4.15

CERTIFIED REAL-TIME SOLUTION OF PARAMETRIZED PARTIAL DIFFERENTIAL EQUATIONS

Nguyen Ngoc Cuong, Karen Veroy, and Anthony T. Patera
Massachusetts Institute of Technology, Cambridge, MA, USA

1. Introduction

Engineering analysis requires the prediction of (say, a single) selected "output" s^{e} relevant to ultimate component and system performance:* typical outputs include energies and forces, critical stresses or strains, flowrates or pressure drops, and various local and global measures of concentration, temperature, and flux. These outputs are functions of system parameters, or "inputs", μ, that serve to identify a particular realization or configuration of the component or system: these inputs typically reflect geometry, properties, and boundary conditions and loads; we shall assume that μ is a P-vector (or P-tuple) of parameters in a prescribed closed input domain $\mathcal{D} \subset \mathbb{R}^P$. The input–output relationship $s^{\mathrm{e}}(\mu)\colon\ \mathcal{D} \to \mathbb{R}$ thus encapsulates the behavior relevant to the desired engineering context.

In many important cases, the input–output function $s^{\mathrm{e}}(\mu)$ is best articulated as a (say) linear functional ℓ of a field variable $u^{\mathrm{e}}(\mu)$. The field variable, in turn, satisfies a μ-parametrized partial differential equation (PDE) that describes the underlying physics: for given $\mu \in \mathcal{D}$, $u^{\mathrm{e}}(\mu) \in X^{\mathrm{e}}$ is the solution of

$$g(u^{\mathrm{e}}(\mu), v; \mu) = 0, \quad \forall\, v \in X^{\mathrm{e}}, \tag{1}$$

where g is the weak form of the relevant partial differential equation† and X^{e} is an appropriate Hilbert space defined over the physical domain $\Omega \subset \mathbb{R}^d$. Note

*Here superscript "e" shall refer to "exact." We shall later introduce a "truth approximation" which will bear no superscript.

†We shall restrict our attention in this paper to *second-order elliptic* partial differential equations; see Outlook for a brief discussion of parabolic problems.

S. Yip (ed.),
Handbook of Materials Modeling, 1529–1564.

in the *linear case*, $g(w, v; \mu) \equiv a(w, v; \mu) - f(v)$, where $a(\cdot, \cdot; \mu)$ and f are continuous bilinear and linear forms, respectively; for any given $\mu \in \mathcal{D}$, $u^{\mathrm{e}}(\mu) \in X^{\mathrm{e}}$ now satisfies

$$a(u^{\mathrm{e}}(\mu), v; \mu) = f(v), \quad \forall\, v \in X^{\mathrm{e}} \ \text{(linear)}. \tag{2}$$

Relevant system behavior is thus described by an implicit "input–output" relationship

$$s^{\mathrm{e}}(\mu) = \ell(u^{\mathrm{e}}(\mu)), \tag{3}$$

evaluation of which necessitates solution of the partial differential equation (1) or (2).

Many problems in materials and materials processing can be formulated as particular instantiations of the abstraction (1) and (3) or perhaps (2) and (3). Typical field variables and associated second-order elliptic partial differential equations include temperature and steady conduction–Poisson; displacement and equilibrium or Helmholtz elasticity; {velocity, temperature} and steady Boussinesq incompressible Navier–Stokes; wavefunction and stationary Schrödinger via (say) Hartree–Fock approximation. The latter two equations are nonlinear, while the former two equations are linear; in subsequent sections we shall provide detailed examples of both nonlinear and linear problems.

Our particular interest – or certainly the best way to motivate our approach – is "deployed" systems: components or processes that are in service, in operation, or in the field. For example, in the materials and materials processing context, we may be interested in assessment, evolution, and accommodation of a crack in a critical component of an in-service jet engine; in real-time characterization and optimization of the heat treatment protocol for a turbine disk; or in online thermal "control" of Bridgman semiconductor crystal growth. Typical computational tasks include robust parameter estimation (inverse problems) and adaptive design (optimization problems): in the former – for example, assessment of current crack length or in-process heat transfer coefficient – we must deduce inputs μ representing system characteristics based on outputs $s^{\mathrm{e}}(\mu)$ reflecting measured observables; in the latter – for example, prescription of allowable load or best thermal environment – we must deduce inputs μ representing "control" variables based on outputs $s^{\mathrm{e}}(\mu)$ reflecting current process objectives. Both of these demanding activities must support an *action* in the presence of continually evolving environmental and mission parameters.

The computational requirements on the forward problem are thus formidable: the evaluation must be *real-time*, since the action must be *immediate*; and the evaluation must be *certified* – endowed with a rigorous error bound – since the action must be *safe* and *feasible*. For example, in our aerospace crack example, we must predict *in the field* – without recourse to a lengthy computational investigation – the load that the potentially damaged structure

can unambiguously *safely* carry. Similarly, in our materials processing examples, we must predict *in operation* – in response to deduced environmental variation – temperature boundary conditions that will preserve the desired material properties.

Classical approaches such as the finite element method cannot typically satisfy these requirements. In the finite element method, we first introduce a piecewise-polynomial "truth" approximation subspace X ($\subset X^{\mathrm{e}}$) of dimension $\mathcal{N}$. The "truth" finite element approximation is then found by (say) Galerkin projection: given $\mu \in \mathcal{D}$,

$$s(\mu) = \ell(u(\mu)), \tag{4}$$

where $u(\mu) \in X$ satisfies

$$g(u(\mu), v; \mu) = 0, \quad \forall\, v \in X, \tag{5}$$

or, in the linear case $g(w, v; \mu) \equiv a(w, v; \mu) - f(v)$,

$$a(u(\mu), v; \mu) = f(v), \quad \forall\, v \in X \ \text{(linear)}. \tag{6}$$

We assume that (5) and (6) are well-posed; we articulate the associated hypotheses more precisely in the context of *a posteriori* error estimation.

We shall assume – hence the appellation "truth" – that X is sufficiently rich that $u(\mu)$ (respectively, $s(\mu)$) is sufficiently close to $u^{\mathrm{e}}(\mu)$ (respectively, $s^{\mathrm{e}}(\mu)$) for all μ in the parameter domain $\mathcal{D}$. Unfortunately, for any reasonable error tolerance, the dimension $\mathcal{N}$ needed to satisfy this condition – even with the application of appropriate (parameter-dependent) adaptive mesh refinement strategies – is typically extremely large, and in particular much too large to provide real-time response in the deployed context. Deployed systems thus present no shortage of unique computational challenges; however, they also provide many unique computational opportunities – opportunities that must be exploited.

We first consider the "approximation opportunity." The critical observation is that, although the field variable $u^{\mathrm{e}}(\mu)$ generally belongs to the infinite-dimensional space X^{e} associated with the underlying partial differential equation, in fact $u^{\mathrm{e}}(\mu)$ resides on a very low-dimensional manifold $\mathcal{M}^{\mathrm{e}} \equiv \{u^{\mathrm{e}}(\mu) \mid \mu \in \mathcal{D}\}$ induced by the parametric dependence; for example, for a single parameter, $\mu \in \mathcal{D} \subset \mathbb{R}^{P=1}$, $u^{\mathrm{e}}(\mu)$ will describe a one-dimensional filament that winds through X^{e}. Furthermore, the field variable $u^{\mathrm{e}}(\mu)$ will typically be extremely regular in μ – the parametrically induced manifold $\mathcal{M}^{\mathrm{e}}$ is very smooth – even when the field variable enjoys only limited regularity with respect to the spatial coordinate $x \in \Omega$.* In the finite element method, the approximation space X is

*The smoothness in μ may be deduced from the equations for the sensitivity derivatives; the stability and continuity properties of the partial differential operator are crucial.

much too general – X can approximate many functions that do not reside on $\mathcal{M}^e$ – and hence much too expensive. This observation presents a clear opportunity: we can effect significant dimension reduction in state space if we restrict attention to $\mathcal{M}^e$; the field variable can then be adequately approximated by a space of dimension $N \ll \mathcal{N}$. However, since manipulation of even one "point" on $\mathcal{M}^e$ is expensive, we must identify further structure.

We thus next consider the "computational opportunities"; here there are two critical observations. The first observation derives from the mathematical formulation: very often, the parameter dependence of the partial differential equation can be expressed as the sum of Q products of (known, easily evaluated) parameter-dependent functions and parameter-independent continuous forms; we shall denote this structure as "affine" parameter dependence. In our linear case, (2), affine parameter dependence reduces to

$$a(w, v; \mu) = \sum_{q=1}^{Q} \Theta^q(\mu)\, a^q(w, v), \tag{7}$$

for $\Theta^q : \mathcal{D} \to \mathbb{R}$ and $a^q : X \times X \to \mathbb{R}$, $1 \le q \le Q$. The second observation derives from our context: rapid deployed response perforce places a predominant emphasis on very low *marginal* cost – we must minimize the additional effort associated with each new evaluation $\mu \to s(\mu)$ "in the field." These two observations present a clear opportunity: we can exploit the underlying affine parametric structure (7) to design effective *offline–online* computational procedures which willingly accept greatly increased initial preprocessing – offline, pre-deployed – expense in exchange for greatly reduced marginal – online, deployed – "in service" cost.*

The two essential components to our approach are (*i*) rapidly, uniformly (over $\mathcal{D}$) convergent reduced-basis (RB) approximations, and (*ii*) associated rigorous and sharp *a posteriori* error bounds. Both components exploit affine parametric structure and offline–online computational decompositions to provide extremely rapid deployed response – real-time prediction and associated error estimation. We next describe these essential ingredients.

2. Reduced-Basis Method

2.1. Approximation

The reduced-basis method was introduced in the late 1970s in the context of nonlinear structural analysis [1, 2] and subsequently abstracted, analyzed,

*Clearly, low marginal cost implies low asymptotic average cost; our methods are thus also relevant to (non real-time) many-query multi-optimization studies – and, in fact, to any situation characterized by extensive exploration of parameter space.

and extended to a much larger class of parametrized PDEs [3, 4] – including the incompressible Navier–Stokes equations [5–7] relevant to many materials processing applications. The RB method explicitly recognizes and exploits the dimension reduction afforded by the low-dimensional and smooth parametrically induced solution manifold. We note that the RB approximation is constructed not as an approximation to the exact solution, $u^{\mathrm{e}}(\mu)$, but rather as an approximation to the (finite element) truth approximation, $u(\mu)$. As already discussed, $\mathcal{N}$, the dimension of X, will be very large; our RB formulation and associated error estimation procedures must be *stable* and (online) *efficient* as $\mathcal{N} \to \infty$.

We shall consider in this section the linear case, $g(w, v; \mu) \equiv a(w, v; \mu) - f(v)$, in which $s(\mu)$ and $u(\mu)$ are given by (4) and (6), respectively; recall that a is bilinear and f, ℓ, are linear. We shall consider a "primal–dual" formulation particularly well-suited to good approximation and error characterization of the output; towards this end, we introduce a dual, or adjoint, problem: given $\mu \in \mathcal{D}$, $\psi(\mu) \in X$ satisfies

$$a(v, \psi(\mu); \mu) = -\ell(v), \quad \forall\, v \in X. \tag{8}$$

Note that if a is symmetric and $\ell = f$, which we shall denote "compliance," $\psi(\mu) = -u(\mu)$.

In the "Lagrangian" [4] RB approach, the field variable $u(\mu)$ is approximated by (typically) Galerkin projection onto a space spanned by solutions of the governing PDE at N selected points in parameter space. For the primal problem, (6), we introduce nested parameter samples $S_N \equiv \{\mu_1^{\mathrm{pr}} \in \mathcal{D}, \ldots, \mu_N^{\mathrm{pr}} \in \mathcal{D}\}$ and associated nested RB approximation subspaces $W_N \equiv \mathrm{span}\{\zeta_n \equiv u(\mu_n^{\mathrm{pr}}), 1 \le n \le N\}$ for $1 \le N \le N_{\max}$; similarly, for the dual problem (8), we define corresponding samples $S_{N^{\mathrm{du}}}^{\mathrm{du}} \equiv \{\mu_1^{\mathrm{du}} \in \mathcal{D}, \ldots, \mu_{N^{\mathrm{du}}}^{\mathrm{du}} \in \mathcal{D}\}$ and RB approximation spaces $W_{N^{\mathrm{du}}}^{\mathrm{du}} \equiv \mathrm{span}\{\zeta_n^{\mathrm{du}} \equiv \psi(\mu_n^{\mathrm{du}}), 1 \le n \le N^{\mathrm{du}}\}$ for $1 \le N^{\mathrm{du}} \le N_{\max}^{\mathrm{du}}$.* (Procedures for selection of good samples S_N, $S_{N^{\mathrm{du}}}^{\mathrm{du}}$ and hence spaces W_N, $W_{N^{\mathrm{du}}}^{\mathrm{du}}$ will be discussed in subsequent sections.) Our RB approximation is thus: given $\mu \in \mathcal{D}$,

$$s_N(\mu) = \ell(u_N(\mu)) + g(u_N(\mu), \psi_{N^{\mathrm{du}}}(\mu); \mu), \tag{9}$$

where $u_N(\mu) \in W_N$ and $\psi_{N^{\mathrm{du}}}(\mu) \in W_{N^{\mathrm{du}}}^{\mathrm{du}}$ satisfy

$$a(u_N(\mu), v; \mu) = f(v), \quad \forall\, v \in W_N, \tag{10}$$

and

$$a(v, \psi_{N^{\mathrm{du}}}(\mu); \mu) = -\ell(v), \quad \forall\, v \in W_{N^{\mathrm{du}}}^{\mathrm{du}}, \tag{11}$$

*In actual practice, the primal and dual bases should be orthogonalized with respect to the inner product associated with the Hilbert space X, $(\cdot, \cdot)_X$; the algebraic systems then inherit the "conditioning" properties of the underlying partial differential equation.

respectively. We emphasize that we are interested in global approximations that are *uniformly valid* over a finite parameter domain $\mathcal{D}$.

We note that, in the compliance case – a symmetric and $\ell = f$ such that $\psi(\mu) = -u(\mu)$ – we may simply take $N^{\rm du} = N$, $S_N^{\rm du} = S_N$, $W_N^{\rm du} = W_N$, and hence $\psi_N(\mu) = -u_N(\mu)$. In practice, in such a case we need never actually form the dual problem – we simply identify $\psi_N(\mu) = -u_N(\mu)$ – with a corresponding 50% reduction in computational effort.

Typically [8, 9], and in some very special cases provably [10], $u_N(\mu)$, $\psi_N(\mu)$, and $s_N(\mu)$ converge to $u(\mu)$, $\psi(\mu)$, and $s(\mu)$ uniformly and extremely rapidly – thanks to the smoothness in μ – and thus we may achieve the desired accuracy for N, $N^{\rm du} \ll \mathcal{N}$. The critical ingredients of the *a priori* theory are (*i*) the optimality properties of Galerkin projection,* and (*ii*) the good approximation properties of W_N (respectively, $W_{N^{\rm du}}^{\rm du}$) for the manifold $\mathcal{M} \equiv \{u(\mu) \mid \mu \in \mathcal{D}\}$ (respectively, $\mathcal{M}^{\rm du} \equiv \{\psi(\mu) \mid \mu \in \mathcal{D}\}$).

2.2. Offline–Online Computational Procedure

Even though N, $N^{\rm du}$ may be small, the elements of (say) W_N are in some sense "large": $\zeta_n \equiv u(\mu_n^{\rm pr})$ will be represented in terms of $\mathcal{N} \gg N$ truth finite element basis functions. To eliminate the $\mathcal{N}$-contamination of the *deployed* performance, we must consider offline–online computational procedures [7–9, 11]. For our purposes here, we continue to assume that our PDE is linear, (6), and furthermore exactly affine, (7), for some modest Q. In future sections we shall consider a nonlinear example as well as the possibility of nonaffine operators.

To begin, we expand our reduced-basis approximation as

$$u_N(\mu) = \sum_{j=1}^{N} u_{N\,j}(\mu)\zeta_j, \quad \psi_{N^{\rm du}}(\mu) = \sum_{j=1}^{N^{\rm du}} \psi_{N^{\rm du}\,j}(\mu)\zeta_j^{\rm du}. \tag{12}$$

It then follows from (9) and (12) that the reduced-basis output can be expressed as

$$\begin{aligned} s_N(\mu) = &\sum_{j=1}^{N} u_{N\,j}(\mu)\,\ell(\zeta_j) - \sum_{j=1}^{N^{\rm du}} \psi_{N^{\rm du}\,j}(\mu) f(\zeta_j^{\rm du}) \\ &+ \sum_{j=1}^{N}\sum_{j'=1}^{N^{\rm du}}\sum_{q=1}^{Q} u_{N\,j}(\mu)\psi_{N^{\rm du}\,j'}(\mu)\Theta^q(\mu)a^q(\zeta_j, \zeta_{j'}^{\rm du}), \end{aligned} \tag{13}$$

*Galerkin optimality relies on stability of the discrete equations. The latter is only assured for coercive problems; for noncoercive problems, Petrov–Galerkin methods may thus be preferred [12].

where the coefficients $u_{N\,j}(\mu)$, $1 \le j \le N$, and $\psi_{N^{\rm du}\,j}$, $1 \le j \le N^{\rm du}$, satisfy the $N \times N$ and $N^{\rm du} \times N^{\rm du}$ linear algebraic systems

$$\sum_{j=1}^{N} \left\{ \sum_{q=1}^{Q} \Theta^q(\mu) a^q(\zeta_j, \zeta_i) \right\} u_{N\,j}(\mu) = f(\zeta_i), \quad 1 \le i \le N, \tag{14}$$

$$\sum_{j=1}^{N^{\rm du}} \left\{ \sum_{q=1}^{Q} \Theta^q(\mu) a^q(\zeta_i^{\rm du}, \zeta_j^{\rm du}) \right\} \psi_{N^{\rm du}\,j}(\mu) = -\ell(\zeta_i^{\rm du}), \quad 1 \le i \le N^{\rm du}. \tag{15}$$

The offline–online decomposition is now clear. For simplicity below we assume that $N^{\rm du} = N$.

In the offline stage – performed *once* – we first solve for the $\zeta_i, \zeta_i^{\rm du}$, $1 \le i \le N$; we then form *and store* $\ell(\zeta_i)$, $f(\zeta_i)$, $\ell(\zeta_i^{\rm du})$, and $f(\zeta_i^{\rm du})$, $1 \le i \le N$, and $a^q(\zeta_j, \zeta_i)$, $a^q(\zeta_i^{\rm du}, \zeta_j^{\rm du})$, $1 \le i, j \le N, 1 \le q \le Q$, and $a^q(\zeta_i, \zeta_j^{\rm du})$, $1 \le i, j \le N, 1 \le q \le Q$.* Note all quantities computed in the offline stage are independent of the parameter μ. In the online stage – performed *many times*, for each new value of μ "in the field" – we first assemble and subsequently invert the $N \times N$ "stiffness matrices" $\sum_{q=1}^{Q} \Theta^q(\mu)\, a^q(\zeta_j, \zeta_i)$ of (14) and $\sum_{q=1}^{Q} \Theta^q(\mu)\, a^q(\zeta_i^{\rm du}, \zeta_j^{\rm du})$ of (15) – this yields the $u_{N\,j}(\mu)$, $\psi_{N^{\rm du}\,j}(\mu)$, $1 \le j \le N$; we next perform the summation (13) – this yields the $s_N(\mu)$. The operation count for the online stage is, respectively, $O(QN^2)$ and $O(N^3)$ to assemble (recall that the $a^q(\zeta_j, \zeta_i)$, $1 \le i, j \le N$, $1 \le q \le Q$, are *pre-stored*) and invert the stiffness matrices, and $O(N) + O(QN^2)$ to evaluate the output (recall that the $\ell(\zeta_j)$ are *pre-stored*); note that the RB stiffness matrix is, in general, *full*.

The essential point is that the online complexity is *independent of* $\mathcal{N}$, the dimension of the underlying truth finite element approximation space. Since $N, N^{\rm du} \ll \mathcal{N}$, we expect – and often realize – significant, orders-of-magnitude computational economies relative to classical discretization approaches.

3. A Posteriori Error Estimation

3.1. Motivation

A posteriori error estimation procedures are very well developed for classical approximations of, and solution procedures for, (say) partial differential equations [13–15] and algebraic systems [16]. However, until quite recently,

*In actual practice, in the offline stage we consider $N = N_{\max}$ and $N^{\rm du} = N^{\rm du}_{\max}$; then, in the online stage, we extract the necessary subvectors and submatrices.

there has been essentially no way to rigorously, quantitatively, sharply, and efficiently assess the accuracy of RB approximations. As a result, for any given new μ, the RB (say, primal) solution $u_N(\mu)$ typically *raises* many more questions than it *answers*. Is there even a solution $u(\mu)$ near $u_N(\mu)$? This question is particularly crucial in the nonlinear context – for which in general we are guaranteed neither existence nor uniqueness. Is $|s(\mu)-s_N(\mu)| \leq \epsilon_{\text{tol}}$, where ϵ_{tol} is the maximum acceptable error? Is a crucial feasibility condition $s(\mu) \leq C$ (say, in a constrained optimization exercise) satisfied – not just for the RB approximation, $s_N(\mu)$, but also for the "true" output, $s(\mu)$? If these questions cannot be affirmatively answered, we may propose the wrong – and unsafe or infeasible – action in the deployed context. A fourth question is also important: Is N too large, $|s(\mu) - s_N(\mu)| \ll \epsilon_{\text{tol}}$, with an associated steep (N^3) penalty on computational efficiency? An overly conservative approximation may jeopardize the real-time response and associated action – with corresponding detriment to the deployed systems.

We may also consider the approximation properties and efficiency of the (say, primal) parameter samples and associated RB approximation spaces, S_N and W_N, $1 \leq N \leq N_{\max}$. Do we satisfy our global "acceptable error level" condition, $|s(\mu) - s_N(\mu)| \leq \epsilon_{\text{tol}}, \forall \mu \in \mathcal{D}$, for (close to) the smallest possible value of N? And a related question: For our given tolerance ϵ_{tol}, are the RB stiffness matrices (or, in the nonlinear case, Newton Jacobians) as well-conditioned as possible – given that *by construction* W_N will be increasingly colinear with increasing N? If the answers are not affirmative, then our RB approximations are more expensive (and unstable) than necessary – and perhaps too expensive to provide real-time response.

In short, the pre-asymptotic and essentially *ad hoc* or empirical nature of reduced-basis discretizations, the strongly superlinear scaling (with N, N^{du}) of the reduced-basis online complexity, and the particular needs of deployed real-time systems virtually demand rigorous *a posteriori* error estimators. Absent such certification, we must either err on the side of computational pessimism – and compromise real-time response – or err on the side of computational optimism – and risk sub-optimal, infeasible, or potentially unsafe decisions.

In Refs. [8, 9, 17, 18], we introduce a family of rigorous error estimators for reduced-basis approximation of a wide class of partial differential equations (see also Ref. [19] for an alternative approach). As in almost all error estimation contexts, the enabling (trivial) observation is that, whereas a 100% error in the *field variable* $u(\mu)$ or output $s(\mu)$ is clearly unacceptable, a 100% or even larger (conservative) error in the *error* is tolerable and not at all useless; we may thus pursue "relaxations" of the equation governing the error and residual that would be bootless for the original equation governing the field variable $u(\mu)$.

We now present further details for the particular case of elliptic linear problems with exact affine parameter dependence (7): the truth solution satisfies

(4), (6), and (8), and the corresponding reduced-basis approximation satisfies (9)–(11). (In subsequent sections we shall consider the extension to nonlinear problems through a detailed example; we shall also briefly discuss nonaffine problems.)

3.2. Error Bounds

We shall need several preliminary definitions. To begin, we denote the inner product and norm associated with our Hilbert space X as $(w, v)_X$ and $\|v\|_X = \sqrt{(v, v)_X}$, respectively; we further define the dual norm (of any bounded linear functional h) as

$$\|h\|_{X'} \equiv \sup_{v \in X} \frac{h(v)}{\|v\|_X}. \tag{16}$$

We recall that we restrict our attention here to second-order elliptic partial differential equations: thus, for a scalar problem (such as heat conduction), $H_0^1(\Omega) \subset X^e \subset H^1(\Omega)$, where $H^1(\Omega)$ (respectively, $H_0^1(\Omega)$) is the usual space of derivative-square-integrable functions (respectively, derivative–square–integrable functions that vanish on $\partial\Omega$, the boundary of Ω) [20]. A typical choice for $(\cdot, \cdot)_X$ is

$$(w, v)_X = \int_\Omega \nabla w \cdot \nabla v + wv, \tag{17}$$

which is simply the standard $H^1(\Omega)$ inner product.

We next introduce [12, 18] the operator $T^\mu : X \to X$ such that, for any w in X, $(T^\mu w, v)_X = a(w, v; \mu)$, $\forall\, v \in X$. We then define

$$\sigma(w; \mu) \equiv \frac{\|T^\mu w\|_X}{\|w\|_X},$$

and note that

$$\beta(\mu) \equiv \inf_{w \in X} \sup_{v \in X} \frac{a(w, v; \mu)}{\|w\|_X \|v\|_X} = \inf_{w \in X} \sigma(w; \mu), \tag{18}$$

$$\gamma(\mu) \equiv \sup_{w \in X} \sup_{v \in X} \frac{a(w, v; \mu)}{\|w\|_X \|v\|_X} = \sup_{w \in X} \sigma(w; \mu); \tag{19}$$

we also recall that

$$\beta(\mu)\|w\|_X \|T^\mu w\|_X \le a(w, T^\mu w; \mu), \qquad \forall\, w \in X. \tag{20}$$

Here $\beta(\mu)$ is the Babuška "inf–sup" stability constant – the minimum singular value associated with our differential operator (and transpose operator) – and

$\gamma(\mu)$ is the standard continuity constant. We suppose that $\gamma(\mu)$ is bounded $\forall\ \mu \in \mathcal{D}$, and that $\beta(\mu) \geq \beta_0 > 0$, $\forall\ \mu \in \mathcal{D}$. We note that for a symmetric, coercive bilinear form, $\beta(\mu) = \alpha_c(\mu)$, where

$$\alpha_c(\mu) \equiv \inf_{w \in X} \frac{a(w, w; \mu)}{\|w\|_X^2},$$

is the standard coercivity constant.

Given our reduced-basis primal solution $u_N(\mu)$, it is readily derived that the error $e(\mu) \equiv u(\mu) - u_N(\mu) \in X$ satisfies

$$a(e(\mu), v; \mu) = -g(u_N(\mu), v; \mu), \quad \forall\ v \in X, \tag{21}$$

where $-g(u_N(\mu), v; \mu) \equiv f(v) - a(u_N(\mu), v; \mu)$ (in this linear case) is the familiar residual. It then follows from (16), (20), and (21) that

$$\|e(\mu)\|_X \leq \frac{\varepsilon_N(\mu)}{\beta(\mu)},$$

where

$$\varepsilon_N(\mu) \equiv \|g(u_N(\mu), \cdot\,; \mu)\|_{X'}, \tag{22}$$

is the dual norm of the residual.

We now assume that we are privy to a nonnegative lower bound for the inf–sup parameter, $\tilde{\beta}(\mu)$, such that $\beta(\mu) \geq \tilde{\beta}(\mu) \geq \epsilon_\beta \beta(\mu)$, $\forall \mu \in \mathcal{D}$, where $\epsilon_\beta \in]0, 1[$. We then introduce our "energy" error bound

$$\Delta_N(\mu) \equiv \frac{\varepsilon_N(\mu)}{\tilde{\beta}(\mu)}, \tag{23}$$

the effectivity of which is defined as

$$\eta_N(\mu) \equiv \frac{\Delta_N(\mu)}{\|e(\mu)\|_X}.$$

It is readily proven [9, 18] that, for any N, $1 \leq N \leq N_{\max}$,

$$1 \leq \eta_N(\mu) \leq \frac{\gamma(\mu)}{\tilde{\beta}(\mu)}, \quad \forall\ \mu \in \mathcal{D}. \tag{24}$$

From the left inequality, we deduce that $\|e(\mu)\|_X \leq \Delta_N(\mu)$, $\forall \mu \in \mathcal{D}$, and hence that $\Delta_N(\mu)$ is a rigorous upper bound for the true error* measured in the $\|\cdot\|_X$ norm – this provides certification: feasibility and "safety" are guaranteed. From the right inequality, we deduce that $\Delta_N(\mu)$ overestimates the true

*Note, however, that these error bounds are relative to our underlying "truth" approximation, $u(\mu) \in X$, not to the exact solution, $u^e(\mu) \in X^e$.

error by at most $\gamma(\mu)/\tilde{\beta}(\mu)$,* *independent of* N – this relates to efficiency: an overly conservative error bound will be manifested in an unnecessarily large N and unduly expensive RB approximation, or (even worse) an overly conservative or expensive decision or action "in the field."

We now turn to error bounds for the output of interest. To begin, we note that the dual satisfies an "energy" error bound very similar to the primal result: for $1 \le N^{\mathrm{du}} \le N^{\mathrm{du}}_{\max}$,

$$\|\psi(\mu) - \psi_{N^{\mathrm{du}}}(\mu)\|_X \le \Delta^{\mathrm{du}}_N(\mu), \quad \forall\, \mu \in \mathcal{D};$$

here $\Delta^{\mathrm{du}}_N \equiv \varepsilon^{\mathrm{du}}_N(\mu)/\tilde{\beta}(\mu)$, and $\varepsilon^{\mathrm{du}}_N(\mu) = \| - \ell(\cdot) - a(\cdot, \psi_{N^{\mathrm{du}}}(\mu); \mu)\|_{X'}$ is the dual norm of the dual residual. It then follows† that

$$|s(\mu) - s_N(\mu)| \le \Delta^s_N(\mu), \quad \forall \mu \in \mathcal{D}, \tag{25}$$

where

$$\Delta^s_N(\mu) \equiv \varepsilon_N(\mu)\Delta^{\mathrm{du}}_N(\mu). \tag{26}$$

It is critical to note that $\Delta^s_N(\mu) = \tilde{\beta}(\mu)\Delta_N(\mu)\Delta^{\mathrm{du}}_N(\mu)$: the output error (and output error bound) vanishes as the *product* of the primal and dual error (bounds), and hence much more rapidly than either the primal or dual error. From the perspective of computational efficiency, a good choice is $\varepsilon_N(\mu) \approx \varepsilon^{\mathrm{du}}_N(\mu)$; the latter also (roughly) ensures that the bound (25), (26) will be quite sharp.

In the compliance case, a symmetric and $\ell = f$, we immediately obtain $\Delta^{\mathrm{du}}_N(\mu) = \Delta_N(\mu)$, and hence (25) obtains for

$$\Delta^s_N(\mu) \equiv \frac{\varepsilon^2_N(\mu)}{\tilde{\beta}(\mu)}, \quad \forall\, \mu \in \mathcal{D} \text{ (compliance)}; \tag{27}$$

here, we obtain the "square" effect even without (explicit) introduction of the dual problem. For a coercive further improvements are possible [9].

The real challenge in *a posteriori* error estimation is not the presentation of these rather classical results, but rather the development of efficient computational approaches for the evaluation of the necessary constituents. In our particular deployed context, "efficient" translates to "online complexity *independent* of $\mathcal{N}$," and "necessary constituents" translates to "dual norm of the primal residual, $\varepsilon_N(\mu) \equiv \|g(u_N(\mu), \cdot; \mu)\|_{X'}$, dual norm of the dual residual, $\varepsilon^{\mathrm{du}}_N(\mu) \equiv \| - \ell(\cdot) - a(\cdot, \psi_{N^{\mathrm{du}}}(\mu); \mu)\|_{X'}$, and lower bound for the inf–sup constant, $\tilde{\beta}(\mu)$." We now turn to these issues.

*The upper bound on the effectivity can be large. In many cases, this effectivity bound is in fact quite pessimistic; in many other cases, the effectivity (bound) may be improved by judicious choice of (multi-point) inner product $(\cdot, \cdot)_X$ – in effect, a "bound conditioner" [21].

†The proof is simple: $|s(\mu) - s_N(\mu)| = |\ell(e) - g(u_N(\mu), \psi_N(\mu); \mu)| = | - a(e(\mu), \psi(\mu); \mu) - g(u_N(\mu), \psi_N(\mu); \mu)| = |g(u_N(\mu), \psi(\mu) - \psi_N(\mu); \mu)| \le \varepsilon_N(\mu)\Delta^{\mathrm{du}}_N(\mu)$.

3.3. Offline–Online Computational Procedures

3.3.1. The dual norm of the residual

We consider only the primal residual; the dual residual admits a similar treatment. To begin, we note from standard duality arguments that

$$\varepsilon_N(\mu) \equiv \|g(u_N(\mu), \cdot; \mu)\|_{X'} = \|\hat{e}(\mu)\|_X, \tag{28}$$

where $\hat{e}(\mu) \in X$ satisfies

$$(\hat{e}(\mu), v)_X = -g(u_N(\mu), v; \mu), \quad \forall\, v \in X. \tag{29}$$

We next observe from our reduced-basis representation (12) and affine assumption (7) that $-g(u_N(\mu), v; \mu)$ may be expressed as

$$-g(u_N(\mu), v; \mu) = f(v) - \sum_{q=1}^{Q} \sum_{n=1}^{N} \Theta^q(\mu) u_{N\,n}(\mu) a^q(\zeta_n, v), \quad \forall v \in X. \tag{30}$$

It thus follows from (29) and (30) that $\hat{e}(\mu) \in X$ satisfies

$$(\hat{e}(\mu), v)_X = f(v) - \sum_{q=1}^{Q} \sum_{n=1}^{N} \Theta^q(\mu)\, u_{N\,n}(\mu)\, a^q(\zeta_n, v), \quad \forall\, v \in X. \tag{31}$$

The critical observation [8, 9] is that the right-hand side of (31) is a sum of products of parameter-dependent functions and parameter-independent linear functionals.

In particular, it follows from linear superposition that we may write $\hat{e}(\mu) \in X$ as

$$\hat{e}(\mu) = \mathcal{C} + \sum_{q=1}^{Q} \sum_{n=1}^{N} \Theta^q(\mu)\, u_{N\,n}(\mu)\, \mathcal{L}_n^q,$$

for $\mathcal{C} \in X$ satisfying $(\mathcal{C}, v)_X = f(v)$, $\forall\, v \in X$, and $\mathcal{L}_n^q \in X$ satisfying $(\mathcal{L}_n^q, v)_X = -\,a^q(\zeta_n, v)$, $\forall\, v \in X$, $1 \le n \le N$, $1 \le q \le Q$; note from (17) that the latter are simple *parameter-independent* (scalar or vector) Poisson, or Poisson-like, problems. It thus follows that

$$\|\hat{e}(\mu)\|_X^2 = (\mathcal{C}, \mathcal{C})_X + \sum_{q=1}^{Q} \sum_{n=1}^{N} \Theta^q(\mu)\, u_{N\,n}(\mu) \Bigg\{ 2(\mathcal{C}, \mathcal{L}_n^q)_X + \sum_{q'=1}^{Q} \sum_{n'=1}^{N} \Theta^{q'}(\mu)\, u_{N\,n'}(\mu)\, (\mathcal{L}_n^q, \mathcal{L}_{n'}^{q'})_X \Bigg\}. \tag{32}$$

The expression (32) – which we relate to the requisite dual norm of the residual through (28) – is the sum of products of parameter-dependent (simple, known) functions and parameter-independent inner products. The offline–online decomposition is now clear.

In the offline stage – performed once – we first solve for $\mathcal{C}$ and $\mathcal{L}_n^q$, $1 \le n \le N$, $1 \le q \le Q$; we then evaluate and save the relevant parameter-independent inner products $(\mathcal{C}, \mathcal{C})_X$, $(\mathcal{C}, \mathcal{L}_n^q)_X$, $(\mathcal{L}_n^q, \mathcal{L}_{n'}^{q'})_X$, $1 \le n, n' \le N$, $1 \le q, q' \le Q$. Note that all quantities computed in the offline stage are independent of the parameter μ. In the online stage – performed many times, for each new value of μ "in the field" – we simply evaluate the sum (32) in terms of the $\Theta^q(\mu)$, $u_{N\,n}(\mu)$ and the precalculated and stored (parameter-independent) $(\cdot, \cdot)_X$ inner products. The operation count for the online stage is only $O(Q^2N^2)$ – again, the essential point is that the online complexity is *independent of* $\mathcal{N}$, the dimension of the underlying truth finite element approximation space. We further note that, unless Q is quite large, the online cost associated with the calculation of the dual norm of the residual is commensurate with the online cost associated with the calculation of $s_N(\mu)$.

3.3.2. Lower bound for the inf–sup parameter

Obviously, from the definition (18), we may readily obtain by a variety of techniques effective *upper bounds* for $\beta(\mu)$; however, lower bounds are much more difficult to construct. We do note that in the case of *symmetric coercive* operators we *can* often determine $\tilde{\beta}(\mu)$ ($\le \beta(\mu) = \alpha_c(\mu), \forall \mu \in \mathcal{D}$) "by inspection." For example, if we verify $\Theta^q(\mu) > 0, \forall\, \mu \in \mathcal{D}$, and $a^q(v, v) \ge 0, \forall\, v \in X$, $1 \le q \le Q$, then we may choose [8, 21] for our coercivity lower bound

$$\tilde{\beta}(\mu) = \left(\min_{q \in \{1,\dots,Q\}} \frac{\Theta^q(\mu)}{\Theta^q(\bar{\mu})} \right) \alpha_c(\bar{\mu}), \tag{33}$$

for some $\bar{\mu} \in \mathcal{D}$. Unfortunately, these hypotheses are rather restrictive, and hence more complicated (and offline-expensive) recipes must often be pursued [17, 18]. We consider here a construction which is valid for general noncoercive operators (and thus also relevant in the nonlinear context [22]); for simplicity, we assume our problem remains well-posed over a convex parameter set that includes $\mathcal{D}$.

To begin, given $\bar{\mu} \in \mathcal{D}$ and $t = (t_{(1)}, \dots, t_{(P)}) \in \mathbb{R}^P$ – note $t_{(j)}$ is the value of the jth component of t – we introduce the bilinear form

$$\mathcal{T}(w, v; t; \bar{\mu}) = (T^{\bar{\mu}} w, T^{\bar{\mu}} v)_X + \sum_{p=1}^{P} t_{(p)} \left\{ \sum_{q=1}^{Q} \frac{\partial \Theta^q}{\partial \mu_{(p)}}(\bar{\mu}) \left[a^q(w, T^{\bar{\mu}} v) + a^q(v, T^{\bar{\mu}} w) \right] \right\} \tag{34}$$

and associated Rayleigh quotient

$$\mathcal{F}(t;\bar{\mu}) = \min_{v\in X} \frac{\mathcal{T}(v,v;t;\bar{\mu})}{\|v\|_X^2}; \tag{35}$$

it is readily demonstrated that $\mathcal{F}(t;\bar{\mu})$ is concave in t [24], and hence $\mathcal{D}^{\bar{\mu}} \equiv \{\mu \in \mathbb{R}^P | \mathcal{F}(\mu-\bar{\mu};\bar{\mu}) \geq 0\}$ is perforce convex. We next introduce semi-norms $|\cdot|_q : X \to \mathbb{R}_{+,0}$ such that

$$\begin{aligned} &|a^q(w,v)| \leq \Gamma^q |w|_q |v|_q, \quad \forall w, v \in X,\ 1 \leq q \leq Q, \\ &C_X = \sup_{w\in X} \frac{\sum_{q=1}^{Q} |w|_q^2}{\|w\|_X^2}, \end{aligned} \tag{36}$$

for positive parameter-independent constants Γ^q, $1 \leq q \leq Q$, and C_X; it is often the case that $\Theta^1(\mu) = \text{Constant}$, in which case the $q = 1$ contribution to the sum in (34) and (36) may be discarded. (Note that C_X is typically *independent of* Q, since the a^q are often associated with non-overlapping subdomains of Ω.) Finally, we define

$$\begin{aligned} \Phi(\mu;\bar{\mu}) \equiv C_X \max_{q\in\{1,\ldots,Q\}} \Bigg(\Gamma^q \Bigg| &\Theta^q(\mu) - \Theta^q(\bar{\mu}) \\ &- \sum_{p=1}^{P} (\mu-\bar{\mu})_{(p)} \frac{\partial \Theta^q}{\partial \mu_{(p)}}(\bar{\mu}) \Bigg| \Bigg), \end{aligned} \tag{37}$$

for $\mu \equiv (\mu_{(1)}, \ldots, \mu_{(P)}) \in \mathbb{R}^P$.

We now introduce points $\bar{\mu}_j$ and associated polytopes $\mathcal{P}^{\bar{\mu}_j} \subset \mathcal{D}^{\bar{\mu}_j}$, $1 \leq j \leq J$, such that

$$\mathcal{D} \subset \bigcup_{j=1}^{J} \mathcal{P}^{\bar{\mu}_j}, \tag{38}$$

$$\min_{v\in\mathcal{V}^{\bar{\mu}_j}} \sqrt{\mathcal{F}(v-\bar{\mu}_j;\bar{\mu}_j)} - \max_{\mu\in\mathcal{P}^{\bar{\mu}_j}} \Phi(\mu;\bar{\mu}_j) \geq \epsilon_\beta \beta(\bar{\mu}^j), \quad 1 \leq j \leq J. \tag{39}$$

Here $\mathcal{V}^{\bar{\mu}_j}$ is the set of vertices associated with the polytope $\mathcal{P}^{\bar{\mu}_j}$ – for example, $\mathcal{P}^{\bar{\mu}_j}$ may be a simplex with $|\mathcal{V}^{\bar{\mu}_j}| = P+1$ vertices; and $\epsilon_\beta \in]0,1[$ is a prescribed accuracy constant. Our lower bound is then given by

$$\tilde{\beta}(\mu) = \max_{j\in\{1,\ldots,J\}|\mu\in\mathcal{P}^{\bar{\mu}_j}} \epsilon_\beta \beta(\bar{\mu}_j). \tag{40}$$

In fact, $\tilde{\beta}(\mu)$ of (40) may not strictly honor our condition $\tilde{\beta}(\mu) > \epsilon_\beta \beta(\mu)$; however, as the latter relates to accuracy, approximate satisfaction suffices.

(Recall that $\tilde{\beta}(\mu)$ appears in the denominator of our *error bound*; hence, even a relative inf–sup discrepancy of 80%, $\epsilon_\beta \approx 1/5$, is acceptable.) It can be easily demonstrated that $\beta(\mu) \geq \tilde{\beta}(\mu) \geq \epsilon_\beta \beta_0 > 0$, $\forall \mu \in \mathcal{D}$, which thus ensures well-posed and rigorous error bounds.

We now turn to the offline–online decomposition. The *offline* stage comprises two parts: the *generation* of a set of points and polytopes–vertices, $\bar{\mu}_j$ and $\mathcal{P}^{\bar{\mu}_j}$, $\mathcal{V}^{\bar{\mu}_j}$, $1 \leq j \leq J$; and the *verification* that (38) (trivial) and (39) (nontrivial) are indeed satisfied. We focus on verification; generation – quite involved – is described in detail in [23]. To verify (39), the essential observation is that the expensive terms – "truth" eigenproblems associated with $\mathcal{F}$, (35), and β, (18) – are limited to a finite set of *vertices*,

$$J + \sum_{j=1}^{J} |\mathcal{V}^{\bar{\mu}_j}|,$$

in total; only for the extremely inexpensive – and typically algebraically very simple – $\Phi(\mu; \bar{\mu}_j)$ terms must we consider minimization over the *polytopes*. The *online* stage (40) is very simple: a search/look-up table, with complexity logarithmic in J and polynomial in P.

We close by remarking on the properties of $\mathcal{F}(\mu - \bar{\mu}; \bar{\mu})$ that play an important role. First, $\mathcal{F}(\mu - \bar{\mu}; \bar{\mu}) \leq \beta^2(\mu)$, $\forall \mu \in \mathcal{D}^{\bar{\mu}}$ (say, for the case in which $\Theta^q(\mu) = \mu_{(q)}$, $1 \leq q \leq Q = P$): this ensures the lower bound result. Second, $\mathcal{F}(t; \bar{\mu})$ is concave in t (note that in general $\beta(\mu)$ is neither (quasi-) concave nor (quasi-) convex in μ [24]): this ensures a tractable offline computation. Third, $\mathcal{F}(\mu - \bar{\mu}; \bar{\mu})$ is "tangent"* to $\beta(\mu)$ at $\mu = \bar{\mu}$ – the cruder estimate $\Phi(\mu; \bar{\mu})$ is a *second-order correction*: this controls the growth of J (for example, relative to simpler continuity bounds [17]).

3.4. Sample Construction: A Greedy Algorithm

Our error estimation procedures also allow us to pursue more rational constructions of our parameter samples S_N, $S_{N^{\mathrm{du}}}^{\mathrm{du}}$ (and hence spaces W_N, $W_{N^{\mathrm{du}}}^{\mathrm{du}}$) [18]. We consider here only the primal problem – in which our error criterion is $\|u(\mu) - u_N(\mu)\|_X \equiv \|e(\mu)\|_X \leq \epsilon_{\mathrm{tol}}$; similar approaches may be developed for the dual – $\|\psi(\mu) - \psi_{N^{\mathrm{du}}}(\mu)\|_X \leq \epsilon_{\mathrm{tol}}^{\mathrm{du}}$, and hence the output – $|s(\mu) - s_N(\mu)| \leq \epsilon_{\mathrm{tol}}^{s}$. We denote the smallest primal error tolerance anticipated as $\epsilon_{\mathrm{tol,\,min}}$ – this must be determined *a priori* offline; we then permit $\epsilon_{\mathrm{tol}} \in [\epsilon_{\mathrm{tol,\,min}}, \infty[$ to be specified online. We also introduce $\Xi_{\mathrm{F}} \in \mathcal{D}^{n_{\mathrm{F}}}$, a very fine random sample over the parameter domain $\mathcal{D}$ of size $n_{\mathrm{F}} \gg 1$.

*To make this third property rigorous we must in general consider non-smooth analysis and also possibly a continuous spectrum as $\mathcal{N} \to \infty$.

We first consider the offline stage. We assume that we are given a sample S_N, and hence space W_N and associated reduced-basis approximation (procedure to determine) $u_N(\mu)$, $\forall \mu \in \mathcal{D}$. We then calculate $\mu_N^* = \arg\max_{\mu \in \Xi_F} \Delta_N(\mu) - \Delta_N(\mu)$ is our "online" error bound (23) that, in the limit of $n_F \to \infty$ queries, may be evaluated (on average) in $O(N^2 Q^2)$ operations; we next append μ_N^* to S_N to form S_{N+1}, and hence W_{N+1}. We now continue this process until $N = N_{\max}$ such that $\epsilon_{N_{\max}}^* = \epsilon_{\text{tol,min}}$, where $\epsilon_N^* \equiv \Delta_N(\mu_N^*)$, $1 \le N \le N_{\max}$.

In the online stage, given any desired $\epsilon_{\text{tol}} \in [\epsilon_{\text{tol, min}}, \infty[$ and any new value of $\mu \in \mathcal{D}$ "in the field", we first choose N from a pre-tabulated array such that $\epsilon_N^* \left(\equiv \Delta_N(\mu_N^*)\right) = \epsilon_{\text{tol}}$. We next calculate $u_N(\mu)$ and $\Delta_N(\mu)$, and then verify that – and if necessary, subsequently increase N *such that* – the condition $\Delta_N(\mu) \le \epsilon_{\text{tol}}$ is indeed satisfied. (We should not and do not rely on the finite sample Ξ_F for either rigor or sharpness.)

The crucial point is that $\Delta_N(\mu)$ is an accurate and "online-inexpensive" – $O(1)$ effectivity and $\mathcal{N}$-independent asymptotic complexity – surrogate for the true (very-expensive-to-calculate) error $\|u(\mu) - u_N(\mu)\|_X$. This surrogate permits us to (*i*) offline – here we exploit low average cost – perform a much more exhaustive ($n_F \gg 1$) and, hence, meaningful search for the best samples S_N and, hence, most rapidly *uniformly* convergent spaces W_N,* and (*ii*) online – here we exploit low marginal cost – determine the smallest N, and hence, the most efficient approximation, for which we *rigorously* achieve the desired accuracy.

4. A Linear Example: Helmholtz-Elasticity

4.1. Problem Description

We consider a two-dimensional thin plate with a horizontal crack at the (say) interface of two lamina: the (original) domain $\Omega^o(z, L) \subset \mathbb{R}^2$, shown in Fig. 1, is defined as $[0, 2] \times [0, 1] \setminus \Gamma_C^o$, where $\Gamma_C^o \equiv \{x_1 \in [b - L/2, b + L/2], x_2 = 1/2\}$ defines the idealized crack. The left surface of the plate is secured; the top and bottom boundaries are stress-free; and the right boundary is subjected to a vertical oscillatory uniform traction at frequency ω. We model the plate as plane-stress linear isotropic elastic with (scaled) density unity, Young's modulus unity, and Poisson ratio 0.25; the latter determine the (parameter-independent) constitutive tensor $E_{ijk\ell}$. Our $P = 3$ input is $\mu \equiv (\mu_{(1)}, \mu_{(2)}, \mu_{(3)}) \equiv (\omega^2, b, L)$; our output is the (oscillatory) amplitude of the average vertical displacement on the right edge of the plate.

*We may in fact view our offline sampling process as a (greedy, parameter space, "$L^\infty(\mathcal{D})$") variant of the POD economization procedure [25] in which – thanks to $\Delta_N(\mu)$ – *we need never construct* the "rejected" snapshots.

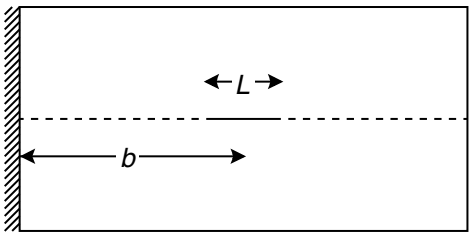

Figure 1. (Original) domain for the Helmholtz elasticity example.

The governing equation for the displacement $u^{\rm o}(x^{\rm o};\mu) \in X^{\rm o}(\mu)$ is therefore $a^{\rm o}(u^{\rm o}(\mu), v;\mu) = f^{\rm o}(v)$, $\forall\, v \in X^{\rm o}(\mu)$, where $X^{\rm o}(\mu)$ is a quadratic finite element truth approximation subspace (of dimension $\mathcal{N} = 14{,}662$) of $X^{\rm e}(\mu) \equiv \{v \in (H^1(\Omega^{\rm o}(b, L)))^2 \mid v|_{x_1^{\rm o}=0} = 0\}$; here

$$a^{\rm o}(w, v;\mu) \equiv \int_{\Omega^{\rm o}(b,L)} w_{i,j} E_{ijk\ell} v_{k,\ell} - \omega^2 w_i v_i,$$

($v_{i,j}$ denotes $\partial v_i/\partial x_j$ and repeated physical indices imply summation), and $f^{\rm o}(v) \equiv \int_{x_1^{\rm o}=2} v_2$. The crack surface is hence modeled extremely simplistically – as a stress-free boundary. The output $s^{\rm o}(\mu)$ is given by $s^{\rm o}(\mu) = \ell^{\rm o}(u^{\rm o}(\mu))$, where $\ell^{\rm o}(v) = f^{\rm o}(v)$; we are thus "in compliance".

We now map $\Omega^{\rm o}(b, L)$ via a continuous piecewise-affine transformation to a fixed domain Ω. This new problem can now be cast precisely in the desired abstract form, in which Ω, X, and $(w, v)_X$ are independent of the parameter μ: as required, all parameter dependence now enters through the bilinear and linear forms; in particular, our affine assumption (7) applies for $Q = 10$. In the Appendix we summarize the $\Theta^q(\mu)$, $a^q(w, v)$, $1 \le q \le Q$; the bound conditioner $(\cdot, \cdot)_X$; and the resulting continuity constants Γ^q and semi-norms $|\cdot|_q$, $1 \le q \le Q$, and norm equivalence parameter C_X.

The (undamped, nonradiating) Helmholtz equation exhibits resonances. Our techniques can treat near resonances, as well as large frequency ranges, quite well [18, 23]. For our illustrative purposes here, we choose the parameter domain $\mathcal{D}$ ($\subset \mathbb{R}^{P=3}$) $\equiv (\omega^2 \in [3.2, 4.8]) \times (b \in [0.9, 1.1]) \times (L \in [0.15, 0.25])$; $\mathcal{D}$ contains no resonances – $\beta(\mu) \ge \beta_0 > 0$, $\forall \mu \in \mathcal{D}$ – however, $\omega^2 = 3.2$ and 4.8 are close to corresponding natural frequencies, and hence the problem is distinctly noncoercive.

4.2. Numerical Results

We first consider the inf–sup lower bound construction. We show in Fig. 2 $\beta^2(\mu)$ and $\mathcal{F}(\mu - \bar{\mu}^1; \bar{\mu}^1)$ for $\bar{\mu}^1 = (4.0, 1.0, 0.2)$; for purposes of presentation we keep $\mu_{(1)} = (\omega^2 = 4.0)$ fixed and vary $\mu_{(2)} (= b)$ and $\mu_{(3)} (= L)$. We observe

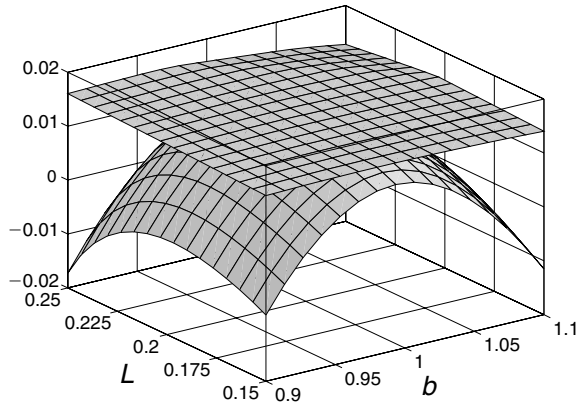

Figure 2. $\beta^2(\mu)$ and $\mathcal{F}(\mu - \bar{\mu}; \bar{\mu})$ for $\bar{\mu} = (4, 1, 0.2)$ as a function of (b, L); $\omega^2 = 4.0$.

that (in this particular case, even without $\Phi(\mu; \bar{\mu})$), $\mathcal{F}(\mu - \bar{\mu}; \bar{\mu})$ is a lower bound for $\beta^2(\mu)$; that $\mathcal{F}(\mu - \bar{\mu}; \mu)$ is concave; and that $\mathcal{F}(\mu - \bar{\mu}; \mu)$ is tangent to $\beta^2(\mu)$ at $\mu = \bar{\mu}$. Thanks to the latter, we can cover $\mathcal{D}$ (for $\bar{\epsilon}_\beta = 0.2$) such that (38) and (39) are satisfied with only $J = 84$ polytopes; in this particular case the $\mathcal{P}^{\bar{\mu}_j}$, $1 \le j \le J$, are hexahedrons such that $|\mathcal{V}^{\mu_j}| = 8$, $1 \le j \le J$.

Armed with the inf–sup lower bound, we can now pursue the adaptive sampling strategy described in the previous section. We recall that our problem is compliant, and hence we need only consider the primal variable (and subsequently set $\psi_{N^{\rm du}=N}(\mu) = -u_N(\mu)$ and $\varepsilon^{\rm du}_{N^{\rm du}=N}(\mu) = \varepsilon_N(\mu)$). For $\epsilon_{\rm tol,\,min} = 10^{-3}$ and $n_{\rm F} = 729$ we obtain $N_{\max} = 32$ such that $\epsilon_{N_{\max}} \equiv \Delta_{N_{\max}}(\mu^{\rm pr}_{N_{\max}}) = 9.03 \times 10^{-4}$.

We present in Table 1 $\Delta_{N,\max,\rm rel}$, $\eta_{N,\rm ave}$, $\Delta^s_{N,\max}$, and $\eta^s_{N,\rm ave}$ as a function of N. Here $\Delta_{N,\max,\rm rel}$ is the maximum over $\Xi_{\rm Test}$ of $\Delta_N(\mu)/\|u_{N_{\max}}\|_{\max}$, $\eta_{N,\rm ave}$ is the average over $\Xi_{\rm Test}$ of $\Delta_N(\mu)/\|u(\mu) - u_N(\mu)\|_X$, $\Delta^s_{N,\max,\rm rel}$ is the maximum over $\Xi_{\rm Test}$ of $\Delta^s_N(\mu)/|s_{N_{\max}}|_{\max}$, and $\eta^s_{N,\rm ave}$ is the average over $\Xi_{\rm Test}$ of $\Delta^s_N(\mu)/|s(\mu) - s_N(\mu)|$. Here $\Xi_{\rm Test} \in (\mathcal{D}^{\rm I})^{343}$ is a random parameter sample of size 343; $\|u_{N_{\max}}\|_{\max} \equiv \max_{\mu \in \Xi_{\rm Test}} \|u_{N_{\max}}(\mu)\|_X = 2.0775$ and $|s_{N_{\max}}|_{\max} \equiv \max_{\mu \in \Xi_{\rm Test}} |s_{N_{\max}}(\mu)| = 0.089966$; and $\Delta_N(\mu)$ and $\Delta^s_N(\mu)$ are given by (23) and (27), respectively. We observe that the RB approximation – in particular, for the output – converges very rapidly, and that our rigorous error bounds are in fact quite sharp. The effectivities are not quite $O(1)$ primarily due to the relatively crude inf–sup lower bound; but note that, thanks to the rapid convergence of the RB approximation, $O(10)$ effectivities do not significantly affect efficiency – the induced increase in RB dimension N is quite modest.

We turn now to computational effort. For (say) $N = 24$ and any given μ (say, $(4.0, 1.0, 0.2)$) – for which the error in the reduced-basis output $s_N(\mu)$

Table 1. Numerical results for Helmholtz elasticity

N	$\Delta_{N,\max,\text{rel}}$	$\eta_{N,\text{ave}}$	$\Delta^s_{N,\max,\text{rel}}$	$\eta^s_{N,\text{ave}}$
12	1.54×10^{-1}	13.41	3.31×10^{-2}	15.93
16	3.40×10^{-2}	12.24	2.13×10^{-3}	14.86
20	1.58×10^{-2}	13.22	4.50×10^{-4}	15.44
24	5.91×10^{-3}	12.56	4.81×10^{-5}	14.45
28	2.42×10^{-3}	12.44	9.98×10^{-6}	14.53

relative to the truth approximation $s(\mu)$ is *certifiably* less than $\Delta^s_N(\mu)$ ($= 4.94 \times 10^{-7}$) – the Online Time (marginal cost) to compute both $s_N(\mu)$ and $\Delta^s_N(\mu)$ is less than 0.0030 the Total Time to directly calculate the truth result $s(\mu) = \ell(u(\mu))$. The savings will be even larger for problems with more complex geometry and solution structure, in particular in three space dimensions.

As desired, we achieve efficiency due to (*i*) our choice of sample, (*ii*) our rigorous stopping criterion $\Delta^s_N(\mu)$, and (*iii*) our affine parameter dependence and associated offline–online computational procedures; and we achieve rigorous certainty – the reduced-basis predictions may serve in "deployed" decision processes with complete confidence (or at least with the same confidence as the underlying physical model and associated truth finite element approximation). The true merit of the approach is best illustrated in the deployed–real-time context of parameter identification (crack assessment) and adaptive mission optimization (load maximization); see Ref. [24] for an example.

5. A Nonlinear Example: Natural Convection

Obviously nonlinear equations do not admit the same degree of generality as linear equations. We thus present our approach to nonlinear equations for a particular quadratically nonlinear elliptic problem: the steady Boussinesq incompressible Navier–Stokes equations. This example permits us to identify the key new computational and theoretical ingredients; then, in Outlook, we contemplate more general (higher-order) nonlinearities.

5.1. Problem Description

We consider Prandtl number $\text{Pr} = 0.7$ Boussinesq natural convection in a square cavity $(x_1, x_2) \in \Omega \equiv [0, 1] \times [0, 1]$; the $\text{Pr} = 0$ limit is described in greater detail in [22, 26]. The governing equations for the velocity $U = (U_1, U_2)$, pressure p, and temperature θ are the (coupled) incompressible steady Navier–Stokes and thermal convection–diffusion equations. Our single parameter

($P = 1$) is the Grashof number, $\mu \equiv \mathrm{Gr}$, which is the ratio of the buoyancy forces (induced by the temperature field) to the momentum dissipation mechanisms; we consider $\mathrm{Gr} \in \mathcal{D} \equiv [1.0, 1.0 \times 10^4]$. This flow is a model problem for Bridgman growth of semi-conductor crystals; future work shall address geometric (angle, aspect ratio) and Pr variation, and higher Gr – all of which are important in actual materials processing applications.

In terms of the general mathematical formulation, (5), $u(\mu) \equiv (U_1, U_2, p, \theta, \lambda)(\mu)$, where λ is a Lagrange multiplier associated with the pressure zero-mean condition. Our solution $u(\mu)$ resides in the space $X \equiv X^U \times X^p \times X^\theta \times \mathbb{R}$, where $X^U \subset (H_0^1(\Omega))^2$, $X^p \subset L^2(\Omega)$ (respectively, $X^\theta \subset \{v \in H^1(\Omega) \; | v|_{x_1=0} = 0\}$) is a classical $\mathbb{P}_2 - \mathbb{P}_1$ Taylor–Hood Stokes (respectively, $\mathbb{P}_2$ scalar) finite element approximation subspace [5]; X is of dimension $\mathcal{N} = 2869$. We associate to X the inner product and norm

$$(w, v)_X = \left(\int_\Omega \frac{\partial W_i}{\partial x_j} \frac{\partial V_i}{\partial x_j} + W_i V_i + rq + \frac{\partial \chi}{\partial x_i} \frac{\partial \phi}{\partial x_i} + \chi \phi \right) + \kappa \alpha$$

and $\|w\|_X = \sqrt{(w, w)_X}$, respectively, where $w = (W_1, W_2, r, \chi, \kappa)$ and $v = (V_1, V_2, q, \phi, \alpha)$.

The strong (or distributional) form of the governing equations is then

$$\sqrt{\mathrm{Gr}}\, u_j \frac{\partial u_i}{\partial x_j} = -\sqrt{\mathrm{Gr}} \frac{\partial p}{\partial x_i} + \sqrt{\mathrm{Gr}}\, \theta \delta_{i2} + \frac{\partial^2 u_i}{\partial x_j \partial x_j}, \quad i = 1, 2,$$

$$\frac{\partial u_i}{\partial x_i} = \lambda,$$

$$\sqrt{\mathrm{Gr}}\, \mathrm{Pr}\, u_j \frac{\partial \theta}{\partial x_j} = \frac{\partial^2 \theta}{\partial x_j \partial x_j},$$

with boundary–normalization conditions $u|_{\partial\Omega} = 0$ on the velocity, $\int_\Omega p = 0$ on the pressure, and $\partial\theta/\partial n|_{\Gamma_1} = 1$, $\theta|_{\Gamma_0} = 0$, $\partial\theta/\partial n|_{\Gamma_s} = 0$ on the temperature; the flow is thus driven by the flux imposed on Γ_1. Here δ_{ij} is the Kronecker-delta, $\partial\Omega$ is the boundary of Ω, and $\Gamma_0 = \{x_1 = 0, x_2 \in [0, 1]\}$ (left side), $\Gamma_1 = \{x_1 = 1, x_2 \in [0, 1]\}$ (right side), and $\Gamma_s = \{x_1 \in]0, 1[\,, x_2 = 0\} \cup \{x_1 \in]0, 1[\,, x_2 = 1\}$ (top and bottom). It is readily derived that $\lambda = 0$; however, we retain this term as a computationally convenient and stable fashion by which to impose the zero-mean pressure condition on the truth finite element solution. Our output of interest is the average temperature over Γ_1: $s(\mathrm{Gr}) = \ell(u(\mathrm{Gr}))$, where

$$\ell(v = (V_1, V_2, q, \phi, \alpha)) \equiv \int_{\Gamma_1} \phi; \tag{41}$$

note that $s^{-1}(\mathrm{Gr})$ is the traditional "Nusselt number".

The weak form of our partial differential equations is then given by (5), where

$$g(w, v; \mathrm{Gr}) \equiv a_0(w, v; \mathrm{Gr}) + \tfrac{1}{2} a_1(w, w, v; \mathrm{Gr}) - f(v), \tag{42}$$

$$a_0(w^1, v; \mathrm{Gr}) \equiv \left[\int_\Omega \frac{\partial W_i^1}{\partial x_j} \frac{\partial V_i}{\partial x_j} - \int_\Omega \frac{\partial W_i^1}{\partial x_i} q + \kappa^1 \int_\Omega q + \alpha \int_\Omega r^1 + \int_\Omega \frac{\partial \chi^1}{\partial x_i} \frac{\partial \phi}{\partial x_i} \right] + \sqrt{\mathrm{Gr}} \left[- \int_\Omega \chi^1 V_2 - \int_\Omega r^1 \frac{\partial V_i}{\partial x_i} \right], \tag{43}$$

$$a_1(w^1, w^2, v; \mathrm{Gr}) \equiv \sqrt{\mathrm{Gr}} \left[- \int_\Omega \left(W_j^1 W_i^2 + W_j^2 W_i^1 \right) \frac{\partial V_i}{\partial x_j} + \mathrm{Pr} \int_\Omega \left(W_j^2 \frac{\partial \chi^1}{\partial x_j} + W_j^1 \frac{\partial \chi^2}{\partial x_j} \right) \phi \right], \tag{44}$$

$$f(v) \equiv \int_{\Gamma_1} \phi; \tag{45}$$

here $w^1 = (W_1^1, W_2^1, r^1, \chi^1, \kappa^1)$, $w^2 = (W_1^2, W_2^2, r^2, \chi^2, \kappa^2)$, and $v = (V_1, V_2, q, \phi, \alpha)$. Note that, even though $\ell = f$, we are not in "compliance" as g is not bilinear, symmetric; however, we are "close" to compliance, and thus might anticipate rapid output convergence.

We next observe that $a_0(w^1, v; \mathrm{Gr})$ and $a_1(w^1, w^2, v; \mathrm{Gr})$ satisfy (a nonlinear version of) our assumption of affine parameter dependence (7). In particular, we may write

$$a_0(w^1, v; \mathrm{Gr}) = \sum_{q=1}^{Q_0} \Theta_0^q(\mathrm{Gr}) a_0^q(w^1, v), \tag{46}$$

$$a_1(w^1, w^2, v; \mathrm{Gr}) = \sum_{q=1}^{Q_1} \Theta_1^q(\mathrm{Gr}) a_1^q(w^1, w^2, v), \tag{47}$$

for $Q_0 = 2$ and $Q_1 = 1$. In particular, $\Theta_0^1(\mathrm{Gr}) = 1$, $\Theta_0^2(\mathrm{Gr}) = \sqrt{\mathrm{Gr}}$, and $\Theta_1^1(\mathrm{Gr}) = \sqrt{\mathrm{Gr}}$; the corresponding parameter-independent bilinear and trilinear forms should be clear from (43) and (44). We shall exploit (46) and (47) in our offline–online decomposition.

We define the derivative (about $z \in X$) bilinear form $\mathrm{d}g(\cdot, \cdot; z; \mathrm{Gr}) : X \times X \to \mathbb{R}$ as

$$\mathrm{d}g(w, v; z; \mathrm{Gr}) \equiv a_0(w, v; \mathrm{Gr}) + a_1(w, z, v; \mathrm{Gr})$$

which clearly inherits the affine structure (46) and (47) of g; we note that, for our simple quadratic nonlinearity, $g(z+w, v; \mathrm{Gr}) = g(z, v; \mathrm{Gr}) + \mathrm{d}g(w, v; z; \mathrm{Gr}) + (1/2)\, a_1(w, w, v; \mathrm{Gr})$. We then associate to $\mathrm{d}g(\cdot, \cdot; z; \mathrm{Gr})$ our Babuška inf–sup and continuity "constants"

$$\beta(z; \mathrm{Gr}) \equiv \inf_{w \in X} \sup_{v \in X} \frac{\mathrm{d}g(w, v; z; \mathrm{Gr})}{\|w\|_X \|v\|_X},$$

$$\gamma(z; \mathrm{Gr}) \equiv \sup_{w \in X} \sup_{v \in X} \frac{\mathrm{d}g(w, v; z; \mathrm{Gr})}{\|w\|_X \|v\|_X},$$

respectively; these constants now depend on the state z about which we linearize. We shall confirm *a posteriori* that a solution to our problem does indeed exist for all Gr in the chosen $\mathcal{D}$; we can further demonstrate [22] that the manifold $\{u(\mathrm{Gr}) | \mathrm{Gr} \in \mathcal{D}\}$ upon which we focus is a nonsingular (isolated) solution branch, and thus $\beta(u(\mathrm{Gr})) \geq \beta_0 > 0$, $\forall\ \mathrm{Gr} \in \mathcal{D}$.* We can also verify $\gamma(z; \mathrm{Gr}) \leq 2\sqrt{\mathrm{Gr}}\,(1 + \rho_U\,(\rho_U + \mathrm{Pr}\rho_\theta)\,\|z\|_X)$, where

$$\rho_U \equiv \sup_{V \in X^U} \frac{\|V\|_{L^4(\Omega)}}{\|V\|_{X^U}}, \quad \rho_\theta \equiv \sup_{\phi \in X^\theta} \frac{\|\phi\|_{L^4(\Omega)}}{\|\phi\|_{H^1(\Omega)}} \tag{48}$$

are Sobolev embedding constants [27, 28]; for $V \in X^U$, $\|V\|_{L^n(\Omega)} \equiv (\int_\Omega (V_i V_i)^{n/2})^{1/n}$, $1 \leq n < \infty$, $(W, V)_{X^U} \equiv \int_\Omega (\partial W_i/\partial x_j)(\partial V_i/\partial x_j) + W_i V_i$, and $\|V\|_{X^U} \equiv (V, V)_{X^U}^{1/2}$.

We present in Fig. 3(a) a plot of $s(\mathrm{Gr})$; as expected, for low Gr we obtain the conduction solution, $s(\mathrm{Gr}) = 1$; at higher Gr, the larger buoyancy terms create more vigorous flows and hence more effective heat transfer. We show in Fig. 3(b) the velocity and temperature distribution at $\mathrm{Gr} = 10^4$; we observe the familiar "S"-shaped natural convection profile.

5.2. Reduced-Basis Approximation

For simplicity of exposition we shall not address here the adjoint in the nonlinear (approximation or error estimation) context [22], and we shall thus only consider RB treatment of the *primal problem*, (5) and (42). Our RB (Galerkin)

*We note that our truth approximation is div-stable in the sense that the "Brezzi" inf–sup parameter, β^{Br}, is bounded from below (independent of $\mathcal{N}$):

$$\beta^{\mathrm{Br}} \equiv \inf_{\{q \in X^p | \int_\Omega q = 0\}} \sup_{V \in X^U} \frac{\int_\Omega q(\partial V_i/\partial x_i)}{\|V\|_{X^U} \|q\|_{L^2(\Omega)}} > 0;$$

this is a necessary condition for "Babuška" inf–sup stability of the linearized operator $dg(\cdot, \cdot, z; \mathrm{Gr})$.

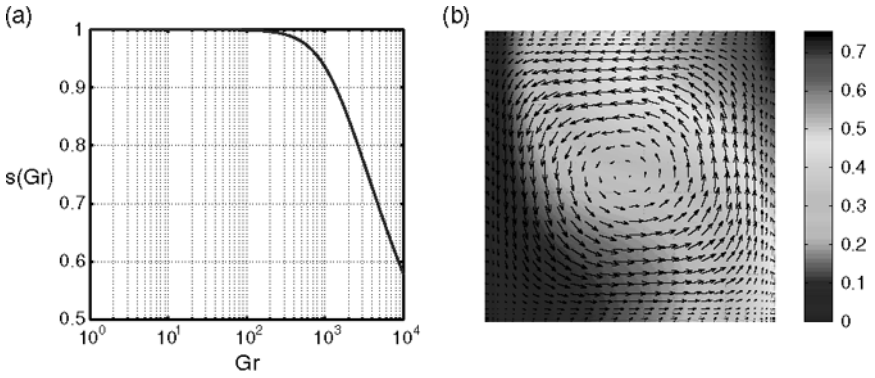

Figure 3. (a) Inverse Nusselt number $s(\mathrm{Gr})$ as a function of Gr; and (b) velocity and temperature field for $\mathrm{Gr} = 10^4$.

approximation is thus: for given $\mathrm{Gr} \in \mathcal{D}$, evaluate $s_N(\mathrm{Gr}) = \ell(u_N(\mathrm{Gr}))$, where $u_N(\mathrm{Gr}) \equiv (U_N, p_N, \theta_N, \lambda_N)(\mathrm{Gr}) \in W_N \equiv W_N^U \times W_N^p \times W_N^\theta \times W_N^\lambda$ satisfies

$$g(u_N(\mathrm{Gr}), v; \mathrm{Gr}) = 0, \quad \forall\, v \in W_N,$$

for ℓ and g defined in (41) and (42)–(45). There are two new ingredients: correct choice of W_N to ensure div-stability; and efficient offline–online treatment of the nonlinearity.

We first address W_N. To begin, we assume that $N = 4m$ for m a positive integer, and we introduce a sequence of nested parameter samples $S_N \equiv \{\mu_1^{\mathrm{pr}} \in \mathcal{D}, \ldots, \mu_{N/4}^{\mathrm{pr}} \in \mathcal{D}\}$ in terms of which we may then define the components of W_N. It is simplest to start with $W^p \equiv \mathrm{span}\{p(\mu_n), 1 \le n \le N/4, \text{ and } \bar{p}\}$, where $\bar{p} = 1$ is the constant function; we then choose $W_N^U \equiv \mathrm{span}\{U(\mu_n^{\mathrm{pr}}), \mathcal{S}p(\mu_n^{\mathrm{pr}}), 1 \le n \le N/4\}$, where for $q \in L^2(\Omega)$, $\mathcal{S}q \in X^U$ satisfies

$$(\mathcal{S}q, V)_{X^U} = \int_\Omega \frac{\partial V_i}{\partial x_i} q, \quad \forall\, V \in X^U;$$

we next define $W_N^\theta \equiv \mathrm{span}\{\theta(\mu_n^{\mathrm{pr}}), 1 \le n \le N/4\}$; and, finally, $W_N^\lambda \equiv \mathbb{R}$. Note that W_N^U must be chosen such that the RB approximation satisfies the Brezzi div-stability condition; for our problem, the domain Ω and hence, the span of the supremizers do not depend on the parameter, and therefore the choice of W_N^U is simple – the more general case is addressed in [29]. We observe that $\dim(W_N^U) = (N/2)$, $\dim(W_N^p) = (N/4) + 1$, $\dim(W_N^\theta) = (N/4)$, and $\dim(W_N^\lambda) = 1$, and hence $\dim(W_N) = N + 2$.*

*In fact, we can explicitly eliminate (the zero coefficient of) $\bar{p}$ and $\lambda_N (= 0)$ from our RB discrete equations, and thus the effective dimension of W_N is N. In the RB context, for which each member $p(\mu_n^{\mathrm{pr}})$ of W_N^p is explicitly zero-mean, the services of the Lagrange multiplier are no longer required.

For our nonlinear problem, the essential computational kernel is the inner Newton update: given a kth iterate $u_N^k(\mathrm{Gr})$, the Newton increment $\delta u_N^k(\mathrm{Gr})$ satisfies $\mathrm{d}g(\delta u_N^k(\mathrm{Gr}), v; u_N^k(\mathrm{Gr}); \mathrm{Gr}) = -g(u_N^k(\mathrm{Gr}), v; \mathrm{Gr}), \forall v \in X$. If we now expand $u_N^k(\mathrm{Gr}) = \sum_{n=1}^N u_{N\,n}^k(\mathrm{Gr})\,\zeta_n$ – where $W_N = \mathrm{span}\{\zeta_n, 1 \le n \le N\}$ – and $\delta u_N^k(\mathrm{Gr}) = \sum_{j=1}^N \delta u_{N\,j}^k(\mathrm{Gr})\,\zeta_j$, we obtain [17] the linear set of equations

$$\sum_{j=1}^{N} \left\{ \sum_{q=1}^{Q_0} \Theta_0^q(\mathrm{Gr})\, a_0^q(\zeta_j, \zeta_i) + \sum_{n=1}^{N} \sum_{q'=1}^{Q_1} \Theta_1^{q'}(\mathrm{Gr}) u_{Nn}^k(\mathrm{Gr}) a_1^{q'}(\zeta_j, \zeta_n, \zeta_i) \right\} \delta_{Nj}^k(\mathrm{Gr}) = -\, g(u_N^k(\mathrm{Gr}), \zeta_i; \mathrm{Gr}), \quad 1 \le i \le N,$$

where (from (42))

$$-g(u_N^k(\mathrm{Gr}), \zeta_i; \mathrm{Gr}) = f(\zeta_i) - \sum_{j=1}^{N} \left\{ \sum_{q=1}^{Q_0} \Theta_0^q(\mathrm{Gr})\, a_0^q(\zeta_j, \zeta_i) + \frac{1}{2} \sum_{n=1}^{N} \sum_{q=1}^{Q_1} u_{N\,n}^k(\mathrm{Gr}) \Theta_1^q(\mathrm{Gr}) a_1^q(\zeta_j, \zeta_n, \zeta_i) \right\} u_{Nj}^k(\mathrm{Gr})$$

is the residual for $v = \zeta_i$.

We can now directly apply the offline–online procedure [7–9] described earlier for linear problems, except now we must perform summations both "over the affine parameter dependence" *and* "over the reduced-basis coefficients" (of the current Newton iterate about which we linearize).* The operation count for the predominant Newton update component of the online stage is then – per Newton iteration – $O(N^3)$ to assemble the residual, $-g(u_N^k(\mathrm{Gr}), \zeta_i; \mathrm{Gr})$, $1 \le i \le N$, and $O(N^3)$ to assemble and invert the $N \times N$ Jacobian. The essential point is that the online complexity is independent of $\mathcal{N}$, thanks to offline generation and storage of the requisite parameter independent quantities (for example, $a_1^q(\zeta_j, \zeta_n, \zeta_i)$).

For this particular nonlinear problem, there is relatively little additional cost associated with the nonlinearity. However, our success depends crucially on the low-order polynomial nature of our nonlinearity: in general, standard Galerkin procedures will yield N^{n+1} complexity for an nth order ($n \ge 2$) polynomial nonlinearity. Although symmetries can be invoked to modestly improve the scaling with N and n [18], in any event new approaches will be

*In essence – we shall see this again in the error estimation context – our quadratic nonlinearity effectively introduces N additional "parameter-dependent functions" and "parameter-independent forms" associated with the coefficients of our field-variable expansion and our trilinear form, respectively; however, these new parameter contributions are correlated in ways that we can gainfully exploit.

required for nonpolynomial nonlinearities; we discuss these new procedures for efficient treatment of general nonaffine and nonlinear operators in Outlook.

5.3. *A Posteriori* Error Estimation

The motivation for rigorous *a posteriori* error estimation is even more self-evident in the case of nonlinear problems. Fortunately, there is a rich mathematical foundation upon which to build the necessary computational structure. We first introduce the former; we then describe the latter. For simplicity, we develop here error bounds only for the primal energy norm, $\|u(\mu)-u_N(\mu)\|_X$; we can also develop error bounds for the output – however, good effectivities will require consideration of the dual [22].

5.3.1. Error bounds

We require some slight modifications to our earlier (linear) preliminaries. In particular, we introduce $T_N^\mu : X \to X$ such that, for any $w \in X$, $(T_N^\mu w, v)_X = \mathrm{d}g(w, v; u_N(\mu); \mu), \forall v \in X$; we then define $\sigma_N(w; \mu) \equiv \|T_N^\mu w\|_X / \|w\|_X$. Our inf–sup and continuity constants – now linearized about the reduced-basis solution – can then be expressed as $\beta_N(\mu) \equiv \beta(u_N(\mu); \mu) = \inf_{w \in X} \sigma_N(w; \mu)$, and $\gamma_N(\mu) \equiv \gamma(u_N(\mu); \mu) = \sup_{w \in X} \sigma_N(w; \mu)$, respectively; as before, we shall need a nonnegative lower bound for the inf–sup parameter, $\tilde{\beta}_N(\mu)$, such that $\beta_N(\mu) \geq \tilde{\beta}_N(\mu) \geq 0, \forall\ \mu \in \mathcal{D}$.

As in the linear case, the dual norm of the residual, $\varepsilon_N(\mu)$ of (22), shall play a central role; the (negative of the) residual for our current nonlinear problem is given by (42) for $w = u_N(\mu)$. We also introduce a new combination of parameters $\tau_N(\mu) \equiv 2\rho(\mu)\varepsilon_N(\mu)/\tilde{\beta}_N^2(\mu)$, where $\rho(\mu) = 2\sqrt{\mathrm{Gr}}\rho_U(\rho_U + \mathrm{Pr}\rho_\theta)$ depends on the Sobolev embedding constants ρ_U and ρ_θ of (48); in essense, $\tau_N(\mu)$ is an appropriately "nondimensionalized" measure of the residual. Finally, we define $N^*(\mu)$ such that $\tau_N(\mu) < 1$ for $N \geq N^*(\mu)$; we require $N^*(\mu) \leq N_{\max}, \forall\ \mu \in \mathcal{D}$. (The latter is a condition on $N_{\max}$ that reflects both the convergence rate of the RB approximation and the quality of our inf–sup lower bound.) We recall that $\mu \equiv \mathrm{Gr} \in \mathcal{D} \equiv [1.0, 1.0 \times 10^4]$.

Our error bound is then expressed, for any $\mu \in \mathcal{D}$ and $N \geq N^*(\mu)$, as

$$\Delta_N(\mu) = \frac{\tilde{\beta}_N(\mu)}{\rho(\mu)} \left(1 - \sqrt{1 - \tau_N(\mu)}\right). \tag{49}$$

The main result can be very simply stated: if $N \geq N^*(\mu)$, there exists a unique solution $u(\mu)$ to (5) in the open ball

$$\mathcal{B}\left(u_N(\mu), \frac{\tilde{\beta}_N(\mu)}{\rho(\mu)}\right) \equiv \left\{ z \in X \;\middle|\; \|z - u_N(\mu)\|_X < \frac{\tilde{\beta}_N(\mu)}{\rho(\mu)} \right\}; \tag{50}$$

furthermore,

$$\|u(\mu) - u_N(\mu)\|_X \leq \Delta_N(\mu). \tag{51}$$

The proof, given in Ref. [22], is a slight specialization of a general abstract result [30, 31] that in turn derives from the Brezzi–Rappaz–Raviart (BRR) framework for the analysis of variational approximations of nonlinear partial differential equations [32]; the central ingredient is the construction of an appropriate contraction mapping which then forms the foundation for a standard fixed-point argument. On the basis of the main proposition (50) and (51) we can further prove several important corollaries related to the well-posedness of the truth approximation (5), and – similar to the linear result (24) – the effectivity of our error bound (49) [22].

We note that, as $\varepsilon_N(\mu) \to 0$, we shall certainly satisfy $N \geq N^*(\mu)$; furthermore the upper bound to the true error, $\Delta_N(\mu)$ of (49), is asymptotic to $\varepsilon_N(\mu)/\tilde{\beta}_N(\mu)$. We may derive these limits directly and rigorously from (49) and (51), or more heuristically from the equation for the error $e(\mu) \equiv u(\mu) - u_N(\mu)$,

$$\mathrm{d}g(e(\mu), v; u_N(\mu); \mu) = -g(u_N(\mu), v; \mu) - \tfrac{1}{2}a_1(e(\mu), e(\mu), v; \mu). \tag{52}$$

We conclude that the nonlinear case shares much in common with the limiting linear case. However, there are also important differences: even for $\tau_N(\mu) < 1$, we must (in general) admit the possibility of other solutions to (5) – solutions *outside* $\mathcal{B}(u_N(\mu), \tilde{\beta}_N/\rho(\mu))$ – that are *not* near $u_N(\mu)$; and for $\tau_N(\mu) \geq 1$, we cannot even be assured that there is indeed any solution $u(\mu)$ near $u_N(\mu)$. This conclusion is not surprising: for "noncoercive" nonlinear problems the error equation (51) may in general admit no or several solutions; we can only be certain that a small (isolated) solution exists, (50) and (51), if the residual is sufficiently small. The theory informs us that the appropriate measure of the residual is $\tau_N(\mu)$, which reflects both the stability of the operator ($\tilde{\beta}_N(\mu)$) and the strength of the nonlinearity ($\rho(\mu)$).

As in the linear case, the real computational challenge is the development of efficient procedures for the calculation of the necessary *a posteriori* quantities:* the dual norm of the residual, $\varepsilon_N(\mu)$; the inf–sup lower bound, $\tilde{\beta}_N(\mu)$; and – new to our nonlinear problem – the Sobolev constants, ρ_U and ρ_θ. We now turn to these considerations.

*Typically, the BRR framework provides a *nonquantitative a priori* or *a posteriori* justification of asymptotic convergence. In our context, there is a unique opportunity to render the BRR theory completely predictive: actual *a posteriori* error estimators that are quantitative, rigorous, sharp, and (online) inexpensive.

5.3.2. Offline-online computational procedures

The dual norm of the residual. Fortunately, the duality relation of the linear case, (29), still applies – $g(w, v; \mu)$ of (42) is nonlinear in w, but of course linear in v. For our nonlinear problem, the negative of the residual, (42), for $w = u_N(\mu)$, may be expressed in terms of the reduced-basis expansion (12) as

$$-g(u_N(\mu), v; \mu) = f(v) - \sum_{n=1}^{N} u_{N\,n}(\mu) \left\{ \sum_{q=1}^{Q_0} \Theta_0^q(\mu) a_0^q(\zeta_n, v) + \frac{1}{2} \sum_{q'=1}^{Q_1} \sum_{n'=1}^{N} \Theta_1^{q'}(\mu)\, u_{N\,n'}(\mu) a_1^{q'}(\zeta_n, \zeta_{n'}, v) \right\}, \tag{53}$$

where we recall that $\mu \equiv \mathrm{Gr}$. If we insert (53) in (29) and apply linear superposition, we obtain

$$\hat{e}(\mu) = \mathcal{C} + \sum_{n=1}^{N} u_{N\,n}(\mu) \left\{ \sum_{q=1}^{Q_0} \Theta_0^q(\mu) \mathcal{L}_n^q + \sum_{q'=1}^{Q_1} \sum_{n'=1}^{N} \Theta_1^{q'}(\mu) u_{N\,n'}(\mu) \mathcal{Q}_{n\,n'}^{q'} \right\},$$

where $\mathcal{C} \in X$ satisfies $(\mathcal{C}, v)_X = f(v)$, $\forall\ v \in X$, $\mathcal{L}_n^q \in X$ satisfies $(\mathcal{L}_n^q, v)_X = -a_0^q(\zeta_n, v)$, $\forall\ v \in X$, $1 \le n \le N$, $1 \le q \le Q_0$, and $\mathcal{Q}_{n\,n'}^q \in X$ satisfies $\mathcal{Q}_{n\,n'}^q = -a_1^q(\zeta_n, \zeta_{n'}, v)/2$, $\forall\ v \in X$, $1 \le n, n' \le N$, $1 \le q \le Q_1$; the latter are again simple (vector) Poisson problems. It thus follows that [22]

$$\begin{aligned} \|\hat{e}(\mu)\|_X^2 = (\mathcal{C}, \mathcal{C})_X + \sum_{n=1}^{N} u_{N\,n}(\mu) \Bigg\{ & 2 \sum_{q=1}^{Q_0} \Theta_0^q(\mu) (\mathcal{C}, \mathcal{L}_n^q)_X + \sum_{n'=1}^{N} u_{N\,n'}(\mu) \\ & \times \Bigg\{ 2 \sum_{q=1}^{Q_1} \Theta_1^q(\mu) (\mathcal{C}, \mathcal{Q}_{n\,n'}^q)_X + \sum_{q=1}^{Q_0} \sum_{q'=1}^{Q_0} \Theta_0^q(\mu) \Theta_0^{q'}(\mu) (\mathcal{L}_n^q, \mathcal{L}_{n'}^{q'})_X \\ & + \sum_{n''=1}^{N} u_{N\,n''}(\mu) \Bigg\{ 2 \sum_{q=1}^{Q_0} \sum_{q'=1}^{Q_1} \Theta_0^q(\mu) \Theta_1^{q'}(\mu) (\mathcal{L}_n^q, \mathcal{Q}_{n'\,n''}^{q'})_X \\ & + \sum_{n'''=1}^{N} u_{N\,n'''}(\mu) \sum_{q=1}^{Q_1} \sum_{q'=1}^{Q_1} \Theta_1^q(\mu) \Theta_1^{q'}(\mu) (\mathcal{Q}_{n\,n'}^q, \mathcal{Q}_{n''n'''}^{q'})_X \Bigg\} \Bigg\} \Bigg\} \end{aligned}$$

from which we can directly calculate the requisite dual norm of the residual through (28).

We can now readily adapt the offline–online procedure developed in the linear case; however, our summation "over the affine dependence" now involves a *double* summation "over the reduced-basis coefficients". The operation count for the online stage is thus (to leading order) $O(Q_1^2 N^4)$; the essential point is that

the online complexity is again independent of $\mathcal{N}$ – thanks to offline generation and storage of the requisite parameter-independent inner products (for example, $(\mathcal{Q}^q_{n\,n'}, \mathcal{Q}^{q'}_{n''\,n'''})_X$, $1 \le n, n', n'', n''' \le N$, $1 \le q, q' \le Q_1$). Although the N^4 online scaling is certainly less than pleasant, the error bound is calculated only once – at the termination of the Newton iteration – and hence in actual practice the additional online cost attributable to the residual dual norm computation is in fact not too large. However, the quartic scaling with N is again a *memento mori* that, for higher order (than quadratic) nonlinearities, standard Galerkin procedures are not viable; we discuss the alternatives further in Outlook. □

Lower bound for the inf–sup parameter. Our procedure for the linear case can be readily adopted: we need "only" incorporate the N additional parameter-dependent "coefficient functions" – in fact, the RB coefficients – that appear in the linearized-about-$u_N(\mu)$ derivative operator. Hence, for our nonlinear problem, the bilinear form $\mathcal{T}$ of (34) and Rayleigh quotient $\mathcal{F}$ of (35) now contain sensitivity derivatives of these additional "coefficient functions"; furthermore, the $\Phi(\mu, \bar{\mu})$ function of (37) – our second-order remainder term – now includes the deviation of the RB coefficients from linear parameter dependence. Further details are provided in Ref. [22] (for Pr = 0) for the case in which $W_N \equiv W_N^U$ is divergence-free. □

Sobolev continuity constant. We present here the procedure for calculation of ρ_U; the procedure for ρ_θ is similar. We first note [27, 28] that $\rho_U = (1/\hat{\delta}_{\min})^{1/2}$, where $(\hat{\delta}, \hat{\xi}) \in (\mathbb{R}_+, X^U)$ satisfies

$$(\hat{\xi}, V)_{X^U} = \hat{\delta} \int_\Omega \hat{\xi}_j \hat{\xi}_j \hat{\xi}_i V_i, \quad \forall V \in X^U, \quad \|\hat{\xi}\|^4_{L^4(\Omega)} = 1,$$

and $(\hat{\delta}_{\min}, \hat{\xi}_{\min})$ denotes the ground state. To solve this eigenproblem, and in particular to ensure that we realize the ground state, we pursue a homotopy procedure.

Towards that end, we introduce a parameter $h \in [0, 1]$ (and associated small increment Δh) and look for $(\delta(h), \xi(h)) \in (\mathbb{R}_+, X^U)$ that satisfies

$$\begin{aligned}(\xi(h), V)_{X^U} = \delta(h) \Bigg(& h \int_\Omega \xi_j(h)\xi_j(h)\xi_i(h) V_i \\ & + (1-h) \int_\Omega \xi_i(h) V_i \Bigg), \forall V \in X^U, \\ & h\|\xi\|^4_{L^4(\Omega)} + (1-h)\|\xi\|^2_{L^2(\Omega)} = 1; \end{aligned} \tag{54}$$

$(\delta_{\min}(h), \xi_{\min}(h))$ denotes the ground state. We observe that $(\delta_{\min}(1), \xi_{\min}(1)) = (\hat{\delta}_{\min}, \hat{\xi}_{\min})$; and that $(\delta_{\min}(0), \xi_{\min}(0))$ is the lowest eigenpair of the standard

(vector) Laplacian "linear" eigenproblem. Our homotopy procedure is simple: we first set $h^{\text{old}} = 0$ and find $(\delta_{\min}(0), \xi_{\min}(0))$ by standard techniques; then, until $h^{\text{new}} = 1$, we set $h^{\text{new}} \leftarrow h^{\text{old}} + \Delta h$, solve (54) for $(\delta_{\min}(h^{\text{new}}), \xi_{\min}(h^{\text{new}}))$ by Newton iteration initialized to $(\delta_{\min}(h^{\text{old}}), \xi_{\min}(h^{\text{old}}))$, and update $h^{\text{old}} \leftarrow h^{\text{new}}$. For our domain, we find (offline) $\rho_U = 0.6008$, $\rho_\theta = 0.2788$; since ρ_U and ρ_θ are parameter-independent, no online computation is required. □

5.3.3. Sample construction

The greedy algorithm developed in the linear case requires some modification in the nonlinear context. The first issue is that, to evaluate our error bound $\Delta_N(\mu)$, we must appeal to our inf–sup lower bound; however, in the nonlinear case, this inf–sup lower bound, $\tilde{\beta}_N(\mu)$, is defined with respect to the linearized state $u_{N_{\max}}(\mu)$ [22]. In short, to determine the "next" sample point μ_{N+1} we must already know $S_{N_{\max}}$ – and hence μ_{N+1}. To avoid this circular reference during the offline sample generation process, we replace our inf–sup lower bound with a crude (for example, piecewise constant over $\mathcal{D}$) approximation to $\beta(u(\mu))$; once the samples are constructed, we revert to our rigorous (and now calculable) lower bound, $\tilde{\beta}_N(\mu)$. The second issue is that, in the nonlinear context, our error bound is not operative until $\tau_N(\mu) < 1$; hence, the greedy procedure must first select on $\arg\max_{\mu \in \Xi_F} \tau_N(\mu)$ – until $\tau_N(\mu) < 1$ over $\mathcal{D}$ – and only subsequently select on $\arg\max_{\mu \in \Xi_F} \Delta_N(\mu)$ [Prud'homme, private communication]. The resulting sample will ensure not only rapid convergence to the exact solution, but also rapid convergence to a *certifiably* accurate solution.

5.4. Numerical Results

We present in Table 2 $\|u(\tilde{\mu}_N) - u_N(\tilde{\mu}_N)\|_X / \|u(\tilde{\mu}_N)\|_X$, $\Delta_{N,\text{rel}}(\tilde{\mu}_N) \equiv \Delta_N(\tilde{\mu}_N)/\|u_N(\tilde{\mu}_N)\|_X$, and $\eta_N(\tilde{\mu}_N) \equiv \Delta_N(\tilde{\mu}_N)/\|e(\tilde{\mu}_N)\|_X$ for $8 \le N \le N_{\max} = 40$; here

$$\tilde{\mu}_N \equiv \arg \max_{\mu \in \Xi_{\text{Test}}} \frac{\|u(\mu) - u_N(\mu)\|_X}{\|u(\mu)\|_X}$$

and Ξ_{Test} is a random parameter grid of size $n_{\text{Test}} = 500$.

We observe very rapid convergence of $u_N(\mu)$ to $u(\mu)$ over $\mathcal{D}$ (more precisely, Ξ_{Test}) – our samples S_N are optimally constructed to provide uniform convergence. The output error decreases even more rapidly: $\max_{\mu \in \Xi_{\text{Test}}} |s(\mu) - s_N(\mu)|/s(\mu) = 1.34 \times 10^{-1}$, 2.80×10^{-4}, and 9.79×10^{-7} for $N = 8$, 16, and 24, respectively; this "superconvergence" is a vestige of near compliance. As regards *a posteriori* error estimation, we observe that $N^*(\tilde{\mu}_N) = 24$

Table 2. Convergence and effectivity results for the natural convection problem; the "*" signifies that $N^*(\tilde{\mu}_N) > N$, which in turn indicate that $\tau_N(\tilde{\mu}_N) \geq 1$

N	$\frac{\|u(\tilde{\mu}_N) - u_N(\tilde{\mu}_N)\|_X}{\|u(\tilde{\mu}_N)\|_X}$	$\Delta_{N,\text{rel}}(\tilde{\mu}_N)$	$\eta_N(\tilde{\mu}_N)$
8	3.28×10^{-1}	*	*
16	1.45×10^{-2}	*	*
24	1.80×10^{-4}	7.47×10^{-4}	4.15
32	8.05×10^{-7}	7.60×10^{-6}	9.44
40	4.60×10^{-8}	8.69×10^{-7}	18.93

is relatively small – we can (respectively, can not) provide a definitive error bound for $N \geq 24$ (respectively, $N < 24$); more generally, we find that $N^*(\mu) \leq 24$, $\forall\ \mu \in \mathcal{D}$. We note that the effectivities are quite good* – in fact, considerably better than the worst-case predictions of our effectivity corollary. (The higher effectivity at $N = 40$ is undoubtedly due to round-off in the online summation.)

The results of Table 2 are based on an inf–sup lower bound construction with $J = 28$ elements: points $\bar{\mu}_j$ and polytopes (here segments) $\mathcal{P}^{\bar{\mu}_j}$, $1 \leq j \leq J$. The accuracy of the resulting lower bound is reflected in the modest $N^*(\mu)$ and the good effectivities reported in Table 2. Most of the points $\bar{\mu}_j$ are clustered at larger Gr, as might be expected.

Finally, we note that the total *online* computational time on a Pentium M 1.6 GHz processor to predict $u_N(\text{Gr})$, $s_N(\text{Gr})$, and $\Delta_N(\text{Gr})$ to a relative accuracy (in the energy norm) of 10^{-3} is – $\forall\ \text{Gr} \in \mathcal{D}$ – 300 ms; this should be compared to 50 s for direct finite element calculation of the truth solution, $u(\text{Gr})$, $s(\text{Gr})$. We achieve computational savings of $O(100)$: N is very small thanks to (*i*) the good convergence properties of S_N and hence W_N, and (*ii*) the rigorous and sharp stopping criterion provided by $\Delta_N(\text{Gr})$; and the marginal computational complexity to evaluate $s_N(\text{Gr})$ and $\Delta_N(\text{Gr})$ depends only on N and *not* on $\mathcal{N}$ – thanks to the offline–online decomposition. The computational savings will be even more significant for more complex problems particularly in three spatial dimensions; it is critical to recall that we realize these savings *without* compromising rigorous certainty.†

*It is perhaps surprising that the BRR theory – not really designed for quantitative service – yields such sharp results. However, it is important to note that, as $\varepsilon_N(\mu) \to 0$, $\Delta_N(\mu) \sim \varepsilon_N(\mu)/\tilde{\beta}_N(\mu)$, and thus the more pessimistic bounds (in particular ρ) are absent – except in $\tau_N(\mu)$.

†We admit that the extension of our results to much larger Gr is not without difficulty. The more complex flow structures and the stronger nonlinearity will degrade the convergence rate and *a posteriori* error bounds – and increase N and J; and (inevitable) limit points and bifurcations will require special precautions.

6. Outlook

We address here some of the more obvious questions that arise in reviewing the current state of affairs. As a first question: How many parameters P can we consider – for P how large are our techniques still viable? It is undeniably the case that ultimately we should anticipate exponential scaling (of both N and certainly J) as P increases, with a concomitant unacceptable increase certainly in offline but also perhaps in online computational effort. Fortunately, for smaller P, the growth in N is rather modest, as (good) sampling procedures will automatically identify the more interesting regions of parameter space. Unfortunately, the growth in J is more problematic: we shall require more efficient construction and verification procedures for our inf–sup lower bound samples. In any event, *treatment of hundreds (or even many tens) of truly independent parameters by the global methods described in this chapter is clearly not practicable; in such cases, more local approaches must be pursued.*[*]

A second question: How can we efficiently treat problems with non-affine parameter dependence and (more than quadratic) state-space nonlinearity? Both these issues are satisfactorily addressed by a new "empirical interpolation" approach [33]. In this approach, we replace a general nonaffine nonlinear function of the parameter μ, spatial coordinate x, and field variable $u(x;\mu)$, $\mathcal{H}(u;x;\mu)$, by a collateral RB expansion: in particular, we approximate $\mathcal{H}(u_N(x;\mu);x;\mu)$ – as required in our RB projection for $u_N(\mu)$ – by $\mathcal{H}_M(x;\mu) = \sum_{m=1}^{M} d_m(\mu)\xi_m(x)$. The critical ingredients of the approach are (*i*) a "good" collateral RB sample, $S_M^{\mathcal{H}} = \{\mu_1^{\mathcal{H}}, \ldots, \mu_M^{\mathcal{H}}\}$, and approximation space, $\text{span}\{\xi_m = \mathcal{H}(u(\mu_m^{\mathcal{H}});x;\mu_m^{\mathcal{H}}),\ 1 \le m \le M\}$, (*ii*) a stable and *inexpensive* interpolation procedure by which to determine (online) the $d_m(\mu)$, $1 \le m \le M$, and (*iii*) effective *a posteriori* error bounds with which to quantify the effect of the newly introduced truncation. It is perhaps only in the latter that the technique is somewhat disappointing: *the error estimators – though quite sharp and very efficient – are completely (provably) rigorous upper bounds only in certain restricted situations*.

Finally, a third question, again related to generality: What class of PDEs can be treated? In addition to the *elliptic* equations discussed in this paper, *parabolic* equations can also be addressed satisfactorily from both the approximation and error estimation points of view [24, 34, 35]:[†] much of the elliptic technology directly applies, except that time now appears as an additional parameter; this parabolic framework can be viewed as an extension of

[*] We do note that at least some problems with ostensibly many parameters in fact involve *highly coupled or correlated* parameters: certain classes of shape optimization certainly fall into this category. In these situations, global progress can be made.

[†] To date we have experience with only stable parabolic systems such as the heat equation; *unstable systems present considerable difficulty, in particular if long-time solutions are desired*.

time-domain model reduction procedures [19, 25, 36]. Unfortunately, *treatment of hyperbolic problems does not look promising*: although RB methods can perform quite well anecdotally, in general the underlying smoothness (in parameter μ) and stability will no longer obtain; as a result, both the approximation properties and error estimators will suffer.

We close by noting that *the offline aspects of the approaches described are both complicated and computationally expensive*. The former can be at least partially addressed by appropriate software and architectures [37]; however, the latter will in any event remain. It follows that these techniques will really only be viable in situations in which there is truly an imperative for real-time certified response: a real premium on (*i*) greatly reduced marginal cost (or asymptotic average cost), and (*ii*) rigorous characterization of certainty; or equivalently, a very high (opportunity) cost associated with (*i*) slow response – long latency times, and (*ii*) incorrect (or unsafe) decisions or actions. There are many classes of materials and materials processing problems and contexts for which the methods are appropriate; and certainly there are many classes of materials and materials processing problems and contexts for which more classical techniques remain distinctly preferred.

Appendix A Helmholtz Elasticity Example

We first define a reference domain corresponding to the geometry $b = b_r = 1$ and $L = L_r = 0.2$. We then map $\Omega^o(b, L) \rightarrow \Omega \equiv \Omega^o(b_r, L_r)$ by a continuous piecewise-affine (in fact, piecewise-dilation-in-x_1) transformation. We define three subdomains, $\Omega_1 \equiv]\,0, b_r - L_r/2\,[\times]\,0, 1\,[$, $\Omega_2 \equiv]\,b_r - L_r/2, b_r + L_r/2\,[\times]\,0, 1[$, $\Omega_3 \equiv]b_r + L_r/2, 2\,[\times]\,0, 1\,[$, such that $\bar{\Omega} = \bar{\Omega}_1 \cup \bar{\Omega}_2 \cup \bar{\Omega}_3$.

We may then express the resulting bilinear form $a(w, v; \mu)$ as an affine sum (7) for $Q = 10$; the particular $\Theta^q(\mu)$, $a^q(w, v)$, $1 \le q \le 10$, as shown in Table 3. (Recall that $w = (w_1, w_2)$ and $v = (v_1, v_2)$.) The constitutive constants in Table 3 are given by

$$c_{11} = \frac{1}{1-\nu^2}, \qquad c_{22} = c_{11}, \qquad c_{12} = \frac{\nu}{1-\nu^2}, \qquad c_{66} = \frac{1}{2(1+\nu)},$$

where $\nu = 0.25$ is the Poisson ratio (and the normalized Young's modulus is unity); recall that we consider plane stress and a linear isotropic solid.

We now define our inner product-cum-bound conditioner as

$$\begin{aligned}(w, v)_X \equiv & \int_\Omega c_{11} \frac{\partial v_1}{\partial x_1} \frac{\partial w_1}{\partial x_1} + c_{22} \frac{\partial v_2}{\partial x_2} \frac{\partial w_2}{\partial x_2} + c_{66} \frac{\partial v_2}{\partial x_1} \frac{\partial w_2}{\partial x_1} + c_{66} \frac{\partial v_1}{\partial x_2} \frac{\partial w_1}{\partial x_2} \\ & + w_1 v_1 + w_2 v_2 \\ = & \sum_{q=2}^{Q} a^q(w, v)\ ;\end{aligned}$$

Table 3. Parametric functions $\Theta^q(\mu)$ and parameter-independent bilinear forms $a^q(w, v)$ for the two-dimensional crack problem

q	$\Theta^q(\mu)$	$a^q(w, v)$
1	1	$c_{12}\int_{\Omega}\left(\frac{\partial v_1}{\partial x_1}\frac{\partial w_2}{\partial x_2}+\frac{\partial v_2}{\partial x_2}\frac{\partial w_1}{\partial x_1}\right)$ $+c_{66}\int_{\Omega}\left(\frac{\partial v_1}{\partial x_2}\frac{\partial w_2}{\partial x_1}+\frac{\partial v_2}{\partial x_1}\frac{\partial w_1}{\partial x_2}\right)$
2	$\frac{b_r - L_r/2}{b - L/2}$	$c_{11}\int_{\Omega_1}\left(\frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1}\right)+c_{66}\int_{\Omega_1}\left(\frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1}\right)$
3	$\frac{L_r}{L}$	$c_{11}\int_{\Omega_2}\left(\frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1}\right)+c_{66}\int_{\Omega_2}\left(\frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1}\right)$
4	$\frac{2 - b_r - L_r/2}{2 - b - L/2}$	$c_{11}\int_{\Omega_3}\left(\frac{\partial v_1}{\partial x_1}\frac{\partial w_1}{\partial x_1}\right)+c_{66}\int_{\Omega_3}\left(\frac{\partial v_2}{\partial x_1}\frac{\partial w_2}{\partial x_1}\right)$
5	$\frac{b - L/2}{b_r - L_r/2}$	$c_{22}\int_{\Omega_1}\left(\frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2}\right)+c_{66}\int_{\Omega_1}\left(\frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2}\right)$
6	$\frac{L}{L_r}$	$c_{22}\int_{\Omega_2}\left(\frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2}\right)+c_{66}\int_{\Omega_2}\left(\frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2}\right)$
7	$\frac{2 - b - L/2}{2 - b_r - L_r/2}$	$c_{22}\int_{\Omega_3}\left(\frac{\partial v_2}{\partial x_2}\frac{\partial w_2}{\partial x_2}\right)+c_{66}\int_{\Omega_3}\left(\frac{\partial v_1}{\partial x_2}\frac{\partial w_1}{\partial x_2}\right)$
8	$-\omega^2\frac{b - L/2}{b_r - L_r/2}$	$\int_{\Omega_1} w_1 v_1 + w_2 v_2$
9	$-\omega^2\frac{L}{L_r}$	$\int_{\Omega_2} w_1 v_1 + w_2 v_2$
10	$-\omega^2\frac{2 - b - L/2}{2 - b_r - L_r/2}$	$\int_{\Omega_3} w_1 v_1 + w_2 v_2$

thanks to the Dirichlet conditions at $x_1 = 0$ (and also the $w_i v_i$ term), $(\cdot, \cdot)_X$ is appropriately coercive. We now observe that $\Theta(\mu) = 1$ ($\Gamma^1 = 0$) and we can thus disregard the $q = 1$ term in our continuity bounds. We may then choose $|v|_q^2 = a^q(v, v)$, $2 \le q \le Q$, since the $a^q(\cdot, \cdot)$ are positive semi-definite; it thus follows from the Cauchy–Schwarz inequality that $\Gamma^q = 1$, $2 \le q \le Q$; furthermore, from (36), we directly obtain $C_X = 1$.

Acknowledgments

We would like to thank Professor Yvon Maday of University Paris VI for his many invaluable contributions to this work. We would also like to thank

Dr Christophe Prud'homme of EPFL, Mr Martin Grepl of MIT, Mr Gianluigi Rozza of EPFL, and Professor Liu Gui-Rong of NUS for many helpful recommendations. This work was supported by DARPA and AFOSR under Grant F49620-03-1-0356, DARPA/GEAE and AFOSR under Grant F49620-03-1-0439, and the Singapore-MIT Alliance.

References

[1] B.O. Almroth, P. Stern, and F.A. Brogan, "Automatic choice of global shape functions in structural analysis," *AIAA J.*, 16, 525–528, 1978.

[2] A.K. Noor and J.M. Peters, "Reduced basis technique for nonlinear analysis of structures," *AIAA J.*, 18, 455–462, 1980.

[3] J.P. Fink, and W.C. Rheinboldt, "On the error behavior of the reduced basis technique for nonlinear finite element approximations," *Z. Angew. Math. Mech.*, 63, 21–28, 1983.

[4] T.A. Porsching, "Estimation of the error in the reduced basis method solution of nonlinear equations," *Math. Comput.*, 45, 487–496, 1985.

[5] M.D. Gunzburger, *Finite Element Methods for Viscous Incompressible Flows: A Guide to Theory, Practice, and Algorithms*, Academic Press, Boston, 1989.

[6] J.S. Peterson, "The reduced basis method for incompressible viscous flow calculations," *SIAM J. Sci. Stat. Comput.*, 10, 777–786, 1989.

[7] K. Ito and S.S. Ravindran, "A reduced-order method for simulation and control of fluid flows," *Journal of Computational Physics*, 143, 403–425, 1998.

[8] L. Machiels, Y. Maday, I.B. Oliveira, A.T. Patera, and D. Rovas, "Output bounds for reduced-basis approximations of symmetric positive definite eigenvalue problems," *C. R. Acad. Sci. Paris, Série I*, 331, 153–158, 2000.

[9] C. Prud'homme, D. Rovas, K. Veroy, Y. Maday, A.T. Patera, and G. Turinici, "Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods," *J. Fluids Eng.*, 124, 70–80, 2002.

[10] Y. Maday, A.T. Patera, and G. Turinici, "Global *a priori* convergence theory for reduced-basis approximation of single-parameter symmetric coercive elliptic partial differential equations," *C. R. Acad. Sci. Paris, Série I*, 335, 289–294, 2002.

[11] E. Balmes, "Parametric families of reduced finite element models: Theory and applications," *Mech. Syst. Signal Process.*, 10, 381–394, 1996.

[12] Y. Maday, A.T. Patera, and D.V. Rovas, "A blackbox reduced-basis output bound method for noncoercive linear problems," In: D. Cioranescu and J. Lions (eds.), *Nonlinear Partial Differential Equations and Their Applications, Collége de France Seminar Volume XIV*, Elsevier Science B.V, pp. 533–569, 2002.

[13] R. Becker and R. Rannacher, "Weighted *a posteriori* error control in finite element methods," *ENUMATH 95 Proceedings* World Science Publications, Singapore, 1997.

[14] M. Paraschivoiu and A.T. Patera, "A hierarchical duality approach to bounds for the outputs of partial differential equations," *Comp. Meth. Appl. Mech. Eng.*, 158, 389–407, 1998.

[15] M. Ainsworth and J.T. Oden, *A Posteriori Error Estimation in Finite Element Analysis. Pure and Applied Mathematics.*, Wiley-Interscience, New York, 2000.

[16] J.W. Demmel, *Applied Numerical Linear Algebra*, SIAM, Philadelphia, 1997.

[17] K. Veroy, C. Prud'homme, and A.T. Patera, "Reduced-basis approximation of the viscous Burgers equation: Rigorous *a posteriori* error bounds," *C. R. Acad. Sci. Paris, Série I*, 337, 619–624, 2003.

[18] K. Veroy, C. Prud'homme, D.V. Rovas, and A.T. Patera, "*A posteriori* error bounds for reduced-basis approximation of parametrized noncoercive and nonlinear elliptic partial differential equations (AIAA Paper 2003-3847)," *Proceedings of the 16th AIAA Computational Fluid Dynamics Conference*, 2003.

[19] M. Meyer and H.G. Matthies, "Efficient model reduction in non-linear dynamics using the Karhunen–Loève expansion and dual-weighted-residual methods," *Comput. Mech.*, 31, 179–191, 2003.

[20] A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations*, 2nd edn. Springer, 1997.

[21] K. Veroy, D. Rovas, and A.T. Patera, "*A posteriori* error estimation for reduced-basis approximation of parametrized elliptic coercive partial differential equations: "Convex inverse" bound conditioners," *Control, Optim. Calculus Var.*, 8, 1007–1028, Special Volume: A tribute to J.-L. Lions, 2002.

[22] 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," *Submitted to International Journal for Numerical Methods in Fluids* (Special Issue — Proceedings for 2004 ICFD Conference on Numerical Methods for Fluid Dynamics, Oxford), 2004.

[23] N.C. Nguyen, *Reduced-Basis Approximation and* A Posteriori *Error Bounds for Nonaffine and Nonlinear Partial Differential Equations: Application to Inverse Analysis*, PhD Thesis, Singapore-MIT Alliance, National University of Singapore, In progress, 2005.

[24] M.A. Grepl, N.C. Nguyen, K. Veroy, A.T. Patera, and G.R. Liu, " Certified rapid solution of parametrized partial differential equations for real-time applications," *Proceedings of the* 2nd *Sandia Workshop of PDE-Constrained Optimization: Towards Real-Time and On-Line PDE-Constrained Optimization*, SIAM Computational Science and Engineering Book Series. Submitted, 2004.

[25] L. Sirovich, "Turbulence and the dynamics of coherent structures, Part 1: Coherent structures," *Q. Appl. Math.*, 45, 561–571, 1987.

[26] B. Roux (ed.), *Numerical Simulation of Oscillatory Convection in Low-Pr Fluids: A GAMM Workshop*, vol. 27 of *Notes on Numerical Fluids Mechanics*, Vieweg, 1990.

[27] N. Trudinger, "On imbedding into Orlicz spaces and some applications," *J. Math. Mech.*, 17, 473–484, 1967.

[28] G. Talenti, "Best constant in Sobolev inequality," *Ann. Mat. Pura Appl.*, 110, 353–372, 1976.

[29] G. Rozza, "Proceedings of the Third M.I.T. Conference on Computational Fluid and Solid Mechanics," June 14–17, 2005. In: K. Bathe (ed.), *Computational Fluid and Solid Mechanics.*, Elsevier, Submitted, 2005.

[30] G. Caloz and J. Rappaz, "Numerical analysis for nonlinear and bifurcation problems," In: P. Ciarlet and J. Lions (eds.), *Handbook of Numerical Analysis, vol. V, Techniques of Scientific Computing* (Part 2), Elsevier Science B.V, pp. 487–637, 1997.

[31] K. Ito and S.S. Ravindran, "A reduced basis method for control problems governed by PDEs," In: W. Desch, F. Kappel, and K. Kunisch (eds.), *Control and Estimation of Distributed Parameter Systems*, Birkhäuser, pp. 153–168, 1998.

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

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

[34] D. Rovas, *Reduced-Basis Output Bound Methods for Parametrized Partial Differential Equations*, PhD Thesis, Massachusetts Institute of Technology, Cambridge, MA, 2002.

[35] M.A. Grepl and A.T. Patera, *A posteriori* error bounds for reduced-basis approximations of parametrized parabolic partial differential equations, *M2AN Math. Model. Numer. Anal.*, To appear, 2005.

[36] Z.J. Bai, "Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems.", *Appl. Numer. Math.*, 43, 9–44, 2002.

[37] C. Prud'homme, D.V. Rovas, K. Veroy, and A.T. Patera, "A mathematical and computational framework for reliable real-time solution of parametrized partial differential equations," *M2AN Math. Model. Numer. Anal.*, 36, 747–771, 2002.