Newton update: find $\Delta u^{h,i} \in \mathcal{V}^h$ such that

$$(r^h)'(u^{h,i}, \Delta u^{h,i}, v; \mu) = -r^h(u^{h,i}, v; \mu) \quad \forall v \in \mathcal{V}^h,$$

where $(r^h)'(u^{h,i}, \Delta u^{h,i}, v; \mu)$ is the Gâteaux derivative of $r^h(\cdot, v; \mu) : \mathcal{V}^h \to \mathbb{R}$ at $u^{h,i} \in \mathcal{V}^h$ in the direction $\Delta u^{h,i} \in \mathcal{V}^h$; note the bilinear from $(r^h)'(u^{h,i}, \cdot, \cdot; \mu) : \mathcal{V}^h \times \mathcal{V}^h \to \mathbb{R}$ induces the Jacobian. We next update the solution according to

$$u^{h,i+1} = u^{h,i} + \alpha^i \Delta u^{h,i}$$

for a damping parameter $\alpha^i \in (0, 1]$. We choose the damping parameter such that the total potential energy decreases after each damped Newton update: we require $\Psi^h_{\text{total}}(u^{h,i+1}(\mu); \mu) < \Psi^h_{\text{total}}(u^{h,i}(\mu); \mu)$, where the ("truth"-quadrature approximation of the) total potential energy is given by

$$\Psi^h_{\text{total}}(w; \mu) \equiv \sum_{k=1}^{K^h} \rho^h_k \left[\Psi(F(w); \mu) - w \cdot f(\mu)\right](x^h_k), \quad \forall w \in \mathcal{V},\ \forall \mu \in \mathcal{D},$$

for $\Psi : \mathbb{R}^N \times \mathcal{D} \to \mathbb{R}$ the strain energy density function (10). We terminate the Newton iteration when the residual is sufficiently small.

We remark on the computational cost and storage for the "truth" solve. (We here distinguish between $\mathcal{O}(\mathcal{N})$ and $\mathcal{O}(K^h)$ to differentiate the requirements which originate in the dimension of the approximation space, $\mathcal{N}$, and in the number of quadrature points, K^h; of course, in practice $\mathcal{O}(\mathcal{N}) = \mathcal{O}(K^h)$.) Throughout the FE computation, we exploit the sparsity of the Jacobian matrix. In each Newton iteration, the evaluation of the residual and Jacobian requires $\mathcal{O}(K^h)$ operations. Then, the solution of the linear system requires $\mathcal{O}(\mathcal{N}^\cdot)$ operations: the exponent is unity in the most favorable case (e.g. multigrid); more generally, the exponent is greater than unity (but less than 2 for $d \leq 3$ for a sparse direct solver). Finally, the evaluation of the total potential energy to find the damping parameter requires $\mathcal{O}(K^h)$ operations. These three steps are repeated several times — typically 5 to 15 times — to achieve convergence. The storage requirement is $\mathcal{O}(\mathcal{N}^\cdot)$, where the dominant storage is the Jacobian: $\mathcal{O}(\mathcal{N})$ for an iterative solver; $\mathcal{O}(\mathcal{N}^\cdot)$, with an exponent greater than unity (but less than 4/3 for $d \leq 3$), for a sparse direct solver.

3.3. "Truth"-quadrature RB approximation

We now consider a reduced basis approximation of (15). Towards this end, we introduce a reduced basis space $\mathcal{V}_N \equiv \text{span}\{u^h(\mu)\}_{\mu \in \Xi^{\text{RB}}_N} \subset \mathcal{V}^h$ associated with a snapshot parameter set $\Xi^{\text{RB}}_N \subset \mathcal{D}$ of size N. We choose snapshot parameter sets and the associated reduced basis approximations that are hierarchical: $\Xi^{\text{RB}}_{N=1} \subset \cdots \subset \Xi^{\text{RB}}_{N=N_{\max}}$ and $\mathcal{V}_{N=1} \subset \cdots \subset \mathcal{V}_{N=N_{\max}}$; we defer to Section 3.6 the discussion of a systematic procedure to select a hierarchical parameter set. We can now state the "truth"-quadrature reduced basis problem: given $\mu \in \mathcal{D}$, find $u_N(\mu) \in \mathcal{V}_N$ such that

$$r^h(u_N(\mu), v; \mu) = 0 \quad \forall v \in \mathcal{V}_N. \tag{16}$$

We *assume* the problem is well posed and refer to $u_N(\mu) \in \mathcal{V}_N$ as the RB solution or, more explicitly, the "truth"-quadrature RB solution.

We now introduce a discrete form of the RB problem. By way of preliminaries, we introduce a $\mathcal{V}$-orthonormal basis $\{\phi_i\}_{i=1}^{N_{\max}}$ such that $\mathcal{V}_N = \text{span}\{\phi\}_{i=1}^N$, $N = 1, \ldots, N_{\max}$; the $\mathcal{V}$-orthonomality implies $(\phi_i, \phi_j)_\mathcal{V} = \delta_{ij}$ for δ_{ij} the Kroncker delta. We then introduce an associated operator $Z_N : \mathbb{R}^N \to \mathcal{V}_N$ which maps a generalized coordinate $\mathbf{v} \in \mathbb{R}^N$ to the associated field $v = Z_N \mathbf{v} \equiv \sum_{i=1}^N \mathbf{v}_i \phi_i \in \mathcal{V}_N$. Throughout this work, we denote the generalized coordinate associated with a function in a reduced basis space by a corresponding boldface letter; e.g. $\mathbf{v} \in \mathbb{R}^N$, $\mathbf{w} \in \mathbb{R}^N$, and

$\mathbf{z} \in \mathbb{R}^N$ are associated with $v = Z_N\mathbf{v} \in \mathcal{V}_N$, $w = Z_N\mathbf{w} \in \mathcal{V}_N$, and $z = Z_N\mathbf{z} \in \mathcal{V}_N$, respectively. We note that because the basis is $\mathcal{V}$-orthonormal, $(w,v)_{\mathcal{V}} = (Z_N\mathbf{w}, Z_N\mathbf{v})_{\mathcal{V}} = (\mathbf{w},\mathbf{v})_2$ and $\|v\|_{\mathcal{V}} = \|Z_N\mathbf{v}\|_{\mathcal{V}} = \|\mathbf{v}\|_2$; here $(\cdot,\cdot)_2$ and $\|\cdot\|_2$ are the Euclidean ℓ^2 inner product and norm given by $(\mathbf{w},\mathbf{v})_2 = \mathbf{v}^T\mathbf{w}$ and $\|\mathbf{w}\|_2 = \sqrt{(\mathbf{w},\mathbf{w})_2}$, respectively.

We next define a discrete residual operator associated with the "truth"-quadrature RB approximation (16), $R_N^h : \mathbb{R}^N \times \mathcal{D} \to \mathbb{R}^N$, such that

$$(R_N^h(\mathbf{w};\mu))_i = r^h(Z_N\mathbf{w}, \phi_i;\mu) \quad \forall i = 1,\ldots,N,\ \forall \mathbf{w} \in \mathbb{R}^N,\ \forall \mu \in \mathcal{D}, \tag{17}$$

where $(R_N^h(\mathbf{w};\mu))_i$ denotes the i-th component of $R_N^h(\mathbf{w};\mu) \in \mathbb{R}^N$. We can then state the discrete counterpart of the "truth"-quadrature RB problem (16): given $\mu \in \mathcal{D}$, find $\mathbf{u}_N(\mu) \in \mathbb{R}^N$ such that

$$R_N^h(\mathbf{u}_N(\mu);\mu) = 0 \quad \text{in } \mathbb{R}^N; \tag{18}$$

we note $u_N(\mu) = Z_N\mathbf{u}_N(\mu)$. We find the solution to (18) using the damped Newton's method described in Section 3.2 adapted to the RB approximation. The latter requires the RB Jacobian: the (i,j)-entry of the RB Jacobian $J_N^h(\mathbf{w};\mu) \in \mathbb{R}^{N\times N}$ is given by

$$(J_N^h(\mathbf{w};\mu))_{(i,j)} \equiv \left.\frac{\partial (R_N^h(\mathbf{v};\mu))_i}{\partial \mathbf{v}_j}\right|_{(\mathbf{w},\mu)} = (r^h)'(Z_N\mathbf{w}, \phi_j, \phi_i;\mu), \quad \forall \mathbf{w} \in \mathbb{R}^N,\ \forall \mu \in \mathcal{D}. \tag{19}$$

Note that in both (18) and (19) we invoke the "truth" FE quadrature.

We remark on the online computational cost and storage for the "truth"-quadrature RB solve. (There are several implementations, all quite similar, of the "truth"-quadrature RB approach; we present the simplest complexity estimates, in particular since in actual practice the "truth"-quadrature RB serves only for purposes of exposition and is never implemented.) In each Newton iteration, the evaluation of the RB residual and Jacobian by the "truth" quadrature requires $\mathcal{O}(N^2K^h)$ operations. Then, the solution of the (dense) linear system requires $\mathcal{O}(N^3) \ll \mathcal{O}(\mathcal{N}^\cdot)$ operations. Finally, the evaluation of the total potential energy to find the damping parameter requires $\mathcal{O}(K^h)$ operations. As in the FE case in Section 3.2, these three steps are repeated several times to achieve convergence. We note that the cost to evaluate the residual, Jacobian, and total potential energy for the "truth"-quadrature RB approximation scales as K^h — hence expensive; on the other hand, the linear solve cost is reduced, since the RB solution is sought in the $N(\ll \mathcal{N})$ dimensional RB space. The online storage is dominated by the RB basis function values $\{\phi_i\}_{i=1}^N$ and gradients $\{\nabla\phi_i\}_{i=1}^N$ evaluated at the "truth" quadrature points; the online storage is thus $\mathcal{O}((d+1)NK^h)$ for $N \ll K^h$. In summary, we observe that the "truth"-quadrature RB approximation does not meet the online complexity goal of model reduction set forth in the Introduction: the online computational cost and storage are both $\mathcal{O}(\mathcal{N}, K^h)$.

3.4. RB-EQP approximation

We now consider an EQP approximation of (18). To this end, we first introduce an empirical quadrature operator now specialized for the RB residual: $\hat{R}_N^h : \mathbb{R}^{K^h} \times \mathbb{R}^N \times \mathcal{D} \to \mathbb{R}^N$ is given by

$$(\hat{R}_N^h(\rho;\mathbf{w};\mu))_i \equiv \sum_{k=1}^{K^h} \rho_k \eta(w \equiv Z_N\mathbf{w}, \phi_i;\mu;x_k^h), \quad \forall i = 1,\ldots,N,\ \forall \rho \in \mathbb{R}_{+0}^{K^h},\ \forall \mathbf{w} \in \mathbb{R}^N,\ \forall \mu \in \mathcal{D}; \tag{20}$$

note that, for the "truth" quadrature weight $\rho^h \in \mathbb{R}^{K^h}$, $\hat{R}_N^h(\rho^h; \mathbf{w}; \mu) = R_N^h(\mathbf{w}; \mu)$, $\forall \mathbf{w} \in \mathbb{R}^N$, $\forall \mu \in \mathcal{D}$. Our goal is to find a sparse quadrature rule $\{x_k^h, \rho_k^\star\}_{k=1}^{K^h}$ such that the associated RB solution satisfies our error tolerance (in the $\|\cdot\|_\mathcal{V}$ norm) for any $\mu \in \mathcal{D}$. We emphasize that we place our error tolerance on the RB solution, and not on the intermediate integrals required to form the RB system.

Our LP EQP for the RB approximation is a specialization of the general procedure outlined in Section 2.1. To this end, we specify i) an accuracy parameter $\delta \in \mathbb{R}_{0+}$, ii) a parameter training set $\Xi_J^{\text{train}} \equiv \{\mu_j^{\text{train}} \in \mathcal{D}\}_{j=1}^J$ of size J, and iii) the associated state training set $U_J^{\text{train}} \equiv \{u_N^{\text{train}}(\mu) \in \mathcal{V}_N\}_{\mu \in \Xi_J^{\text{train}}} \equiv \{Z_N \mathbf{u}_N^{\text{train}}(\mu)\}_{\mu \in \Xi_J^{\text{train}}}$; we identify the hyperparameter ν as $\nu \equiv \{h, \delta, \Xi_J^{\text{train}}, U_J^{\text{train}}\}$. We now define our linear program $\mathtt{LP}_{\text{EQP}}^{\nu,N}$: find a basic feasible solution $\rho^\star \in \mathbb{R}^{K^h}$ such that

$$\rho^\star = \underset{\rho_k \in \mathbb{R}^{K^h}}{\arg\min} \sum_{k=1}^{K^h} \rho_k$$

subject to K^h non-negativity constraints

$$\rho_k \geq 0, \quad \forall k = 1, \ldots, K^h,$$

the constant-function accuracy constraint

$$||\Omega| - \sum_{k=1}^{K^h} \rho_k| \leq \delta, \tag{21}$$

and NJ manifold accuracy constraints

$$\|(J_N^h(\mathbf{u}_N^{\text{train}}(\mu); \mu))^{-1} \hat{R}_N^h(\rho; \mathbf{u}_N^{\text{train}}(\mu); \mu)\|_\infty \leq \delta, \quad \forall \mu \in \Xi_{\text{train}}. \tag{22}$$

We then extract the (strictly) positive quadrature weights $\{x_k^\nu, \rho_k^\nu\}_{k=1}^{K^\nu} \equiv \{\{x_i^h, \rho_i^\star\} \mid i \in \{k \mid \rho_k^\star > 0\}\}$. The LP EQP approximation to the RB residual (17) is given by

$$(R_N^\nu(\mathbf{w}; \mu))_i \equiv (\hat{R}_N^h(\rho^\star; \mathbf{w}; \mu))_i \equiv \sum_{k=1}^{K^\nu} \rho_k^\nu \eta(Z_N \mathbf{w}, \phi_i; \mu; x_k^\nu), \quad \forall i = 1, \ldots, N, \ \forall \mathbf{w} \in \mathbb{R}^N, \ \forall \mu \in \mathcal{D}. \tag{23}$$

Note that M of the general LP EQP formulation of Section 2 instantiates here to N.

We make a few remarks. First, the motivation behind the form of the accuracy constraints (22) may not be clear at the moment; the constraints are in fact informed by an error analysis presented in Section 3.5. Second, the accuracy constraints (22) are linear in $\rho \in \mathbb{R}^{K^h}$ — as required for the LP — since $\hat{R}_N^h(\rho; u_N^{\text{train}}(\mu); \mu)$ is linear in $\rho \in \mathbb{R}^{K^h}$; a more explicit representation is provided in Appendix A. Third, we have not addressed the selection of the state training set $U_J^{\text{train}} \equiv \{u_N^{\text{train}}(\mu)\}_{\mu \in \Xi_J^{\text{train}}}$; the issue will be addressed in Section 3.6. Fourth, we may control the relative, as opposed to absolute, error by replacing (21) and (22) by

$$||\Omega| - \sum_{k=1}^{K^h} \rho_k| \leq \delta|\Omega|,$$

$$|(J_N^h(\mathbf{u}_N^{\text{train}}(\mu); \mu))^{-1} \hat{R}_N^h(\rho; \mathbf{u}_N^{\text{train}}(\mu); \mu)|_i \leq \delta \|\mathbf{u}_N^{\text{train}}(\mu)\|_2, \quad \forall i = 1, \ldots, N, \ \forall \mu \in \Xi_{\text{train}},$$

respectively; in our numerical results, we appeal to this relative from of the LP EQP. Finally, we note that the LP has K^h unknowns, K^h positivity constraints, and $NJ + 1$ absolute-value bound constraints.

We can now state the RB-EQP problem: given $\mu \in \mathcal{D}$, find $\mathbf{u}_N(\mu) \in \mathbb{R}^N$ such that

$$R_N^\nu(\mathbf{u}_N^\nu(\mu); \mu) = 0 \quad \text{in } \mathbb{R}^N. \tag{24}$$

The equivalent problem in functional form is as follows: given $\mu \in \mathcal{D}$, find $u_N^\nu(\mu) \in \mathcal{V}$ such that

$$r^\nu(u_N^\nu(\mu), v; \mu) = 0 \quad \forall v \in \mathcal{V}_N, \tag{25}$$

where

$$r^\nu(w, v; \mu) \equiv \sum_{k=1}^{K^\nu} \rho_k^\nu \eta(w, v; \mu; x_k^\nu) \quad \forall w, v \in \mathcal{V},\ \forall \mu \in \mathcal{D}. \tag{26}$$

We note that $u_N^\nu(\mu) = Z_N \mathbf{u}_N^\nu(\mu)$. We find the solution using the damped Newton's method described in Section 3.2, but with all the ingredients — the residual, Jacobian, and total potential energy — evaluated using the EQP. The Jacobian $J_N^\nu(\mathbf{w}; \mu) \in \mathbb{R}^{N \times N}$ and the total potential energy are given by

$$(J_N^\nu(\mathbf{w}; \mu))_{(i,j)} \equiv \left. \frac{\partial (R_N^\nu(\mathbf{v}; \mu))_i}{\partial \mathbf{v}_j} \right|_{(\mathbf{w}, \mu)} = (r^\nu)'(Z_N \mathbf{w}, \phi_j, \phi_i; \mu) = \sum_{k=1}^{K^\nu} \rho_k^\nu \eta'(Z_N \mathbf{w}, \phi_j, \phi_i; \mu; x_k^\nu),$$

and

$$\Psi_{\text{total}}^\nu(w; \mu) \equiv \sum_{k=1}^{K^\nu} \rho_k^\nu \left[\Psi(F(w); \mu) - w \cdot f(\mu) \right] (x_k^\nu),$$

respectively.

We now remark on the online storage requirement. The evaluation of the residual (26) (and the associated Jacobian and total potential energy) requires i) the EQP weights $\{\rho_k^\nu\}_{k=1}^{K^\nu}$, ii) the RB values evaluated at the EQP quadrature points $\{\{\phi_i(x_k^\nu)\}_{i=1}^N\}_{k=1}^{K^\nu}$, and iii) the RB gradient values evaluated at the EQP quadrature points $\{\{\{\frac{\partial \phi_i}{\partial x_j}(x_k^\nu)\}_{i=1}^N\}_{k=1}^{K^\nu}\}_{j=1}^d$. The total online storage is hence $K^\nu(1 + N(d+1)) \ll \mathcal{O}(\mathcal{N}, K^h)$; we note a significant reduction in the storage requirement relative to the "truth"-quadrature RB approximation, which requires the RB values and gradients evaluated at all "truth" quadrature points.

We next remark on the online computational cost. In each Newton iteration, the evaluation of the RB-EQP residual and Jacobian requires $\mathcal{O}(N^2 K^\nu) \ll \mathcal{O}(K^h)$ operations. Then, the solution of the (dense) linear system requires $\mathcal{O}(N^3) \ll \mathcal{O}(\mathcal{N}^\cdot)$ operations. Finally, the evaluation of the total potential energy to find the damping parameter requires $\mathcal{O}(K^\nu) \ll \mathcal{O}(K^h)$ operations. In short, with respect to the "truth"-quadrature RB approximation, the RB-EQP online cost, thanks to the EQP quadrature rule, replaces all appearances of K^h with $K^\nu \ll K^h$. The RB-EQP method hence achieves the online complexity goal of model reduction set forth in the Introduction: the online computational cost and storage are $\mathcal{O}(N^\cdot, K^\nu)$ and in particular independent of $\mathcal{N}$ and K^h.

3.5. *A priori error analysis*

Our *a priori* error analysis builds on the Brezzi-Rappaz-Raviart theorem [4]. In earlier work [20], the Brezzi-Rappaz-Raviart framework is applied within the RB context to develop *a posteriori* error estimators relative to the "truth" finite element approximation; in our current context, we apply Brezzi-Rappaz-Raviart to understand and control the RB error associated with empirical quadrature. Towards that end, we first reproduce the Brezzi-Rappaz-Raviart theorem in the form presented in [6] and with specialization to the $\ell^2(\mathbb{R}^N)$ Euclidean space relevant to our context:

Lemma 3.1. *[Brezzi-Rappaz-Raviart theorem [4, 6] for $\ell^2(\mathbb{R}^N)$] We introduce a C^1 mapping $G: \mathbb{R}^N \to \mathbb{R}^N$, $\mathbf{v} \in \mathbb{R}^N$ such that the Jacobian $DG(\mathbf{v}) \in \mathbb{R}^{N\times N}$ is non-singular, and constants ϵ, γ, and $L(\alpha)$ such that*

$$\|G(\mathbf{v})\|_2 \le \epsilon, \tag{27}$$

$$\|DG(\mathbf{v})^{-1}\|_2 \le \gamma, \tag{28}$$

$$\sup_{\mathbf{w}\in\bar{B}(\mathbf{v},\alpha)} \|DG(\mathbf{v}) - DG(\mathbf{w})\|_2 \le L(\alpha), \tag{29}$$

for $\bar{B}(\mathbf{v},\alpha) \equiv \{\mathbf{z} \mid \|\mathbf{z}-\mathbf{v}\|_2 \le \alpha\}$. Suppose

$$2\gamma L(2\gamma\epsilon) \le 1. \tag{30}$$

Then, for all $\beta \ge 2\gamma\epsilon$ such that $\gamma L(\beta) < 1$, there exists a unique solution $\mathbf{u} \in \mathbb{R}^N$ that satisfies $G(\mathbf{u}) = 0$ in the ball $\bar{B}(\mathbf{v},\beta)$. Moreover, $\forall \mathbf{w} \in \bar{B}(\mathbf{v}, 2\gamma\epsilon)$,

$$\|\mathbf{w} - \mathbf{u}\|_2 \le 2\gamma\|G(\mathbf{w})\|_2. \tag{31}$$

Proof. See [6]. □

We now specialize Lemma (3.1) to analyze the error in the RB-EQP solution $\mathbf{u}_N^\nu(\mu)$ with respect to the "truth"-quadrature RB solution $\mathbf{u}_N(\mu)$ and, hence, $u_N^\nu(\mu)$ with respect to $u_N(\mu)$.

Proposition 3.2. *We first fix $\mu \in \mathcal{D}$. We then introduce $\hat{\mathbf{u}}_N(\mu) \in \mathbb{R}^N$ such that*

$$\|\mathbf{u}_N(\mu) - \hat{\mathbf{u}}_N(\mu)\|_2 \le \epsilon^{\text{train}} \tag{32}$$

for some $\epsilon^{\text{train}} \in \mathbb{R}_{+0}$ and such that $J_N^h(\hat{\mathbf{u}}_N(\mu))$ is non-singular. Suppose the EQP-approximated residual form and Jacobian satisfy

$$\|J_N^h(\hat{\mathbf{u}}_N(\mu);\mu)^{-1} R_N^\nu(\hat{\mathbf{u}}_N(\mu);\mu)\|_\infty \le \delta_R, \tag{33}$$

$$\|J_N^h(\hat{\mathbf{u}}_N(\mu);\mu)^{-1} J_N^\nu(\hat{\mathbf{u}}_N(\mu);\mu) - I\|_{\max} \le \delta_J, \tag{34}$$

for some $\delta_R \in \mathbb{R}_{+0}$ and

$$\delta_J \in [0, 1/N). \tag{35}$$

We in addition define

$$L(\alpha) \equiv 2 \sup_{\mathbf{w}\in\bar{B}(\hat{\mathbf{u}}_N(\mu),\alpha)} \|J_N^h(\hat{\mathbf{u}}_N(\mu);\mu)^{-1} J_N^\nu(\mathbf{w};\mu) - I\|_2. \tag{36}$$

Suppose

$$L\left(\frac{2N^{1/2}\delta_R}{1-N\delta_J}\right) \le \frac{1-N\delta_J}{2}. \tag{37}$$

Then, for all $\beta \ge 2N^{1/2}\delta_R/(1-N\delta_J)$ such that $L(\beta) \le 1 - N\delta_J$, there exists a unique solution $\mathbf{u}_N^\nu(\mu) \in \mathbb{R}^N$ to (24) in the ball $\bar{B}(\hat{\mathbf{u}}_N, \beta)$. Moreover,

$$\|u_N(\mu) - u_N^\nu(\mu)\|_{\mathcal{V}} = \|\mathbf{u}_N(\mu) - \mathbf{u}_N^\nu(\mu)\|_2 \le \frac{2N^{1/2}\delta_R}{1-N\delta_J} + \epsilon^{\text{train}}. \tag{38}$$

Proof. Throughout the proof, we suppress the parameter μ in the forms and solutions for notational brevity. Our plan is to identify appropriate conditions of Lemma 3.1 given by (27)–(29) in the context of the RB-EQP approximation, and then invoke Lemma 3.1. Towards this end, we set $G(\cdot)$ and $\mathbf{v}$ in Lemma 3.1 equal to the preconditioned RB-EQP residual operator and $\hat{\mathbf{u}}_N$ which satisfies (32), respectively: $G(\cdot) \equiv J_N^h(\hat{\mathbf{u}}_N)^{-1} R_N^\nu(\cdot)$ and $\mathbf{v} \equiv \hat{\mathbf{u}}_N$. We then invoke a) $\forall \mathbf{v} \in \mathbb{R}^N$, $\|\mathbf{v}\|_2 \le N^{1/2} \|\mathbf{v}\|_\infty$, and b) the assumption (33), to obtain ϵ in (27):

$$\|G(\mathbf{v})\|_2 = \|J_N^h(\hat{\mathbf{u}}_N)^{-1} R_N^\nu(\hat{\mathbf{u}}_N)\|_2 \le N^{1/2} \|J_N^h(\hat{\mathbf{u}}_N)^{-1} R_N^\nu(\hat{\mathbf{u}}_N)\|_\infty \le N^{1/2} \delta_R \equiv \epsilon.$$

We similarly invoke i) the assumption (34), ii) condition (35), iii) $\forall A \in \mathbb{R}^{N \times N}$, $\|A\|_2 \le N \|A\|_{\max}$ and iv) $\forall A \in \mathbb{R}^{N \times N}$ such that $\|A\|_2 < 1$, $(I + A)^{-1}$ exists, and $\|(I + A)^{-1}\|_2 \le (1 - \|A\|_2)^{-1}$, to obtain γ in (28):

$$\begin{aligned}
\|DG(\mathbf{v})^{-1}\|_2 &= \|(J_N^h(\hat{\mathbf{u}}_N)^{-1} J_N^\nu(\hat{\mathbf{u}}_N))^{-1}\|_2 = \|(I + J_N^h(\hat{\mathbf{u}}_N)^{-1} J_N^\nu(\hat{\mathbf{u}}_N) - I)^{-1}\|_2 \\
&\le (1 - \|J_N^h(\hat{\mathbf{u}}_N)^{-1} J_N^\nu(\hat{\mathbf{u}}_N) - I\|_2)^{-1} \le (1 - N \|J_N^h(\hat{\mathbf{u}}_N)^{-1} J_N^\nu(\hat{\mathbf{u}}_N) - I\|_{\max})^{-1} \\
&\le (1 - N\delta_J)^{-1} \equiv \gamma;
\end{aligned}$$

here, i), ii), and iii) imply $\|J_N^h(\hat{\mathbf{u}}_N)^{-1} J_N^\nu(\hat{\mathbf{u}}_N) - I\|_2 \le 1$, which in turn allows us to invoke iv) to obtain the first inequality. We then invoke a) the triangle inequality, and b) $\hat{\mathbf{u}}_N \in \bar{B}(\hat{\mathbf{u}}_N, \alpha)$, to specialize (29) to (36):

$$\begin{aligned}
\sup_{\mathbf{w} \in \bar{B}(\mathbf{v}, \alpha)} \|DG(\mathbf{v}) - DG(\mathbf{w})\|_2 &= \sup_{\mathbf{w} \in \bar{B}(\hat{\mathbf{u}}_N, \alpha)} \|J_N^h(\hat{\mathbf{u}}_N)^{-1} J_N^\nu(\hat{\mathbf{u}}_N) - J_N^h(\hat{\mathbf{u}}_N)^{-1} J_N^\nu(\mathbf{w})\|_2 \\
&\le \|J_N^h(\hat{\mathbf{u}}_N)^{-1} J_N^\nu(\hat{\mathbf{u}}_N) - I\|_2 + \sup_{\mathbf{w} \in \bar{B}(\hat{\mathbf{u}}_N, \alpha)} \|J_N^h(\hat{\mathbf{u}}_N)^{-1} J_N^\nu(\mathbf{w}) - I\|_2 \\
&\le 2 \sup_{\mathbf{w} \in \bar{B}(\hat{\mathbf{u}}_N, \alpha)} \|J_N^h(\hat{\mathbf{u}}_N)^{-1} J_N^\nu(\mathbf{w}) - I\|_2 \equiv L(\alpha).
\end{aligned}$$

Then, (37) is a restatement of (30) for the above specializations of ϵ, γ, and $L(\cdot)$. We next note that $\mathbf{u}_N^\nu \in \mathbb{R}^N$ is a solution to $G(\mathbf{u}_N^\nu) \equiv J_N^h(\hat{\mathbf{u}}_N)^{-1} R_N^\nu(\mathbf{u}_N^\nu) = 0$ by (24). The uniqueness of the solution $\mathbf{u}_N^\nu$ in $\bar{B}(\hat{\mathbf{u}}_N, \beta)$ for $\beta \ge 2\gamma\epsilon = 2N^{1/2}\delta_R/(1 - N\delta_J)$ such that $L(\beta) \le 1/\gamma = 1 - N\delta_J$ then follows directly from the uniqueness statement of Lemma 3.1. Finally, we invoke a) the triangle inequality, b) condition (32), and c) (31) for $\mathbf{w} \equiv \hat{\mathbf{u}}_N \in \bar{B}(\hat{\mathbf{u}}_N, 2\gamma\epsilon)$, to obtain

$$\|\mathbf{u}_N - \mathbf{u}_N^\nu\|_2 \le \|\mathbf{u}_N - \hat{\mathbf{u}}_N\|_2 + \|\hat{\mathbf{u}}_N - \mathbf{u}_N^\nu\|_2 \le \epsilon^{\text{train}} + 2\gamma \|G(\hat{\mathbf{u}}_N)\|_2 \le \epsilon^{\text{train}} + \frac{2N^{1/2}\delta_R}{1 - N\delta_J},$$

which is the desired bound (38). □

Corollary 3.3. *Suppose all conditions of the Proposition 3.2 are satisfied with* (32)*,* (33)*, and* (37) *replaced by*

$$\begin{aligned}
\|\mathbf{u}_N(\mu) - \hat{\mathbf{u}}_N(\mu)\|_2 &\le \epsilon^{\text{train}} \|\hat{\mathbf{u}}_N(\mu)\|_2, \\
\|J_N^h(\hat{\mathbf{u}}_N(\mu); \mu)^{-1} R_N^\nu(\hat{\mathbf{u}}_N(\mu); \mu)\|_\infty &\le \delta_R \|\hat{\mathbf{u}}_N(\mu)\|_2, \\
L\left(\frac{2N^{1/2}\delta_R \|\hat{\mathbf{u}}_N(\mu)\|_2}{1 - N\delta_J}\right) &\le \frac{1 - N\delta_J}{2}.
\end{aligned}$$

respectively; we now further assume $\epsilon^{\text{train}} < 1$. Then, for $\|\hat{\mathbf{u}}_N(\mu)\| > 0$ and $\epsilon^{\text{train}} < 1$, the relative error in the RB-EQP solution with respect to the "truth"-quadrature RB solution is bounded by

$$\frac{\|u_N(\mu) - u_N^\nu(\mu)\|_{\mathcal{V}}}{\|u_N(\mu)\|_{\mathcal{V}}} = \frac{\|\mathbf{u}_N(\mu) - \mathbf{u}_N^\nu(\mu)\|_2}{\|\mathbf{u}_N(\mu)\|_2} \leq \frac{1}{1-\epsilon^{\text{train}}}\left[\frac{2N^{1/2}\delta_R}{1-N\delta_J} + \epsilon^{\text{train}}\right]. \tag{39}$$

Proof. In Proposition 3.2, we first replace δ_R and ϵ^{train} with $\delta_R\|\hat{\mathbf{u}}_N(\mu)\|_2$ and $\epsilon^{\text{train}}\|\hat{\mathbf{u}}_N(\mu)\|_2$, respectively, to obtain $\|\mathbf{u}_N - \mathbf{u}_N^\nu\|_2/\|\hat{\mathbf{u}}_N\|_2 \leq 2N^{1/2}\delta_R/(1-N\delta_J) + \epsilon^{\text{train}}$. We then invoke $\|\mathbf{u}_N\|_2 \geq \|\hat{\mathbf{u}}_N\|_2 - \|\hat{\mathbf{u}}_N - \mathbf{u}_N\|_2 \geq \|\hat{\mathbf{u}}_N\|_2(1-\epsilon^{\text{train}})$ which implies, for $\|\hat{\mathbf{u}}_N\|_2 > 0$ and $\epsilon^{\text{train}} < 1$, $\|\mathbf{u}_N - \mathbf{u}_N^\nu\|_2/\|\mathbf{u}_N\|_2 \leq (1-\epsilon^{\text{train}})^{-1}\|\mathbf{u}_N - \mathbf{u}_N^\nu\|_2/\|\hat{\mathbf{u}}_N\|_2$ to obtain (39). □

We note that $\hat{\mathbf{u}}_N(\mu)$ plays the role of $\mathbf{u}_N^{\text{train}}(\mu)$ in (22). Since we wish to control the error of $\mathbf{u}_N^\nu(\mu)$ relative to $\mathbf{u}_N(\mu)$, we know from Proposition 3.2 that we prefer the choice ($\hat{\mathbf{u}}_N(\mu)$ and hence) $\mathbf{u}_N^{\text{train}}(\mu) = \mathbf{u}_N(\mu)$ such that $\epsilon^{\text{train}} = 0$. In actual practice, and as discussed in Section 3.6, we choose a slightly different training strategy which improves computational performance at apparently very little compromise in EQP error control.

Proposition 3.2 (and its relative-error counterpart Corollary 3.3) identifies sufficient conditions to control the difference between the RB-EQP solution (25) and the "truth"-quadrature RB solution (16). We first consider the case in which the empirical quadrature is the "truth" quadrature (i.e., $\rho = \rho^h$) and furthermore $\epsilon^{\text{train}} = 0$. In this case, the conditions (33) and (34) are trivially satisfied with $\delta_R = 0$ and $\delta_J = 0$, respectively. Moreover, the condition (37) is satisfied because $L(2N^{1/2}\delta_R/(1-N\delta_J)) = L(0) = 0 \leq 1/2$. Hence $\|u_N(\mu) - u_N^\nu(\mu)\|_{\mathcal{V}} = 0$, $\forall \mu \in \mathcal{D}$, as expected.

We next observe that $\texttt{LP}_{\text{EQP}}^{\nu,N}$ outlined in Section 3.4 incorporates (33) as the accuracy constraint (22), but does not include the conditions (34) and (37). We exclude these two conditions from the $\texttt{LP}_{\text{EQP}}^{\nu,N}$ for the following reasons. First, the condition (34) would correspond in our LP to N^2J constraints; we thus prefer to omit (34) and hence retain many fewer constraints, in particular $NJ(+1)$. By way of justification, we note that the tolerance δ_J in (34) only weakly influences the error bound (38) for $N\delta_J \ll 1$; at least for the parameterized PDEs we have considered, numerical evidence suggests that the condition $\delta_R \ll 1$ is sufficient to control the error $\|u_N(\mu) - u_N^\nu(\mu)\|_{\mathcal{V}}$. Second, the condition (37) does not directly influence the error bound: the condition only identifies a neighborhood of $\hat{u}_N(\mu)$ over which our Brezzi-Rappaz-Raviart error bound, based on a linearized analysis, holds. We finally note that, although the conditions (34) and (37) are omitted in our formulation, (34) could be readily incorporated into the LP problem (but at the considerable expense of $\mathcal{O}(N^2)$ scaling), and (37) can be confirmed *a posteriori* for certain nonlinearities.

We also observe that $\texttt{LP}_{\text{EQP}}^{\nu,N}$ outlined in Section 3.4 incorporates the constant-function condition (21) even though the latter does not appear in Proposition 3.2. The constant-function accuracy constraint may contribute to satisfaction of (34) and is furthermore a plausible condition for any quadrature scheme; in the context of our residual integration objective we require this (or similar) additional constraint to ensure a nontrivial (generic) solution to the LP.

3.6. Greedy algorithm: simultaneous RB and EQP training

We presented in Section 3.4 a procedure to construct an RB-EQP approximation *assuming* we have already identified a reduced basis $\{\phi_i\}_{i=1}^N$ and EQP state training set $U_J^{\text{train}} = \{u_N^{\text{train}}(\mu)\}_{\mu\in\Xi_J^{\text{train}}}$. *If* the offline computational cost could be entirely neglected, we could identify $\{\phi_i\}_{i=1}^{N_{\max}}$ and U_J^{train} as follows. We first identify an effective reduced basis $\{\phi_i\}_{i=1}^{N_{\max}}$ in two steps: solve the "truth" FE problem (15) for each $\mu \in \Xi_J^{\rm train} \subset \mathcal{D}$; apply the POD or the strong greedy algorithm to the "truth" data. We then identify the EQP state training set for each $N \in \{1,\dots,N_{\max}\}$ by solving the "truth"-quadrature RB problem (16) and setting $U_J^{\rm train} = \{u_N^{\rm train}(\mu) = u_N(\mu)\}_{\mu\in\Xi_J^{\rm train}}$. This approach requires $J \equiv |\Xi_J^{\rm train}|$ "truth" FE solves to identify $\{\phi_i\}_{i=1}^N$ and NJ "truth"-quadrature RB solves to identify $U_J^{\rm train}$. (We recall from Section 3.3 that the "truth"-quadrature RB solve is almost as expensive as the "truth" FE solve.) While the emphasis of model reduction is often on the online efficiency, the offline training cost cannot be neglected in engineering practice, and J "truth" FE solves and NJ "truth"-quadrature RB solves may be prohibitive. Here we propose an algorithm that identifies $\{\mathcal{V}_N\}_{N=1}^{N_{\max}}$ using only $N_{\max}$ "truth" solves (and no "truth"-quadrature RB solves).

Our greedy algorithm is outlined in Algorithm 1. Algorithm 1 considers control of the quadrature contribution to absolute solution error. We may also choose to control the quadrature contribution to relative solution error, assuming $u^h(\mu) \neq 0$, $\forall \mu \in \mathcal{D}$: we need only replace line 6 with $\mu^{(N)} = \arg\sup_{\mu\in\Xi_J^{\rm train}} \|r^h(u_{N-1}^\nu(\mu),\cdot;\mu)\|_{(V^h)'}/\|u_{N-1}^\nu(\mu)\|_{\mathcal{V}}$, and invoke in line 16 the relative-error version of $\texttt{LP}_{\rm EQP}^{\nu,N}$; we invoke this relative-error version of the greedy algorithm in our numerical example of Section 4. The inputs to the greedy algorithm, for either absolute error control or relative error control, are the parameter training set $\Xi_J^{\rm train} \equiv \{\mu_j^{\rm train}\}_{j=1}^J \subset \mathcal{D}$ of size J, the EQP tolerance $\delta \in \mathbb{R}_{+0}$, and the RB residual tolerance $\epsilon_{\rm RB} \in \mathbb{R}_+$; the parameter training set $\Xi_J^{\rm train}$ is used for both RB and EQP training. The outputs are the reduced basis $\{\phi_n\}_{n=1}^{N_{\max}}$ and an empirical quadrature rule $\{x_k^\nu, \rho_k^\nu\}_{k=1}^{K^\nu}$.

In each iteration we perform two distinct tasks: the identification of the next snapshot parameter $\mu^{(N)} \in \Xi_J^{\rm train}$, and the evaluation of the associated "truth" snapshot; the construction of the associated EQP quadrature rule. The first task, the identification of an appropriate snapshot parameter $\mu^{(N)}$, is performed as follows (line 6). For each $\mu \in \Xi_J^{\rm train}$, we compute the RB-EQP solution $u_{N-1}^\nu(\mu)$ and the associated dual norm of the residual with respect to the "truth" space $\mathcal{V}^h$, $\|r^h(u_{N-1}^\nu(\mu),\cdot;\mu)\|_{(\mathcal{V}^h)'} \equiv \sup_{v\in\mathcal{V}^h} |r^h(u_{N-1}^\nu(\mu),v;\mu)|/\|v\|_{\mathcal{V}}$. The evaluation of J "truth" residual dual norms makes the procedure more expensive than the (original) weak greedy algorithm [18] which appeals to an online-efficient *a posteriori* error bound. However, in the context of nonlinear equations, the evaluation of a "truth" residual dual norm is substantially less expensive than the evaluation of the "truth" *solution*. In particular, the former requires a residual evaluation and solution of a (sparse and parameter-independent) SPD linear system. In contrast, the latter requires for each Newton iteration the evaluation of the residual *and* Jacobian as well as the solution of a non-symmetric and perhaps indefinite linear system; furthermore, a substantial number of Newton iterations may be required for highly nonlinear systems. Hence, the RB-EQP greedy algorithm, while not as efficient as the (original) weak greedy algorithm, achieves significant computational savings relative to the strong greedy algorithm.

The second task, the update of the EQP quadrature rule (line 16), is based on the LP EQP described in Section 3.4. We recall that the LP procedure requires as inputs a parameter training set $\Xi_J^{\rm train}$ and a state training set $U_J^{\rm train} \equiv \{u_N^{\rm train}(\mu)\}_{\mu\in\Xi_J^{\rm train}}$. It is clear from Proposition 3.2 that the best choice for error control is the "truth"-quadrature RB solution, which in fact corresponds to the following modifications to Algorithm 1: in line 12 set $N_{\rm eqp,smooth} = 1$; in line 14 set $u_N^{\rm train}(\mu) = u_N(\mu)$. But, as previously noted, this choice is rather expensive. As such, in the greedy algorithm, we set $u_N^{\rm train}(\mu) = u_N^\nu(\mu)$ — the current RB-EQP approximation based on the empirical quadrature rule. The state training set and the empirical quadrature rule itself are then simultaneously updated using $N_{\rm eqp,smooth}$ smoothing iterations (lines 12–17); we have found $N_{\rm eqp,smooth} \in \{2,3\}$ is sufficient.

Algorithm 1: Greedy algorithm for simultaneous RB and EQP training.

inputs : Parameter training set: $\Xi_J^{\text{train}} \subset \mathcal{D}$
EQP tolerance: $\delta \in \mathbb{R}_{+0}$
RB residual tolerance: $\epsilon_{\text{RB}} \in \mathbb{R}_+$
outputs: Reduced basis: $\{\phi_i\}_{i=1}^{N_{\max}}$
EQP rule: $\{x_k^\nu, \rho_k^\nu\}_{k=1}^{K^\nu}$

Initialize the EQP state training set: set $U_J^{\text{train}} \equiv \{u_N^{\text{train}}(\mu) = 0\}_{\mu \in \Xi_J^{\text{train}}}$ and $\Xi^{\text{train,exact}} = \emptyset$.
for $N = 1, \dots, N_{\max}$ **do**
    if $N = 1$ **then**
        Set $\mu^{(N)} = \arg\inf_{\mu \in \Xi_J^{\text{train}}} \|\bar{\mu} - \mu\|$ for $\bar{\mu} \equiv \frac{1}{N} \sum_{\mu \in \Xi_J^{\text{train}}} \mu$.
    else
        Find the parameter $\mu^{(N)}$ that maximizes the "truth" dual-norm of the residual:

$$\mu^{(N)} = \underset{\mu \in \Xi_J^{\text{train}}}{\arg\sup} \|r^h(u_{N-1}^\nu(\mu), \cdot; \mu)\|_{(\mathcal{V}^h)'}.$$

    end
    If $\|r^h(u_{N-1}(\mu^{(N)}); \mu)\|_{(\mathcal{V}^h)'} < \epsilon_{\text{RB}}$, terminate.
    Find the "truth" FE solution $u^h(\mu^{(N)}) \in \mathcal{V}^h$.
    Update reduced basis

$$\{\phi_i\}_{i=1}^N = \text{Gram-Schmidt}_{\mathcal{V}}\{\phi_1, \dots, \phi_{N-1}, u^h(\mu^{(N)})\}.$$

    Update EQP state training set: $u_N^{\text{train}}(\mu^{(N)}) = u^h(\mu^{(N)})$, $\Xi^{\text{train,exact}} = \Xi^{\text{train,exact}} \cup \mu^{(N)}$.
    for $i = 1, \dots, N_{\text{eqp,smooth}}$ **do**
        if $i \neq 1$ **then**
            For $\mu \in \Xi_J^{\text{train}}$ but $\mu \notin \Xi^{\text{train,exact}}$, solve for $u_N^\nu(\mu)$ and set $u_N^{\text{train}}(\mu) = u_N^\nu(\mu)$.
        end
        Solve $\texttt{LP}_{\text{EQP}}^{\nu,N}$ for $\nu \equiv \{h, \delta, \Xi_J^{\text{train}}, U_J^{\text{train}} \equiv \{u_N^{\text{train}}(\mu)\}_{\mu \in \Xi_J^{\text{train}}}\}$ to update $\{x_k^\nu, \rho_k^\nu\}_{k=1}^{K^\nu}$.
    end
end

We note that a similar simultaneous update strategy has been introduced in the context of RB-EIM [8]. We report in the next section empirical results which demonstrate that (38) and (39) hold even for $\epsilon^{\text{train}} = 0$; however, we have at present no theoretical control over the error induced by the "bootstrap" training set.

We also note a subtle, but important, feature of the greedy algorithm (implemented in line 11); the EQP state training set associated with the RB snapshot parameters are the "truth" solution. This construction guarantees that for $\{\mu^{(n)}\}_{n=1}^{N}$, $\|u_N(\mu) - u_N^\nu(\mu)\|_{\mathcal{V}} = \|u^h(\mu) - u_N^\nu(\mu)\|_{\mathcal{V}}$ is $\mathcal{O}(\delta)$ and hence $\|r^h(u_N^\nu(\mu), \cdot; \mu)\|_{(\mathcal{V}^h)'}$ is $\mathcal{O}(\delta)$. This feature is important for the stability of the greedy procedure; without the feature, the dual-norm of the RB-EQP solution for a snapshot parameter may be large, and the algorithm may attempt to add the associated snapshot that already belongs to the RB.

We finally comment on the computational cost for a single step of the greedy algorithm. We decompose the algorithm into three parts: "truth" residual sampling; "truth" solution; and EQP update. The "truth" residual sampling requires J RB-EQP solves in $\mathcal{O}(JN^\cdot)$ operations and J "truth" residual dual-norm evaluations in $\mathcal{O}(J\mathcal{N}^\cdot)$ operations. The "truth" solution requires $\mathcal{O}(\mathcal{N}^\cdot)$ operations. We again note that although the computation of the "truth" residual dual norm and the computation of the "truth" solution each requires $\mathcal{O}(\mathcal{N}^\cdot)$ operations, the latter is much more expensive than the former; for this reason we avoid all reference to the "truth" solution except for the (relatively few) snapshots which form the RB space. Finally the EQP update requires $JN_{\text{eqp,smooth}}$ RB-EQP solves in $\mathcal{O}(JN_{\text{eqp,smooth}}N^\cdot)$ operations and $N_{\text{eqp,smooth}}$ solves of $\mathtt{LP}_{\text{EQP}}^{\nu,N}$. The $\mathtt{LP}_{\text{EQP}}^{\nu,N}$ has K^h unknowns, K^h non-negativity constraints, and $JN+1$ absolute-value bound constraints; the preparation of the absolute-value bounds in particular requires $\mathcal{O}(JN^2K^h)$ operations as shown in Appendix A. The storage requirement is dominated by the LP constraint matrix and is $\mathcal{O}(K^h(JN+1))$.

3.7. Functional output evaluation

In many engineering scenarios, our interest is certain (integral) quantities. As before, to fix notation and to provide a concrete description, we introduce a model problem: the evaluation of local strain energy associated with a neo-Hookean solid,

$$s(\mu) \equiv \psi(u(\mu); \mu) \in \mathbb{R}, \tag{40}$$

where

$$\psi(w; \mu) \equiv \int_{\Omega^o} \Psi(F(w(x)); \mu) dx \quad \forall w \in \mathcal{V} \tag{41}$$

for $\Psi : \mathbb{R}^N \times \mathcal{D} \to \mathbb{R}$ the strain energy density function (10) and $\Omega^o \subset \Omega$ a region of interest.

We now introduce an associated "truth" FE approximation of the output form (41). As before, we also define a "truth" quadrature rule $\{x_k^{o,h} \in \Omega^o\}$ and the associated weights $\{\rho_k^{o,h} \in \mathbb{R}_+\}$. We then define a "truth"-quadrature output form,

$$\psi^h(w; \mu) \equiv \sum_{k=1}^{K^{o,h}} \rho_k^{o,h} \Psi(F(w(x^{o,h}); \mu), \quad \forall w \in \mathcal{V}, \ \forall \mu \in \mathcal{D}. \tag{42}$$

Our "truth" FE approximation to (40) is as follows: given $\mu \in \mathcal{D}$, find $s^h(\mu) \equiv \psi^h(u^h(\mu); \mu) \in \mathbb{R}$. Similarly, our "truth"-quadrature RB approximation to (40) is as follows: given $\mu \in \mathcal{D}$, find

$s_N(\mu) \equiv \psi^h(u_N(\mu); \mu) \in \mathbb{R}$. Given the solution (field), the computational cost for the evaluation of the output using the "truth" quadrature is $\mathcal{O}(K^h)$; the approach hence does not meet the online complexity goal of model reduction set forth in the Introduction.

We thus introduce an EQP approximation for the output form (41). As before, we introduce an empirical quadrature rule $\{x_k^{o,\nu} \in \Omega^o\} \subset \{x_k^{o,h}\}$ and the associated weights $\{\rho_k^{o,\nu} \in \mathbb{R}_+\}$; we recall that an empirical rule is (hyper)parametrized by $\nu \equiv \{h, \delta, \Xi_J\}$ for h the "truth"-quadrature parameter, δ the threshold tolerance, and Ξ_J the parameter training set. We then define an EQP-quadrature output form,

$$\psi^\nu(w; \mu) \equiv \sum_{k=1}^{K^{o,\nu}} \rho_k^{o,\nu} \Psi(F(w(x_k^{o,\nu})); \mu), \quad \forall w \in \mathcal{V},\ \forall \mu \in \mathcal{D}.$$

Our RB-EQP approximation to (40) is as follows: given $\mu \in \mathcal{D}$, find $s^\nu(\mu) \equiv \psi^\nu(u_N^\nu(\mu); \mu) \in \mathbb{R}$. We invoke the LP EQP described in Section 2.1 on the integrand $\Psi(F(u_N^\nu(\cdot; \mu)); \mu)$, $\mu \in \mathcal{D}$, to identify the quadrature rule; the linear program has $K^{o,h}$ unknowns and $J + 1$ constraints. Given the solution (field) evaluated at the EQP quadrature points, the computational cost for the evaluation of the output using the EQP quadrature is $\mathcal{O}(K^\nu) \ll \mathcal{O}(K^h)$; we hence meet the online complexity goal of model reduction.

4. Numerical example: neo-Hookean beam

4.1. Problem description

We consider a neo-Hookean beam subject to self-weight load. The geometry of the beam is shown in Figure 1; the beam is characterized by a height $\tilde{h}$, length $\tilde{L}$, hole radius $\tilde{r}_{\text{hole}}$, and hole position $(\tilde{x}_{\text{hole}}, \tilde{y}_{\text{hole}})$ with respect to the origin fixed at the bottom left corner. The beam is clamped (homogeneous Dirichlet) at the left end, free (homogeneous Neumann) on all other surfaces, and subject to gravity of magnitude $\tilde{g}$. The beam is modeled as a (plain-strain) neo-Hookean solid with a density $\tilde{\rho}$, Young's module $\tilde{E}$, and Poisson ratio ν. The two parameters of the problem are the gravity angle $\theta_g \in [-\pi/2, \pi/2]$ and the Poisson ratio $\nu \in [0.35, 0.45]$; the former should be interpreted to result from a variation in the orientation of the beam with respect to the vertical. We hence set $\mu \equiv (\theta_g, \nu)$ and $\mathcal{D} \equiv [-\pi/2, \pi/2] \times [0.35, 0.45]$. The output of interest is the strain energy density integrated over the annular region $\tilde{\Omega}_{\text{annulus}}$ of radius $\tilde{r}_{\text{annulus}}$; the output serves as an indicator of stress concentration in the vicinity of the hole.

We introduce characteristic length and pressure scales of $\tilde{h}$ and $\tilde{E}$, respectively. With this normalization, the geometry of the problem is described in terms of $h \equiv \tilde{h}/\tilde{h} = 1$, $L \equiv \tilde{L}/\tilde{h} \equiv 4$, $x_{\text{hole}} = \tilde{x}_{\text{hole}}/\tilde{h} \equiv 1$, $y_{\text{hole}} = \tilde{y}_{\text{hole}}/\tilde{h} \equiv 1/2$, $r_{\text{hole}} = \tilde{r}_{\text{hole}}/\tilde{h} \equiv 0.15$, $r_{\text{annulus}} = \tilde{r}_{\text{annulus}}/\tilde{h} \equiv 0.3$, and body-force density $\rho g = (\tilde{\rho}\tilde{g})/(\tilde{E}/\tilde{h}) \equiv 0.005$. The particular value of body-force density results from typical values of the density and Young's modulus for a soft-rubber-like material. (We however consider the range of Poisson's ratio away from the incompressible limit $\nu = 1/2$; treatment of near-incompressibility introduces complications not specifically related to our LP EQP RB approach.)

Figure 2(a) shows the deformed beam for $(\theta_g, \nu) = (0, 0.4)$ and the associated strain energy density distribution over Ω_{annulus}. Figure 2(b) shows the variation in the output as a function of θ_g for a fixed $\nu \equiv 0.4$; the nonlinear effect of the neo-Hookean model is clearly evident in the response.

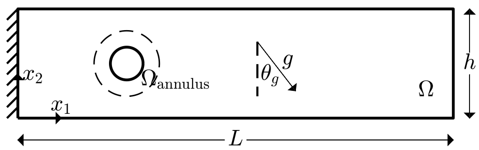
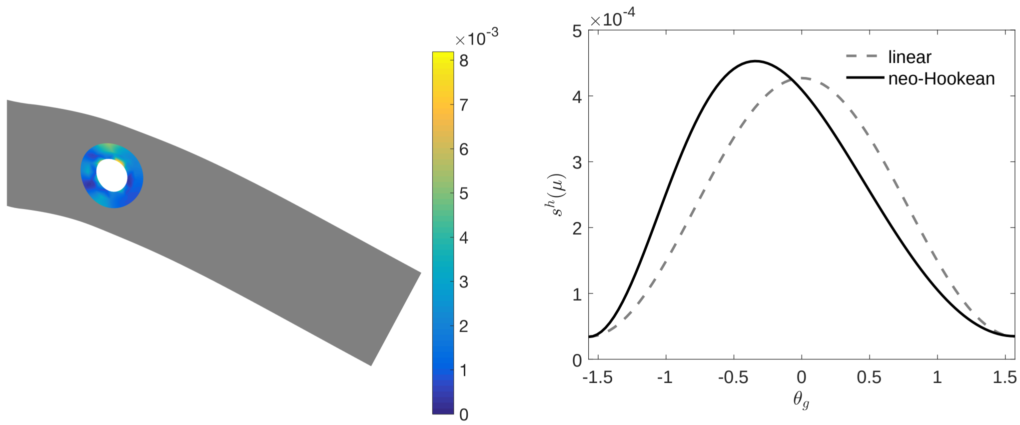


Figure 1: Geometry of the neo-Hookean beam problem.



(a) deformed beam ($\theta_g = 0$, $\nu = 0.4$)

(b) output variation ($\nu = 0.4$)

Figure 2: Response of the neo-Hookean beam problem.

4.2. Results

We apply the (relative-error version of the) greedy algorithm, Algorithm 1, to the neo-Hookean beam problem. The "truth" space of $\mathbb{P}^2$ triangular finite elements has 1568 degrees of freedom; the "truth" quadrature rule contains $K^h = 6878$ points over Ω, of which $K^{o,h} = 1216$ are in Ω_{annulus}. (The p-convergence study for 3×3 parameter points uniformly distributed over $\mathcal{D}$ confirms that the maximum relative error in the $\mathbb{P}^2$ "truth" output is $\approx 1.9 \times 10^{-3}$ with respect to the $\mathbb{P}^4$ solutions with 6032 degrees of freedom; the convergence study also suggests that the continuous (infinite-dimensional) problem is indeed well posed and we remain on a single solution branch.) The inputs to the greedy algorithm are as follows: the parameter training set $\Xi_J^{\text{train}} \subset \mathcal{D}$ which consists of $|\Xi_J^{\text{train}}| = 31 \times 5$ points uniformly distributed over $\mathcal{D}$; the EQP tolerance of $\delta = 10^{-3}$ for both the residual and output; and RB dual-norm tolerance of $\epsilon_{\text{RB}} = 10^{-2}$.

We first test the RB-EQP model for all parameter values in Ξ_J^{train}; this test does *not* account for the "generalization error" associated with the fact that Ξ_J^{train} is a *finite subset* of $\mathcal{D}$. Table 1 summarizes the behavior of the RB-EQP model. We first observe that the dual-norm of the "truth" residual over the parameter training set Ξ_J^{train} decreases rapidly with N, with the algorithm terminating for $N_{\max} = 7$. We next assess the quality of the RB spaces: the solutions (field) and outputs associated with the "truth"-quadrature RB approximation converge rapidly with N.

We now assess the quality of the RB-EQP solutions (field) and make the following observations. The EQP quadrature rules are sparse: $K^\nu \leq 32 \ll 6878 \equiv K^h$ for all values of N. The relative error between the RB and RB-EQP solutions, $\sup_{\mu \in \Xi_J^{\text{train}}} \frac{\|u_N(\mu) - u_N^\nu(\mu)\|_\mathcal{V}}{\|u_N(\mu)\|_\mathcal{V}}$ is bounded by $\approx 2N^{1/2}\delta$ for $N > 2$ as predicted by Proposition 3.2; this error control provided by the RB-EQP procedure is more clearly depicted in Figure 3. We emphasize that the error here is measured relative to

N	residual $\frac{\|r^h(u_N^\nu,\cdot)\|_{(\mathcal{V}_h)'}}{\|u_N^\nu\|_{\mathcal{V}}}$	RB assessment $\frac{\|u_h - u_N\|_{\mathcal{V}}}{\|u_h\|_{\mathcal{V}}}$	$\frac{\|s(u_h)-s(u_N)\|}{\|s(u_h)\|}$	RB-EQP assessment K^ν	$\frac{\|u_N - u_N^\nu\|_{\mathcal{V}}}{\|u_N\|_{\mathcal{V}}}$	$\frac{\|u_h - u_N^\nu\|_{\mathcal{V}}}{\|u_h\|_{\mathcal{V}}}$	$K^{o,\nu}$	$\frac{\|s(u_N^\nu)-s^\nu(u_N^\nu)\|}{\|s(u_N^\nu)\|}$	$\frac{\|s(u_h)-s^\nu(u_N^\nu)\|}{\|s(u_h)\|}$
1	1.50×10^{2}	1.00×10^{0}	10.00×10^{-1}	5	1.35×10^{-1}	1.00×10^{0}	4	1.00×10^{-3}	10.00×10^{-1}
2	1.89×10^{-1}	7.87×10^{-1}	9.26×10^{-1}	14	4.47×10^{-3}	7.87×10^{-1}	11	1.00×10^{-3}	9.26×10^{-1}
3	1.77×10^{-1}	7.69×10^{-1}	9.21×10^{-1}	18	1.67×10^{-3}	7.69×10^{-1}	11	1.00×10^{-3}	9.20×10^{-1}
4	2.91×10^{-2}	1.65×10^{-1}	1.42×10^{-1}	25	1.78×10^{-3}	1.65×10^{-1}	9	1.00×10^{-3}	1.38×10^{-1}
5	2.61×10^{-2}	1.24×10^{-1}	1.60×10^{-1}	29	1.98×10^{-3}	1.24×10^{-1}	13	1.00×10^{-3}	1.66×10^{-1}
6	1.51×10^{-2}	1.23×10^{-1}	1.58×10^{-1}	29	2.43×10^{-3}	1.24×10^{-1}	13	1.00×10^{-3}	1.68×10^{-1}
7	8.87×10^{-3}	1.65×10^{-2}	4.09×10^{-2}	32	2.29×10^{-3}	1.53×10^{-2}	14	1.00×10^{-3}	4.50×10^{-2}

Table 1: The convergence of RB-EQP approximations over all parameter values in Ξ_J^{train}. All entries are supremum over the parameter training set Ξ_J^{train}; e.g., the first column is $\sup_{\mu \in \Xi_J^{\text{train}}} \frac{\|r^h(u_N^\nu(\mu),\cdot;\mu)\|_{(\mathcal{V}_h)'}}{\|u_N^\nu(\mu)\|_{\mathcal{V}}}$.

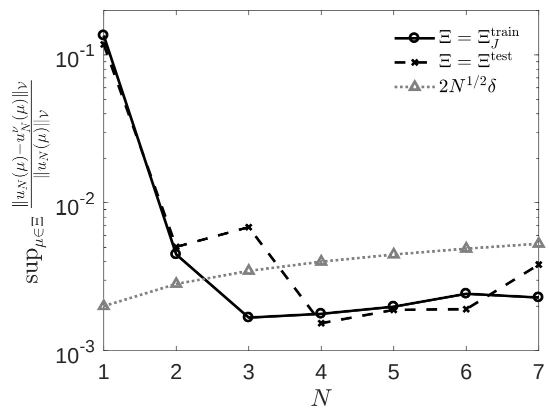
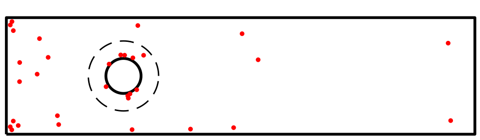


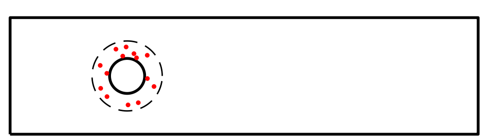
Figure 3: The error in the RB-EQP approximation with respect to the "truth"-quadrature RB solution.

$u_N(\mu)$ and hence implicitly assumes $\epsilon^{\text{train}} = 0$ in (39); this in turn justifies (empirically) the bootstrap approach for our particular problem. For $N = 1$ and 2, it is likely that the conditions that are not explicitly enforced in $\texttt{LP}^{\nu,N}_{\text{EQP}}$, (34) and (37), are violated and hence the bound does not hold. The error in the RB-EQP solution relative to the "truth" is comparable to the error in the "truth"-quadrature RB solution relative to the "truth"; the use of the EQP quadrature, which has over two orders of magnitude fewer quadrature points, does not compromise the accuracy of the approximation. Figure 4(a) shows the EQP quadrature points associated with the $N = 7$ RB-EQP residual form; the EQP points are distributed in a nontrivial manner and are clustered in the regions of high stress concentration at the root and around the hole.

We next assess the quality of the RB-EQP output approximations and make the following



(a) residual quadrature points $\{x_k^\nu\}_{k=1}^{K^\nu=32}$



(b) output quadrature points $\{x_k^{o,\nu}\}_{k=1}^{K^{o,\nu}=14}$

Figure 4: Residual and output quadrature points for $N = 7$.

N	residual	RB assessment		RB-EQP assessment					
	$\frac{\|r^h(u_N^\nu,\cdot)\|_{(\mathcal{V}_h)'}}{\|u_N^\nu\|_\mathcal{V}}$	$\frac{\|u_h-u_N\|_\mathcal{V}}{\|u_h\|_\mathcal{V}}$	$\frac{\|s(u_h)-s(u_N)\|}{\|s(u_h)\|}$	K^ν	$\frac{\|u_N-u_N^\nu\|_\mathcal{V}}{\|u_N\|_\mathcal{V}}$	$\frac{\|u_h-u_N^\nu\|_\mathcal{V}}{\|u_h\|_\mathcal{V}}$	$K^{o,\nu}$	$\frac{\|s(u_N^\nu)-s^\nu(u_N^\nu)\|}{\|s(u_N^\nu)\|}$	$\frac{\|s(u_h)-s^\nu(u_N^\nu)\|}{\|s(u_h)\|}$
1	1.46×10^{-1}	7.80×10^{-1}	9.84×10^{-1}	5	1.18×10^{-1}	7.84×10^{-1}	4	9.59×10^{-4}	9.88×10^{-1}
2	1.51×10^{-1}	7.74×10^{-1}	9.19×10^{-1}	14	5.04×10^{-3}	7.74×10^{-1}	11	1.00×10^{-3}	9.18×10^{-1}
3	1.59×10^{-1}	7.54×10^{-1}	9.14×10^{-1}	18	6.85×10^{-3}	7.54×10^{-1}	11	9.96×10^{-4}	9.13×10^{-1}
4	2.11×10^{-2}	1.58×10^{-1}	1.41×10^{-1}	25	1.53×10^{-3}	1.57×10^{-1}	9	9.83×10^{-4}	1.37×10^{-1}
5	1.40×10^{-2}	1.12×10^{-1}	1.47×10^{-1}	29	1.89×10^{-3}	1.12×10^{-1}	13	9.37×10^{-4}	1.48×10^{-1}
6	1.37×10^{-2}	1.11×10^{-1}	1.46×10^{-1}	29	1.91×10^{-3}	1.11×10^{-1}	13	1.02×10^{-3}	1.51×10^{-1}
7	8.35×10^{-3}	1.47×10^{-2}	4.06×10^{-2}	32	3.82×10^{-3}	1.37×10^{-2}	14	9.36×10^{-4}	4.41×10^{-2}

Table 2: The convergence of RB-EQP approximation for $\Xi^{\text{test}} \neq \Xi_J^{\text{train}}$ comprises 50 random uniformly distributed points over $\mathcal{D}$. All entries are supremum over the parameter training set Ξ^{test}; e.g., the first column is $\sup_{\mu\in\Xi^{\text{test}}} \frac{\|r^h(u_N^\nu(\mu),\cdot;\mu)\|_{(\mathcal{V}_h)'}}{\|u_N^\nu(\mu)\|_\mathcal{V}}$.

observations. As before, $K^{o,\nu} \leq 14 \ll 1216 \equiv K^{o,h}$ for all values of N and hence EQP again achieves sparsity. The integration error due to the use of the EQP quadrature instead of the "truth" quadrature is exactly controlled for this test case (without generalization error, i.e., the supremum is over Ξ_J^{train}). The error in the RB-EQP output relative to the "truth" is comparable to the error in the "truth"-quadrature RB output relative to the "truth" despite the significant sparsity of the RB-EQP quadrature rule. Figure 4(b) shows the EQP quadrature points associated with the output evaluation of the $N = 7$ approximation.

We now test the RB-EQP model for $\Xi^{\text{test}} \subset \mathcal{D}$ which consists of $|\Xi^{\text{test}}| = 50$ random points uniformly sampled over $\mathcal{D}$; this test, unlike the previous test, accounts for the "generalization error." Table 2 shows that the results over Ξ^{test} are largely unchanged from the results over Ξ_J^{train} shown in Table 1. The result is consistent with Proposition 2.2, which predicts a small generalization error for a sufficiently dense parameter training set $\Xi_J^{\text{train}} \subset \mathcal{D}$.

We finally report the online computational savings. Both the "truth" FE solver and the RB-EQP solver are implemented in Matlab and solve the nonlinear problem using the damped Newton strategy described in Section 3.2. All computations are performed on a commodity laptop. Over the 50 test cases defined by Ξ^{test}, the RB-EQP solver ($N = 7$) on average reduces wall-clock time by ≈ 60.

4.3. A perspective on generality

We have observed that the RB-EQP formulation introduced in this work is rather "blackbox" in the sense that i) it requires minimal hyperparameter specification and ii) it is quite robust with respect to applications. As regards i), the greedy algorithm discussed in Section 3.6 requires as input only the parameter training set and EQP and RB tolerances; the algorithm then yields an online-efficient reduced model for the parametrized nonlinear PDE. As regards ii), in addition to the neo-Hookean hyperelasticity problem reported here, we have also successfully applied the formulation to a nonlinear reaction-diffusion equation with a cubic reaction term, nonlinear elasticity based on the Saint Venant-Kirchhoff model, and the compressible Navier-Stokes equations.

Appendix A. Explicit representation of RB-EQP accuracy constraints

We provide an explicit presentation of the RB-EQP accuracy constraints (22). We first reproduce the expression:

$$\|J_N^h(\mathbf{u}_N^{\text{train}}(\mu);\mu)^{-1}\hat{R}_N^h(\rho;\mathbf{u}_N^{\text{train}}(\mu);\mu)\|_\infty \le \delta, \quad \forall \mu \in \Xi_J^{\text{train}}. \tag{A.1}$$

The i-th entry of the empirical quadrature operator is given by (20),

$$(\hat{R}_N^h(\rho;\mathbf{u}_N^{\text{train}}(\mu);\mu))_i = \sum_{k=1}^{K^h} \rho_k \eta(Z_N\mathbf{u}_N^{\text{train}}(\mu),\phi_i;\mu;x_k^h), \quad i=1,\dots,N,$$

where $\eta(Z_N\mathbf{u}_N^{\text{train}}(\mu),\phi_i;\mu;x_k^h)$ is the residual integrand (12) evaluated for the trial function $Z_N\mathbf{u}_N^{\text{train}}(\mu)$, the test function ϕ_i, parameter μ, and point x_k^h. The (i,j)-entry of the "truth"-quadrature Jacobian matrix is given by

$$(J_N^h(\mathbf{u}_N^{\text{train}}(\mu);\mu))_{i,j} = (r^h)'(Z_N\mathbf{u}_N^{\text{train}}(\mu),\phi_j,\phi_i;\mu) = \sum_{k=1}^{K^h} \rho_k^h \eta'(Z_N\mathbf{u}_N^{\text{train}}(\mu),\phi_j,\phi_i;\mu;x_k^h), \quad i,j=1,\dots,N,$$

where $\eta'(Z_N\mathbf{u}_N^{\text{train}}(\mu),\phi_j,\phi_i;\mu;x_k^h)$ is the Gâteaux derivative of $\eta(\cdot,\phi_i;\mu;x_k^h):\mathcal{V}\to\mathbb{R}$ at $Z_N\mathbf{u}_N^{\text{train}}(\mu)\in\mathcal{V}_N$ in the direction $\phi_j\in\mathcal{V}_N$. We denote the (i,j)-entry of the inverse of $J_N^h(\mathbf{u}_N^{\text{train}}(\mu);\mu)$ by $(J_N^h(\mathbf{u}_N^{\text{train}}(\mu);\mu)^{-1})_{i,j}$. Then, the constraints (A.1) (or equivalently (22)) can be explicitly written as

$$\begin{aligned}&\left|\sum_{j=1}^{N}(J_N^h(\mathbf{u}_N^{\text{train}}(\mu);\mu)^{-1})_{i,j}(\hat{R}_N^h(\rho;\mathbf{u}_N^{\text{train}}(\mu);\mu))_j\right|\\ &= \left|\sum_{k=1}^{K^h}\rho_k\left[\sum_{j=1}^{N}(J_N^h(\mathbf{u}_N^{\text{train}}(\mu);\mu)^{-1})_{i,j}\eta(Z_N\mathbf{u}_N^{\text{train}}(\mu),\phi_j;\mu;x_k^h)\right]\right| \le \delta, \quad \forall i=1,\dots,N,\ \forall\mu\in\Xi_J^{\text{train}}.\end{aligned}$$

In practice, we compute the quantity in the bracket for each $\mu\in\Xi_J^{\text{train}}$ as follows: we evaluate the residual kernel $\eta(Z_N\mathbf{u}_N^{\text{train}}(\mu),\phi_j;\mu;x_k^h)$ for $j=1,\dots,N$ and $k=1,\dots,K^h$ and store it as a $N\times K^h$ matrix; we then apply $(J_N^h(\mathbf{u}_N^{\text{train}}(\mu);\mu))^{-1}$ (from left) to compute $N\times K^h$ entries of LP accuracy constraints. The evaluation of the residual kernel requires $\mathcal{O}(NK^h)$ operations, whereas the application of the inverse Jacobian requires $\mathcal{O}(N^2K^h)$ operations. (In practice, the constant associated with the former is typically much greater than the constant associated with the latter.) We repeat the procedure for all $\mu\in\Xi_J^{\text{train}}$ to populate the $(NJ)\times K^h$ entries of the LP accuracy constraints.

Acknowledgments

This work was supported by NSERC Discovery Grant (M Yano) and ONR Grant N00014-17-1-2077 (AT Patera).

References

[1] 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.

[2] P. Astrid, S. Weiland, K. Willcox, and T. Backx. Missing point estimation in models described by proper orthogonal decomposition. *IEEE Trans Autom Control*, 53(10):2237–2251, 2008.

[3] 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.

[4] F. Brezzi, J. Rappaz, and P. A. Raviart. Finite dimensional approximation of nonlinear problems. part I: Branches of nonsingular solutions. *Numerische Mathematik*, 36:1–25, 1980.

[5] T. Bui-Thanh, D. Murali, and K. Willcox. Proper orthogonal decomposition extensions for parametric applications in compressible aerodynamics. AIAA 2003-4213, AIAA, 2003.

[6] G. Caloz and J. Rappaz. Numerical analysis for nonlinear and bifurcation problems. In C. PG and L. JL, editors, *Handbook of Numerical Analysis, Vol. V*, pages 487–637. Elsevier Science B.V., Amsterdam, 1997.

[7] K. Carlberg, C. Bou-Mosleh, and C. Farhat. Efficient non-linear model reduction via a least-squares petrovgalerkin projection and compressive tensor approximations. *Int. J. Numer. Methods Eng.*, 86(2):155–181, 2011.

[8] 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.

[9] R. DeVore, S. Fourcart, G. Petrova, and P. Wojtaszczyk. Computing a quantity of interest from observational data. *preprint*, 2017.

[10] J. L. Eftang, M. A. Grepl, and A. T. Patera. A posteriori error bounds for the empirical interpolation method. *C. R. Acad. Sci. Paris, Ser. I*, 348(9–10):575–579, 2010.

[11] R. Everson and L. Sirovich. Karhunen-Loève procedure for gappy data. *J. Opt. Soc. Am. A*, 12(8):1657–1664, 1995.

[12] 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. *Int. J. Numer. Methods Eng.*, 102(5):1077–1110, 2015. nme.4820.

[13] 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.

[14] N. C. Nguyen, A. T. Patera, and J. Peraire. A best points interpolation method for efficient approximation of parametrized functions. *Int. J. Numer. Methods Eng.*, 73(4):521–543, 2008.

[15] N. C. Nguyen and J. Peraire. An efficient reduced-order modeling approach for non-linear parametrized partial differential equations. *Int. J. Numer. Methods Eng.*, 76(1):27–55, 2008.

[16] M. Ohlberger and S. Rave. Reduced basis methods: Success, limitations and future challenges. In *Proceedings of the Conference Algoritmy*, pages 1–12, 2016.

[17] A. T. Patera and M. Yano. An LP empirical quadrature procedure for parametrized functions. *C. R. Acad. Sci. Paris, Ser. I*, 2017.

[18] 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. *Arch. Comput. Methods Eng.*, 15(3):229–275, 2008.

[19] E. K. Ryu and S. P. Boyd. Extensions of gauss quadrature via linear programming. *Found Comput Math*, 15(4):953–971, 2015.

[20] K. Veroy and A. T. Patera. Certified real-time solution of the parametrized steady incompressible Navier-Stokes equations: rigorous reduced-basis *a posteriori* error bounds. *Int. J. Numer. Methods Fluids*, 47:773–788, 2005.