A LEAST-SQUARES FINITE ELEMENT REDUCED BASIS METHOD *

JEHANZEB HAMEED CHAUDHRY†, LUKE N. OLSON‡, AND PETER SENTZ§

Abstract. We present a reduced basis (RB) method for parametrized linear elliptic partial differential equations (PDEs) in a least-squares finite element framework. A rigorous and reliable error estimate is developed, and is shown to bound the error with respect to the exact solution of the PDE, in contrast to estimates that measure error with respect to a finite-dimensional (high-fidelity) approximation. It is shown that the first-order formulation of the least-squares finite element is a key ingredient. The method is demonstrated using numerical examples.

Key words. least-squares, finite elements, reduced basis

AMS subject classifications. 65N15, 65N30

1. Introduction.

In this work, we formulate a reduced basis method for the solution of linear elliptic partial differential equations (PDEs) based on the least-squares finite element method (LSFEM). In many engineering and scientific applications, PDEs often depend on one or more parameters, which reflect either physical properties (e.g., the viscosity of a fluid, the heat conductivity of a medium), source terms and boundary conditions, or the geometry of the domain in which the problem is posed. In the case of parametrized geometry, transformation techniques [17, 39, 44] are used to obtain a PDE on a parameter-independent reference domain Ω. Letting $\boldsymbol{\mu}$ be a vector containing the relevant parameters, we study linear elliptic PDEs of the form:

$$\mathcal{L}_{\boldsymbol{\mu}} u_{\boldsymbol{\mu}} = f_{\boldsymbol{\mu}}, \qquad \boldsymbol{x} \in \Omega, \tag{1.1}$$

where Ω is a bounded subset of $\mathbb{R}^d$, $d = 2, 3$. The subscript $\boldsymbol{\mu}$ conveys the fact that the operator $\mathcal{L}_{\boldsymbol{\mu}}$ and the functions $u_{\boldsymbol{\mu}}$ and $f_{\boldsymbol{\mu}}$ depend on the value of the parameter(s) contained in $\boldsymbol{\mu}$. In this work, we consider elliptic problems in (1.1) — e.g., the Poisson's Equation with different values for the thermal conductivity of a medium, or the Stokes Equations with a varying Reynolds number.

LSFEMs are widely used for the solution of PDEs arising in many applications in science and engineering like fluid flow, transport, hyperbolic equations, quantum chromodynamics, magnetohydrodynamics, biomolecular simulation, plasma, elasticity, liquid crystals etc. [1, 2, 5, 6, 9, 12, 14, 20, 21, 29, 34, 35, 38]. LSFEMs are based on minimizing the residual of the PDE in an appropriate norm, and have a number of attractive properties. The finite element discretization of the weak form yields symmetric positive definite linear systems that are often suitable for optimal multigrid solvers. Moreover, the bilinear form arising from LSFEM is coercive and continuous, thus allowing flexibility in the choice of finite element (FE) spaces. This is in contrast to a mixed method which requires that the FE spaces satisfy the inf-sup or the Ladyzhenskaya-Babuška-Brezzi condition [8]. An additional advantage of LSFEMs is that complex boundary conditions may be handled weakly by incorporating

*Submitted to the editors March 4, 2020.

Funding: J. Chaudhry's work is supported by the NSF-DMS 1720402.

†Department of Mathematics, University of New Mexico (jehanzeb@unm.edu, https://math.unm.edu/~jehanzeb).

‡Department of Computer Science (lukeo@illinois.edu, http://lukeo.cs.illinois.edu).

§Department of Computer Science (sentz2@illinois.edu).

them into the definition of the least-squares residual.

Least-squares finite element methods provide a robust and inexpensive *a posteriori* error estimate. This is a crucial ingredient in our approach to constructing a reduced basis method for LSFEMs. Moreover, while the additional auxiliary variables and resulting large linear systems is a potential drawback to LSFEMs, a reduced basis approach which preserves the accuracy of the full finite element discretization while being inexpensive to compute is especially appealing for this class of discretizations.

In many applications, solutions are computed for a wide range of parameter values (many-query context), or must be computed cheaply following a parameter measurement or estimation (real-time context) [10, 24, 40, 48, 54]. In the case of a finite element discretization, a system of linear equations is obtained that involves a large number of unknowns. If solutions must be obtained quickly or for many parameter sets, the solution of these linear systems is prohibitively expensive. Reduced basis methods are a form of model order reduction that offers the potential to decrease the dimension of the problem, exploiting the low dimensionality of the solution manifold through parametric dependence [43]. As a result, solutions based on the low order representation are constructed with low computational cost.

RB methods are separated into two stages: "offline" and "online" [22, 28, 44, 48]. During the offline stage, a set of representative solutions is constructed by sampling the parameter domain and computing high dimensional finite element solutions called full-order model (FOM) solutions or snapshots. Two standard approaches for the offline basis construction include Proper Orthogonal Decomposition (POD) [36, 55] and greedy sampling methods [28]. Greedy sampling methods often lead to a more computationally efficient offline stage and are used in numerous applications [22, 25, 28, 31, 44]. This work is thus restricted to reduced basis methods with a greedy sampling procedure. Details of POD applied to parametrized elliptic systems is found in [33].

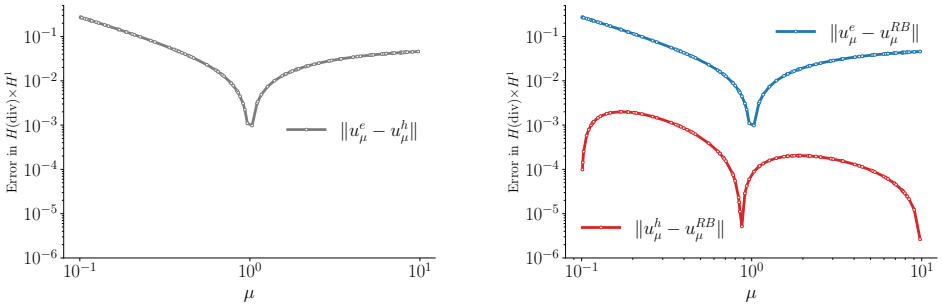
During the online stage, the previously constructed reduced basis is used to generate an inexpensive yet accurate solution for an estimated or measured set of parameters. The accuracy of this solution strongly depends on the sampling strategy and as well as the selection criteria for choosing the reduced basis.

The accuracy of a reduced basis solution is typically measured in reference to a full-order finite element solution [23, 25, 31, 49]. The error $\|u^h_{\boldsymbol{\mu}} - u^{\mathrm{RB}}_{\boldsymbol{\mu}}\|$ under an appropriate norm is heuristically minimized, where $u^h_{\boldsymbol{\mu}}$ and $u^{\mathrm{RB}}_{\boldsymbol{\mu}}$ are the full-order and reduced basis solutions, respectively. In essence, the full-order FE solution is treated as the exact solution for every parameter value; it is used as the benchmark for accuracy of the reduced order solution. However, the accuracy of the full order finite element solution is itself heavily dependent on the value of parameters for certain problems, resulting in an error estimate for the reduced basis solution that is often overly optimistic. In this article, a sharp error estimate with respect to the *exact* solution of the PDE is used in the construction of the reduced basis during the offline stage. This error estimate is provided by the relaxed smoothness requirements afforded by a first-order formulation, as well as a posteriori error estimate provided naturally by the LSFEM, and is inexpensive to compute, and provides an attractive feature of a LSFEM-based RB method.

To demonstrate the utility of measuring the accuracy of the reduced basis solution in terms of the exact solution, we consider a variable coefficient Poisson's problem, see § 6.1 for the detailed setup. The problem is dependent on a single parameter $\mu \in [10^{-1}, 10^{1}]$, which represents the thermal conductivity of one-half of an inhomogeneous

material. The solution is benign for $\mu = 1$, but features a discontinuous gradient for all other values. Thus, high accuracy requires a very fine mesh.

The left plot of Figure 1 shows the error[1] between a discrete solution and an "true" solution for different values of the parameter μ. The discrete solution u_μ^h is computed on a mesh with 1,065 degrees of freedom and a reference or "true" solution u_μ^e is computed on a mesh with 122,497 degrees of freedom. The error is particularly large for $\mu = 10^{-1}$. A reduced basis solution u_μ^{RB} is constructed on the same mesh as u_μ^h; the right plot of Figure 1 shows the error of this reduced-order solution with respect to both the reference solution and the full-order solution.



(a) Full-order error.

(b) Reduced basis error.

Fig. 1: The $H(\mathrm{div})\times H^1$-norm of the error between a full-order solution and a reference solution (left), and the error in the reduced-basis solution with respect to both the full-order solution and the reference (right).

In Figure 1, the error between the reduced-basis solution and the reference solution is four orders of magnitude greater than the error with respect to the discrete solution. Thus, a sharp, rigorous error bound based on $\|u_\mu^{RB} - u_\mu^h\|$ would significantly underestimate the error with respect to the true solution. For this reason, an error estimate is not reliable without first ensuring that the full-order solution is sufficiently accurate.

The paper is organized as follows. In § 2, we describe parametrized equations in a Hilbert space setting, and describe a rigorous error estimate for approximate solutions. In § 3, we introduce least-squares finite element methods and how they fit into the abstract Hilbert space context. In § 4, we review standard reduced basis methods, in particular, those based on residual error indicators and greedy-sampling. In § 5, we propose a LSFEM-based reduced basis method and in § 6 we provide several numerical examples. § 7 consists of conclusions and possibilities for future work.

2. Parameterized Equations and Error Bounds. In this section we set up the parameterized equations in a Hilbert space setting. In § 2.1 we discuss error bounds in this context and in § 2.2 we detail issues that arise when considering elliptic problems in a standard Galerkin setting.

Let X and Y be Hilbert spaces, and let $\mathcal{D}$ be a compact subset of $\mathbb{R}^P$, with $P \geq 1$.

[1]The error is measured in the $H(\mathrm{div}) \times H^1$-norm, which is the appropriate norm for the least-squares setup for the Poisson's problem. See § 6.1 for details.

For any $\boldsymbol{\mu} \in \mathcal{D}$, we assume the existence of a linear operator

$$\mathcal{L}_{\boldsymbol{\mu}} : X \to Y. \tag{2.1}$$

For a fixed $f_{\boldsymbol{\mu}} \in Y$, we seek $u_{\boldsymbol{\mu}} \in X$ that satisfies

$$\mathcal{L}_{\boldsymbol{\mu}} u_{\boldsymbol{\mu}} = f_{\boldsymbol{\mu}}. \tag{2.2}$$

We further assume that for any $\boldsymbol{\mu} \in \mathcal{D}$ there exists a parameter dependent *coercivity* constant $\alpha(\boldsymbol{\mu})$ and a *continuity* constant $\gamma(\boldsymbol{\mu})$ with $0 < \alpha(\boldsymbol{\mu}) \leq \gamma(\boldsymbol{\mu}) < \infty$ such that

$$\alpha(\boldsymbol{\mu})\|v\|_X^2 \leq \|\mathcal{L}_{\boldsymbol{\mu}} v\|_Y^2 \leq \gamma(\boldsymbol{\mu})\|v\|_X^2, \qquad \forall v \in X. \tag{2.3}$$

That is, $\mathcal{L}_{\boldsymbol{\mu}}$ and its inverse are bounded.

In order to approximate $u_{\boldsymbol{\mu}}$, we introduce a finite-dimensional subspace $X^h \subset X$ and seek a function $u_{\boldsymbol{\mu}}^h \in X^h$. The subspace X^h may correspond to any general discretization procedure, e.g., finite differences, finite elements, or from a reduced order model.

For a particular parameter $\boldsymbol{\mu}$, we define the error to be

$$e_{\boldsymbol{\mu}}^h \coloneqq u_{\boldsymbol{\mu}} - u_{\boldsymbol{\mu}}^h, \tag{2.4}$$

which is a measure of the quality of this approximation. Developing a rigorous and strict upper bound for the norm of the error $\|e_{\boldsymbol{\mu}}^h\|_X$ is important for assessing the quality of the numerical approximation. Likewise, the residual is defined as

$$r_{\boldsymbol{\mu}}^h \coloneqq f_{\boldsymbol{\mu}} - \mathcal{L}_{\boldsymbol{\mu}} u_{\boldsymbol{\mu}}^h, \tag{2.5}$$

and we see that $e_{\boldsymbol{\mu}}^h$ satisfies the error equation

$$\mathcal{L}_{\boldsymbol{\mu}} e_{\boldsymbol{\mu}}^h = r_{\boldsymbol{\mu}}^h. \tag{2.6}$$

2.1. Error Bounds. Our approach to developing rigorous upper bounds on the error is to begin with (2.3) and (2.6), which leads to

$$\begin{aligned} \alpha(\boldsymbol{\mu})\|e_{\boldsymbol{\mu}}^h\|_X^2 &\leq \|\mathcal{L}_{\boldsymbol{\mu}} e_{\boldsymbol{\mu}}^h\|_Y^2 = \|r_{\boldsymbol{\mu}}^h\|_Y^2, \\ \Rightarrow \|e_{\boldsymbol{\mu}}^h\|_X &\leq \frac{\|r_{\boldsymbol{\mu}}^h\|_Y}{\sqrt{\alpha(\boldsymbol{\mu})}}. \end{aligned} \tag{2.7}$$

Unfortunately, this upper bound proves to be extremely pessimistic, especially for problems for which the coercivity constant $\alpha(\boldsymbol{\mu})$ is relatively small, a common scenario. This is illustrated with a simple finite dimensional example.

Let $X = Y = \mathbb{R}^n$ under the standard Euclidean norm. Consider the operator $A : X \to Y$ represented by the matrix

$$A = \begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix}, \tag{2.8}$$

which is positive-definite with smallest eigenvalue $\lambda_1 = 4\sin^2\left(\frac{\pi}{2(n+1)}\right)$. In addition consider the right-hand side $f = [1, 0, \ldots, 0, 1]^T$, which yields a solution to $Au = f$ of $u = [1, 1, \ldots, 1]^T$. Then, consider the perturbation $\hat{u} \in \mathbb{R}^n$ given by

$$\hat{u}_i = 1 + \frac{(-1)^i}{n}. \tag{2.9}$$

The error in this case is given by $\|u - \hat{u}\|_X = \frac{1}{\sqrt{n}}$ and the residual by $\|r\|_X = \|f - Au\|_X = \frac{\sqrt{16n-14}}{n}$. Thus, both the error and residual converge to zero as $n \to \infty$. However, the ratio

$$\frac{\|r\|_X}{\sqrt{\alpha}} = \frac{\|r\|_X}{\sqrt{\lambda_1}} = \frac{\sqrt{16n-14}}{2n\sin\left(\frac{\pi}{2(n+1)}\right)} > \frac{4\sqrt{n-1}}{\pi} \tag{2.10}$$

is unbounded for large n. The error $u - \hat{u}$ has no component in the span of the eigenvector of A corresponding to λ_1. Thus, reflecting on (2.7), the ratio of the residual to the square root of the coercivity constant is not an accurate predictor of the norm of the error.

As a consequence, our goal is to improve the error bound in (2.7). We do so by computing an approximation to the error in a finite-dimensional subspace $Z^h \subset X$ (we do not exclude the possibility that $Z^h = X^h$ or $X^h \cap Z^h = \{0\}$), and denote this approximation by $\hat{e}^h_{\boldsymbol{\mu}}$.

We introduce the *auxiliary* or *error residual*

$$\rho^h_{\boldsymbol{\mu}} \coloneqq r^h_{\boldsymbol{\mu}} - \mathcal{L}_{\boldsymbol{\mu}}\hat{e}^h_{\boldsymbol{\mu}}. \tag{2.11}$$

Analogous to the previous bound (2.7), with this form we arrive at

$$\begin{aligned} \alpha(\boldsymbol{\mu})\|e^h_{\boldsymbol{\mu}} - \hat{e}^h_{\boldsymbol{\mu}}\|^2_X \le \left\|\mathcal{L}_{\boldsymbol{\mu}}\left(e^h_{\boldsymbol{\mu}} - \hat{e}^h_{\boldsymbol{\mu}}\right)\right\|^2_Y = \|r^h_{\boldsymbol{\mu}} - \mathcal{L}_{\boldsymbol{\mu}}\hat{e}^h_{\boldsymbol{\mu}}\|^2_Y = \|\rho^h_{\boldsymbol{\mu}}\|^2_Y, \\ \Rightarrow \|e^h_{\boldsymbol{\mu}} - \hat{e}^h_{\boldsymbol{\mu}}\|_X \le \frac{\|\rho^h_{\boldsymbol{\mu}}\|_Y}{\sqrt{\alpha(\boldsymbol{\mu})}}. \end{aligned} \tag{2.12}$$

In the case that the approximation to the error is simply taken to be $\hat{e}^h_{\boldsymbol{\mu}} = 0$, then $\|\rho^h_{\boldsymbol{\mu}}\|_Y = \|r^h_{\boldsymbol{\mu}}\|_Y$. However, if a reasonable approximation to the error is computed, it is often the case that $\|\rho^h_{\boldsymbol{\mu}}\|_Y \ll \|r^h_{\boldsymbol{\mu}}\|_Y$, resulting is less sensitivity to a small coercivity constant.

We use (2.12) and the triangle inequality to develop an alternative upper bound for $\|e^h_{\boldsymbol{\mu}}\|_X$:

$$\|e^h_{\boldsymbol{\mu}}\|_X \le \|\hat{e}^h_{\boldsymbol{\mu}}\|_X + \|e^h_{\boldsymbol{\mu}} - \hat{e}^h_{\boldsymbol{\mu}}\|_X \le \|\hat{e}^h_{\boldsymbol{\mu}}\|_X + \frac{\|\rho^h_{\boldsymbol{\mu}}\|_Y}{\sqrt{\alpha(\boldsymbol{\mu})}} =: M^h(\boldsymbol{\mu}). \tag{2.13}$$

With this form of the error bound we monitor its effectiveness with the so-called *effectivity* ratio, defined as

$$\frac{M^h(\boldsymbol{\mu})}{\|e^h_{\boldsymbol{\mu}}\|_X}, \tag{2.14}$$

which we seek as close to one as possible. The effectivity ratio is bounded in the following, which is adapted from [50]:

THEOREM 2.1. *Fix $\delta \in [0,1)$ and $\boldsymbol{\mu} \in \mathcal{D}$. Let $u_{\boldsymbol{\mu}}$ be the solution to (2.2), and let $u^h_{\boldsymbol{\mu}}$ be its discrete approximation in X^h, with residual $r^h_{\boldsymbol{\mu}}$ as defined in (2.5). Denote the error as $e^h_{\boldsymbol{\mu}}$ (cf. (2.4)) and consider $\hat{e}^h_{\boldsymbol{\mu}}$ to be any approximation of this error. Finally, let $\rho^h_{\boldsymbol{\mu}}$ denote the auxiliary residual (cf. (2.11)). If*

$$\frac{\|\rho^h_{\boldsymbol{\mu}}\|_Y}{\sqrt{\alpha(\boldsymbol{\mu})}\,\|\hat{e}^h_{\boldsymbol{\mu}}\|_X} \leq \delta, \tag{2.15}$$

then the effectivity satisfies the following bound:

$$\frac{M^h(\boldsymbol{\mu})}{\|e^h_{\boldsymbol{\mu}}\|_X} \leq \frac{1+\delta}{1-\delta} \tag{2.16}$$

Proof. Assume (2.15) holds. By the reverse triangle inequality, (2.12), and (2.15), we have

$$\left|\frac{\|e^h_{\boldsymbol{\mu}}\|_X - \|\hat{e}^h_{\boldsymbol{\mu}}\|_X}{\|\hat{e}^h_{\boldsymbol{\mu}}\|_X}\right| \leq \frac{\|\hat{e}^h_{\boldsymbol{\mu}} - e^h_{\boldsymbol{\mu}}\|_X}{\|\hat{e}^h_{\boldsymbol{\mu}}\|_X} \leq \frac{\|\rho^h_{\boldsymbol{\mu}}\|_Y}{\sqrt{\alpha(\boldsymbol{\mu})}\,\|\hat{e}^h_{\boldsymbol{\mu}}\|_X} \leq \delta. \tag{2.17}$$

If $\|\hat{e}^h_{\boldsymbol{\mu}}\|_X > \|e^h_{\boldsymbol{\mu}}\|_X$, then it follows from (2.17) that

$$\|\hat{e}^h_{\boldsymbol{\mu}}\|_X - \|e^h_{\boldsymbol{\mu}}\|_X \leq \delta\|\hat{e}^h_{\boldsymbol{\mu}}\|_X \implies (1-\delta)\|\hat{e}^h_{\boldsymbol{\mu}}\|_X \leq \|e^h_{\boldsymbol{\mu}}\|_X. \tag{2.18}$$

If $\|\hat{e}^h_{\boldsymbol{\mu}}\|_X \leq \|e^h_{\boldsymbol{\mu}}\|_X$, then $(1-\delta)\|\hat{e}^h_{\boldsymbol{\mu}}\|_X \leq \|e^h_{\boldsymbol{\mu}}\|_X$ follows immediately since $\delta \geq 0$. In either case,

$$(1-\delta)\|\hat{e}^h_{\boldsymbol{\mu}}\|_X \leq \|e^h_{\boldsymbol{\mu}}\|_X, \tag{2.19}$$

holds. Using (2.19), (2.12), and (2.15), it follows that

$$\frac{\|\hat{e}^h_{\boldsymbol{\mu}} - e^h_{\boldsymbol{\mu}}\|_X}{\|e^h_{\boldsymbol{\mu}}\|_X} \leq \frac{\|\hat{e}^h_{\boldsymbol{\mu}} - e^h_{\boldsymbol{\mu}}\|_X}{(1-\delta)\|\hat{e}^h_{\boldsymbol{\mu}}\|_X} \leq \frac{\|\rho^h_{\boldsymbol{\mu}}\|_Y}{\sqrt{\alpha(\boldsymbol{\mu})}(1-\delta)\|\hat{e}^h_{\boldsymbol{\mu}}\|_X} \leq \frac{\delta}{1-\delta}. \tag{2.20}$$

Finally, using the triangle inequality, (2.15), (2.19), and (2.20), we have

$$\begin{aligned} M^h(\boldsymbol{\mu}) = \|\hat{e}^h_{\boldsymbol{\mu}}\|_X + \frac{\|\rho^h_{\boldsymbol{\mu}}\|_Y}{\sqrt{\alpha(\boldsymbol{\mu})}} &\leq \|e^h_{\boldsymbol{\mu}}\|_X + \|\hat{e}^h_{\boldsymbol{\mu}} - e^h_{\boldsymbol{\mu}}\|_X + \frac{\|\rho^h_{\boldsymbol{\mu}}\|_Y}{\sqrt{\alpha(\boldsymbol{\mu})}} \\ &\leq \|e^h_{\boldsymbol{\mu}}\|_X + \|\hat{e}^h_{\boldsymbol{\mu}} - e^h_{\boldsymbol{\mu}}\|_X + \delta\|\hat{e}^h_{\boldsymbol{\mu}}\| \\ &= \left(1 + \frac{\|\hat{e}^h_{\boldsymbol{\mu}} - e^h_{\boldsymbol{\mu}}\|_X}{\|e^h_{\boldsymbol{\mu}}\|_X} + \delta\frac{\|\hat{e}^h_{\boldsymbol{\mu}}\|_X}{\|e^h_{\boldsymbol{\mu}}\|_X}\right)\|e^h_{\boldsymbol{\mu}}\|_X \\ &\leq \left(1 + \frac{\delta}{1-\delta} + \frac{\delta}{1-\delta}\right)\|e^h_{\boldsymbol{\mu}}\|_X \\ &= \left(\frac{1+\delta}{1-\delta}\right)\|e^h_{\boldsymbol{\mu}}\|_X, \end{aligned} \tag{2.21}$$

which completes the proof. ∎

2.2. Application to Poisson's equation: Galerkin setting. Bound (2.13) is only useful if the inner products associated with the spaces X and Y are computable, and if there are easily constructed conforming subspaces X^h, $Z^h \subset X$. We will demonstrate possible consequences by considering an example of the parameter-independent Poisson's equation with homogeneous boundary conditions:

$$\begin{aligned} -\Delta u &= f, \qquad \boldsymbol{x} \in \Omega, \\ u &= 0, \qquad \boldsymbol{x} \in \partial\Omega, \end{aligned} \tag{2.22}$$

with $\mathcal{L} \coloneqq -\Delta$. From here, we have a number of different choices for the domain and range. One option is $X = H^2(\Omega) \cap H_0^1(\Omega)$, and $Y = L^2(\Omega)$; both norms are easily computable. However, to compute discrete approximations u^h and $\hat{e}^h$, we must construct finite element spaces X^h and Z^h that contain functions that are class $C^1(\Omega)$ across element boundaries, which are difficult to construct [8], and are not usually used in the numerical approximation of Poisson's equation.

Alternatively, consider $X = H_0^1(\Omega)$ and $Y = H^{-1}(\Omega) = \left(H_0^1(\Omega)\right)'$. In this setting, the Poisson equation is often solved using variational methods, resulting in the Galerkin weak form of the equation:

$$a(u,v) \coloneqq \int_\Omega \nabla u \cdot \nabla v \; dx = \int_\Omega f v \; dx =: F(v), \qquad \forall v \in H_0^1(\Omega). \tag{2.23}$$

Using the language of duality pairings (see for example [11]), it is possible to express this through a mapping $\mathcal{L} : X \to Y$, with $X = H_0^1(\Omega)$ and $Y = H^{-1}(\Omega)$, via

$$\mathcal{L}u\,[\cdot] \coloneqq \int_\Omega \nabla u \cdot \nabla\,[\cdot]\,\mathrm{d}x. \tag{2.24}$$

That is, $\mathcal{L}u = F \in H^{-1}(\Omega)$. Standard conforming finite element spaces are readily constructed for X, and the norm on X is easily computable. However, the $Y = H^{-1}(\Omega)$ norm requires inversion of the Laplacian operator [6]:

$$\|F\|_Y = \left((-\Delta)^{-1/2}F, (-\Delta)^{-1/2}F\right)_0^{1/2}. \tag{2.25}$$

Consequently, to compute the Y-norm of the auxiliary residual $\rho_{\boldsymbol{\mu}}^h = f_{\boldsymbol{\mu}} + \Delta u_{\boldsymbol{\mu}}^h + \Delta \hat{e}_{\boldsymbol{\mu}}^h$, we would need to compute $(-\Delta)^{-1} f$, which is exactly the equation for which we seek an error bound.

In the finite element setting, the infinite dimensional space $H_0^1(\Omega)$ is not dealt with directly, but instead a finite dimensional *test* subspace $V^h \subset H_0^1(\Omega)$ is introduced. The restriction of F to the subspace V^h is a bounded linear functional on V^h, so that F is identified with an element in $\left(V^h\right)'$. Thus, $\mathcal{L}u = F \in \left(V^h\right)'$, allowing us to associate Y with $\left(V^h\right)'$. While the norm for $Y = \left(V^h\right)'$ is more complex than either the L^2 or H^1 norms, it is still computable due to its finite dimension [48].

Unfortunately, the operator fails to be coercive in this case, which is seen either by using the fact that X is infinite dimensional and Y is finite dimensional, or by observing the standard Galerkin orthogonality condition:

$$a(u - u^h, v^h) = 0, \qquad \forall v^h \in V^h. \tag{2.26}$$

This implies that

$$\mathcal{L}(u - u^h) = 0 \in \left(V^h\right)'. \tag{2.27}$$

That is, even if $u - u^h \neq 0$, the image $\mathcal{L}(u - u^h)$ is zero when considered as an element of the space $\left(V^h\right)'$.

Defining a finite-dimensional *trial* subspace $W^h \subset H_0^1(\Omega)$ (where W^h coincides with V^h in the standard Galerkin method), standard ellipticity results [11] show that $\mathcal{L}$ in (2.24) is coercive on $X = W^h$. However, the exact solution to (2.2) does not belong to W^h in general. For this reason, reduced basis approximations for standard Galerkin methods typically consider the "true" solution as a discrete solution in the finite dimensional subspace $W^h \subset X$. As a result, it is not possible to apply the error bounds to an exact solution $u \notin W^h$ in the Ritz-Galerkin finite element setting. In the next section, we show that this problem does not arise in a LSFEM context.

3. The Least-Squares Finite Element Method. The least-squares finite element method (LSFEM) reformulates the PDE as a system of first-order equations and then defines the solution as the minimizer of a functional in an appropriate norm. See [6, 32] for a complete description; a brief overview, with application to parametrized equations, is presented in this section.

3.1. Abstract Formulation. In addition to the assumptions of the previous section, we consider $\mathcal{L}_{\boldsymbol{\mu}}$ to be a bounded, linear first-order differential operator. We wish to solve (2.2). Under the assumptions given by (2.3), any solution to (2.2) is the unique minimizer of the following problem:

$$\arg\min_{v \in X} J_{\boldsymbol{\mu}}(v; f_{\boldsymbol{\mu}}) \coloneqq \|\mathcal{L}_{\boldsymbol{\mu}} v - f_{\boldsymbol{\mu}}\|_Y^2. \tag{3.1}$$

Conversely, (3.1) is guaranteed to have a unique minimizer $u_{\boldsymbol{\mu}} \in X$, and if $f_{\boldsymbol{\mu}}$ belongs to the range of $\mathcal{L}_{\boldsymbol{\mu}}$, this minimizer also solves (2.2). $u_{\boldsymbol{\mu}}$ necessarily satisfies the first-order optimality condition:

$$(\mathcal{L}_{\boldsymbol{\mu}} u_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} v)_Y = (f_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} v)_Y, \qquad \forall v \in X. \tag{3.2}$$

For the remainder of the paper, we denote $u_{\boldsymbol{\mu}}$ as the unique solution to (3.1) and (3.2); i.e., $\mathcal{L}_{\boldsymbol{\mu}}$ is surjective.

A LSFEM is defined by choosing a finite element subspace $X^h \subset X$, and seeking the minimum to (3.1) over this subspace instead. The first-order optimality condition is now: find $u_{\boldsymbol{\mu}}^h \in X^h$ such that

$$(\mathcal{L}_{\boldsymbol{\mu}} u_{\boldsymbol{\mu}}^h, \mathcal{L}_{\boldsymbol{\mu}} v^h)_Y = (f_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} v^h)_Y, \qquad \forall v^h \in X^h. \tag{3.3}$$

Since $X^h \subset X$, coercivity of the bilinear form $a(\cdot, \cdot; \boldsymbol{\mu}) \coloneqq (\mathcal{L}_{\boldsymbol{\mu}} u_{\boldsymbol{\mu}}^h, \mathcal{L}_{\boldsymbol{\mu}} v^h)_Y$ and continuity of $F(\cdot) \coloneqq (f_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} v^h)_Y$ on X^h follow immediately. Thus, (3.3) admits a unique solution $u_{\boldsymbol{\mu}}^h \in X^h$.

The corresponding linear system of algebraic equations

$$A_{\boldsymbol{\mu}}^h \boldsymbol{u}_{\boldsymbol{\mu}}^h = \boldsymbol{b}_{\boldsymbol{\mu}}^h \tag{3.4}$$

that is solved for the unknown vector of degrees of freedom $\boldsymbol{u}_{\boldsymbol{\mu}}^h$, is also symmetric positive-definite. This follows from the symmetry and coercivity of $a(\cdot, \cdot; \boldsymbol{\mu})$.

3.2. Error Approximation. To approximate the error $e_{\boldsymbol{\mu}}^h = u_{\boldsymbol{\mu}} - u_{\boldsymbol{\mu}}^h$ for the LSFEM, first an approximation $u_{\boldsymbol{\mu}}^h \in X^h \subset X$ is computed via (3.3), and then the residual $r_{\boldsymbol{\mu}}^h = f_{\boldsymbol{\mu}} - \mathcal{L}_{\boldsymbol{\mu}} u_{\boldsymbol{\mu}}^h$ is computed. Because of the form of (3.3), the corresponding error satisfies

$$(\mathcal{L}_{\boldsymbol{\mu}} e_{\boldsymbol{\mu}}^h, \mathcal{L}_{\boldsymbol{\mu}} v^h)_Y = 0, \qquad \forall v^h \in X^h. \tag{3.5}$$

As a result, if we attempt to compute an approximate error $\hat{e}^h_{\boldsymbol{\mu}} \in X^h$, we obtain zero. To alleviate this, we introduce an additional subspace Z^h that satisfies $X^h \subset Z^h \subset X$. In the context of a finite element method, Z^h may represent a refinement of the mesh, an increase in the polynomial order of the elements, or both. We then solve for an approximation $Z^h \ni \hat{e}^h_{\boldsymbol{\mu}} \approx e^h_{\boldsymbol{\mu}}$ through the variational problem:

$$(\mathcal{L}_{\boldsymbol{\mu}} \hat{e}^h_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} w^h)_Y = (r^h_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} w^h)_Y, \qquad \forall w^h \in Z^h. \tag{3.6}$$

Given the refinement of the space with $X^h \subset Z^h$, the auxiliary residual is expected to satisfy $\|\rho^h_{\boldsymbol{\mu}}\|_Y \ll \|r^h_{\boldsymbol{\mu}}\|_Y$. Thus, the rigorous error bound $M^h(\boldsymbol{\mu})$ is applicable, and if the hypotheses of Theorem 2.1 hold, then the bounds on the effectivity are computable as well.

It is the first-order formulation of the PDE that allows us to extend the theory from § 2 to form a practical method. Any first-order formulation that leads to a practical LSFEM will lead to a space X that is approximated by easily constructed finite element spaces, and a space Y with an easily computable inner product. This leads to practical and computable norms $\|\cdot\|_X$ and $\|\cdot\|_Y$. We demonstrate this by continuing the example of the Poisson equation from § 2.2. An equivalent first-order system of PDEs is given by

$$\begin{aligned} \boldsymbol{q} + \nabla u &= 0, && \boldsymbol{x} \in \Omega, \\ \nabla \cdot \boldsymbol{q} &= 0, && \boldsymbol{x} \in \Omega, \\ u &= 0, && \boldsymbol{x} \in \partial\Omega. \end{aligned} \tag{3.7}$$

The corresponding first-order differential operator is

$$\begin{aligned} &\mathcal{L} : H(\mathrm{div}) \times H^1(\Omega) \to \left[L^2(\Omega)\right]^d \times L^2(\Omega), \\ &\mathcal{L}\left[(\boldsymbol{q}, u)\right] \coloneqq \begin{pmatrix} \boldsymbol{q} + \nabla u \\ \nabla \cdot \boldsymbol{q} \end{pmatrix}. \end{aligned} \tag{3.8}$$

The operator $\mathcal{L}$ satisfies (2.3) on the Hilbert spaces $X = H(\mathrm{div}) \times H^1$, $Y = \left(L^2\right)^d \times L^2$ [42], so the theory of § 2 applies. The norm of $Y = \left(L^2\right)^d \times L^2$ is easily computable; moreover, simple conforming finite element spaces exist for $X = H(\mathrm{div}) \times H^1$ [8, 47]. Thus, the computational difficulties associated with the pairings $X = H^1_0$, $Y = H^{-1}$ and $X = H^2 \cap H^1_0$, $Y = L^2$ from § 2.2 are not present.

Furthermore, the LSFEM method minimizes the norm of the auxiliary residual $\|\rho^h_{\boldsymbol{\mu}}\|_Y$ by design. This is a desirable property in light of the discussion in § 2.1.

4. Reduced Basis Methods. In this section, we provide a brief overview of reduced basis (RB) methods for elliptic equations. See [48] for an extensive overview. As described in (3.4), the LSFEM discretization of a linear elliptic PDE leads to a parameter-dependent system of linear equations. A Galerkin finite element method will also lead to a system of linear equations of the same form. Thus, the algebraic considerations of reduced basis methods carry over from Galerkin methods to LSFEMs in a straightforward way.

While we restrict our attention to steady-state problems, LSFEMs have also been successfully applied to time-dependent parabolic problems [6, 53, 56]. Thus, for these two classes of problems, standard projection-based reduced order modeling approaches (e.g., Galerkin and Petrov-Galerkin [13, 15, 16]) can be applied. Hyperbolic problems have proved to be more challenging for LSFEMs, although work has been done in this area [4, 20, 41].

4.1. Galerkin Projection. A parametrized elliptic PDE solved by a Galerkin variational method (e.g., a finite element method), leads to the equation:

$$a(u^h_{\boldsymbol{\mu}}, v^h; \boldsymbol{\mu}) = F(v^h; \boldsymbol{\mu}), \qquad \forall v^h \in X^h. \tag{4.1}$$

Here, $a(\cdot, \cdot; \boldsymbol{\mu}) : X \times X \to \mathbb{R}$ is a continuous and coercive bilinear form for all $\boldsymbol{\mu} \in \mathcal{D} \subset \mathbb{R}^d$, and $F(\cdot; \boldsymbol{\mu}) : X \to \mathbb{R}$ is a bounded linear functional for all $\boldsymbol{\mu}$.

Let $N^h \coloneqq \dim(X^h)$ and consider $\{\eta_j\}_{j=1}^{N^h}$ to be a basis for X^h. For any $\boldsymbol{\mu}$, the discrete solution has a representation $u^h_{\boldsymbol{\mu}} = \sum_{j=1}^{N^h} u_j(\boldsymbol{\mu})\eta_j$, where $u_j(\boldsymbol{\mu})$ denotes the coefficient to basis function η_j and depending on $\boldsymbol{\mu}$. Substitution into (4.1), results in a linear system of the form

$$\sum_{j=1}^{N^h} a(\eta_j, \eta_i; \boldsymbol{\mu}) u_j(\boldsymbol{\mu}) = F(\eta_i; \boldsymbol{\mu}) \qquad i = 1, \ldots, N^h. \tag{4.2}$$

In a many query or real-time context, (4.2) must be solved repeatedly or very quickly. Even more, a large linear system must be assembled for *each* parameter instance, which is prohibitively expensive for discretizations with many degrees of freedom. Reduced basis methods are intended to help alleviate this cost. By introducing a subspace $X^N \subset X$ with dimension $N \ll N^h$ and basis $\{\xi_j\}_{j=1}^N$, a reduced solution $u^N_{\boldsymbol{\mu}} = \sum_{j=1}^N c_j(\boldsymbol{\mu})\xi_j$ is sought instead. This leads to the much smaller linear system

$$\sum_{j=1}^{N} a(\xi_j, \xi_i; \boldsymbol{\mu}) c_j(\boldsymbol{\mu}) = F(\xi_i; \boldsymbol{\mu}) \qquad i = 1, \ldots, N. \tag{4.3}$$

There are a number of features that distinguish a RB method. First, an RB method must specify *how* the reduced basis $\{\xi_j\}$ is constructed, which is part of the "offline" stage. This "offline-online" decomposition is found throughout the reduced basis literature [44, 48]. Typically, the basis functions are linear combinations of the high-fidelity basis functions η_j. We review the greedy sampling strategy for constructing the reduced basis in § 4.4.

Second, a RB method requires the construction of an efficient error indicator $\widetilde{M}^N(\boldsymbol{\mu})$ that quantifies the quality of the RB solution $u^N_{\boldsymbol{\mu}}$ in some manner. This is used both to assess the quality of the computed RB solution in the online stage, and to guide the construction of the reduced basis when using a greedy sampling strategy in the offline stage. We review the standard error indicator used in reduced basis literature in § 4.3 and discuss our improved error indicator for the least-squares reduced basis method in § 5.

Finally, a RB method is distinguished by the handling of the resulting reduced system (4.3), which still requires considerable cost in the assembly process, despite the reduction, because each new value of $\boldsymbol{\mu}$ requires a new linear system and right-hand side. The cost of this assembly is, in general, dependent on the dimension N^h, which is unacceptable for the many-query or real-time context. Either additional assumptions on $a(\cdot, \cdot; \boldsymbol{\mu})$ and $F(\cdot; \boldsymbol{\mu})$ must be made, or an algorithm to remove this N^h dependency must be specified. This issue is addressed by considering affinely parametrized equations.

4.2. Affinely Parametrized Equations. A critical feature of an effective RB method is that the assembly of (4.3) should be independent of the dimension of

the full-order problem N^h to be useful in a many-query or real-time context. A certain class of variational problems exist where an N^h-independent assembly process is readily obtained. A variational problem is said to be *affinely parametrized* if it can be expressed in the form

$$\begin{aligned} a(u,v;\boldsymbol{\mu}) &= \sum_{k=1}^{Q_a} \theta_k^a(\boldsymbol{\mu}) a_k(u,v), \\ F(v;\boldsymbol{\mu}) &= \sum_{k=1}^{Q_F} \theta_k^F(\boldsymbol{\mu}) F_k(v). \end{aligned} \tag{4.4}$$

Here, $\{\theta_k^a\}_{k=1}^{Q_a}$ and $\{\theta_k^F\}_{k=1}^{Q_F}$ are a set of Q_a (respectively Q_F) scalar functions of $\boldsymbol{\mu}$, the $\{a_k(u,v)\}_{k=1}^{Q_a}$ are continuous, parameter-independent, bilinear forms, and the $\{F_k\}_{k=1}^{Q_F}$ are continuous, parameter-independent, linear functionals. When this is satisfied, equation (4.3) takes the form:

$$\sum_{j=1}^{N} \left(\sum_{k=1}^{Q_a} \theta_k^a(\boldsymbol{\mu}) a_k(\xi_j, \xi_i) \right) c_j(\boldsymbol{\mu}) = \sum_{\ell=1}^{Q_F} \theta_\ell^F(\boldsymbol{\mu}) F_\ell(\xi_i) \qquad i = 1, 2, \ldots, N. \tag{4.5}$$

That is, the system matrix and right hand side are simply linear combinations of the matrices and vectors

$$\begin{aligned} (A_k)_{ij} &\coloneqq a_k(\xi_j, \xi_i) \\ (\boldsymbol{b}_k)_i &\coloneqq F_k(\xi_i). \end{aligned} \tag{4.6}$$

These are assembled in the offline stage, leading to an online stage that is independent of the problem size N_h. While there there are RB methods that do not satisfy this property — e.g., using the empirical interpolation method [3] — the work here is restricted to affinely parametrized problems as in a host of other works [22, 23, 25, 27, 31, 37, 51, 57].

The requirement for affinely parametrized equations is no more restrictive for the least-squares method than it is for the Galerkin case. An example is in the case of the time-harmonic Maxwell's equation for the calculation of the electric field, $\boldsymbol{E}$ [27]. Let $\boldsymbol{J}$ be a known source term, μ the permeability, σ the conductivity, ϵ the permittivity, ω the frequency, and $\beta = i\omega\sigma - \omega^2\epsilon$, where $i = \sqrt{-1}$. The vector of parameters is thus $\boldsymbol{\mu} = (\mu, \sigma, \epsilon, \omega)$. Introducing a test function $\boldsymbol{v}$, the variational equation becomes

$$\frac{1}{\mu}(\nabla \times \boldsymbol{E}, \nabla \times \boldsymbol{v})_0 + \beta(\boldsymbol{E}, \boldsymbol{v})_0 = i\omega(\boldsymbol{J}, \boldsymbol{v})_0, \qquad \forall \boldsymbol{v} \in H(\mathrm{curl}). \tag{4.7}$$

where $(\cdot,\cdot)_0$ is the $L^2(\Omega)$ inner-product for vector valued functions. The resulting weak equation is affinely parametrized.

A least-squares discretization is be obtained by introducing the variable $\boldsymbol{q} = \mu^{-1}\nabla \times \boldsymbol{E}$. Introducing test functions $\boldsymbol{r}$ and $\boldsymbol{v}$, one obtains the weak formulation

$$\begin{aligned} &\left[(\nabla \times \boldsymbol{q}, \nabla \times \boldsymbol{r})_0 + (\boldsymbol{q}, \boldsymbol{r})_0\right] + \beta\left[(\nabla \times \boldsymbol{q}, \boldsymbol{v})_0 + (\boldsymbol{E}, \nabla \times \boldsymbol{r})_0\right] + \beta^2(\boldsymbol{E}, \boldsymbol{v})_0 \\ &- \frac{1}{\mu}\left[(\boldsymbol{q}, \nabla \times \boldsymbol{v})_0 + (\nabla \times \boldsymbol{E}, \boldsymbol{r})_0\right] + \frac{1}{\mu^2}(\nabla \times \boldsymbol{E}, \nabla \times \boldsymbol{v})_0 \\ &= i\omega(\boldsymbol{J}, \nabla \times \boldsymbol{r})_0 + i\beta\omega(\boldsymbol{J}, \boldsymbol{v})_0. \end{aligned} \tag{4.8}$$

We see that the least-squares discretization also leads to an affinely parametrized variational equation.

4.3. Error Indicator. For a reduced basis of dimension N and for every $\boldsymbol{\mu}$, there is a corresponding RB solution $u^N_{\boldsymbol{\mu}}$ and a corresponding weak residual $R^N(\cdot;\boldsymbol{\mu}) \in \left(X^h\right)'$ defined as

$$R^N(v^h;\boldsymbol{\mu}) \coloneqq F(v^h;\boldsymbol{\mu}) - a(u^N_{\boldsymbol{\mu}}, v^h;\boldsymbol{\mu}), \qquad \forall v^h \in X^h. \tag{4.9}$$

Reduced basis methods typically construct error indicators of the form

$$\widetilde{M}^N(\boldsymbol{\mu}) \coloneqq \frac{\|R^N(\cdot;\boldsymbol{\mu})\|_{\left(X^h\right)'}}{\beta^{\mathrm{LB}}_{\boldsymbol{\mu}}}, \tag{4.10}$$

where $\beta^{\mathrm{LB}}_{\boldsymbol{\mu}}$ is a lower bound of a coercivity or stability constant, which is computed via the Successive Constraint Method (SCM) [18, 19, 30, 31, 48, 49, 52]. SCM constructs a linear program of complexity independent of the problem size in the offline stage, similar to the construction of the reduced basis itself.

The indicator in (4.10) is an analogous quantity to

$$\frac{\|\rho^h_{\boldsymbol{\mu}}\|_Y}{\sqrt{\alpha(\boldsymbol{\mu})}}. \tag{4.11}$$

In [50], the error indicator was improved upon by introducing an auxiliary error residual as in § 2. However, as shown in § 2.2, an indicator based on the residual in (4.9) cannot be applied to the error with respect to an arbitrary function in H^1. We refer to [44] for a detailed explanation on the construction of R^N and its corresponding dual norm.

4.4. Offline and Online Stages using a Greedy Sampling Strategy. The task of the offline stage in the reduced basis method is to construct the actual basis $\{\xi_i\}$. A finite subset $\mathcal{D}_{\mathrm{train}} \subset \mathcal{D}$ is chosen to represent the space of possible parameter values. A parameter vector $\boldsymbol{\mu}_1 \in \mathcal{D}_{\mathrm{train}}$ is chosen arbitrarily.

Define $\tilde{\xi}_1 \in X^h$ to be the solution of

$$a(\tilde{\xi}_1, v^h;\boldsymbol{\mu}_1) = F(v^h;\boldsymbol{\mu}_1), \qquad \forall v^h \in X^h. \tag{4.12}$$

Then the first reduced basis function is

$$\xi_1 = \frac{\tilde{\xi}_1}{\|\tilde{\xi}_1\|_X}. \tag{4.13}$$

Suppose for $N \geq 1$, an orthonormal basis $\{\xi_1, \ldots, \xi_N\}$ has been constructed corresponding to parameters $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_N \in \mathcal{D}_{\mathrm{train}}$. For each $\boldsymbol{\mu} \in \mathcal{D}_{\mathrm{train}} \setminus \{\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_N\}$, let $u^N_{\boldsymbol{\mu}}$ be the solution to the projected variational problem

$$a(u^N_{\boldsymbol{\mu}}, \xi_i;\boldsymbol{\mu}) = F(\xi_i;\boldsymbol{\mu}), \qquad i = 1, \ldots, N. \tag{4.14}$$

Using the expression for $R^N \in (X^h)'$ given by (4.9), the next parameter value is chosen through

$$\boldsymbol{\mu}_{N+1} = \underset{\boldsymbol{\mu} \in \mathcal{D}_{\mathrm{train}} \setminus \{\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_N\}}{\arg\max} \widetilde{M}^N(\boldsymbol{\mu}), \tag{4.15}$$

where $\widetilde{M}^N(\boldsymbol{\mu})$ is defined in (4.10). The next basis function ξ_{N+1} is found after computing the full-order solution $u^h_{\boldsymbol{\mu}_{N+1}}$ to equation (4.14), and orthonormalizing against the existing basis functions in the appropriate inner product.

The algorithm terminates after either the dimension of the basis has reached an upper bound or $\widetilde{M}^N(\boldsymbol{\mu})$ is smaller than a preset tolerance. At this point, the matrices and vectors from (4.6) are computed.

In the subsequent online stage, having constructed a basis $\{\xi_1, \dots, \xi_N\}$, a reduced-order solution is easily obtained by solving the $N \times N$ linear system corresponding to the projected variational problem. The computational cost is thus independent of the dimension of X^h, an essential component of a computationally efficient online stage.

5. A Least-Squares Finite Element Method with Reduced Basis. We now develop a least-squares based reduced basis method, which we label LSFEM-RB. First, recall the *improved* error estimate

$$\|e^h_{\boldsymbol{\mu}}\|_X \leq \|\hat{e}^h_{\boldsymbol{\mu}}\|_X + \frac{\|\rho^h_{\boldsymbol{\mu}}\|_Y}{\sqrt{\alpha(\boldsymbol{\mu})}} = M^h(\boldsymbol{\mu}), \tag{5.1}$$

which is a rigorous upper bound for the error; its effectivity is also bounded by Theorem 2.1.

Next, we make use of two finite-dimensional finite element spaces, $X^h \subset Z^h \subset X$, to compute the numerical approximation to the PDE and to the error equation. To this end, we define

$$\begin{aligned} a(u, v; \boldsymbol{\mu}) &\coloneqq \left(\mathcal{L}_{\boldsymbol{\mu}} u, \mathcal{L}_{\boldsymbol{\mu}} v\right), \\ F(v; \boldsymbol{\mu}) &\coloneqq \left(f_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} v\right), \\ R(w, u; \boldsymbol{\mu}) &\coloneqq F(w; \boldsymbol{\mu}) - a(u, w; \boldsymbol{\mu}). \end{aligned} \tag{5.2}$$

Recall the assumption of affine parametric dependence from § 4.2,

$$\begin{aligned} a(u, v; \boldsymbol{\mu}) &= \sum_{k=1}^{Q_a} \theta^a_k(\boldsymbol{\mu}) a_k(u, v), \\ F(v; \boldsymbol{\mu}) &= \sum_{k=1}^{Q_f} \theta^F_k(\boldsymbol{\mu}) F_k(v), \end{aligned} \tag{5.3}$$

which is key for a computationally efficient online stage.

Finally, using the error estimate (5.1) requires knowledge of the coercivity constant $\alpha(\boldsymbol{\mu})$; replacing this by a lower bound $0 < \alpha_{\mathrm{LB}}(\boldsymbol{\mu}) \leq \alpha(\boldsymbol{\mu})$ also results in a rigorous upper bound for the error. Computationally, we make use of a lower bound computed via the Successive Constraint Method [18]. While this method guarantees a rigorous lower bound for the coercivity constant of a finite-dimensional subspace, it is still possible that it returns a value that is larger than the true coercivity constant $\alpha(\boldsymbol{\mu})$. This is addressed in more detail in § 5.3.

Remark. For the solution of the approximate error $\hat{e}^h_{\boldsymbol{\mu}}$, we use a space Z^h that contains the original finite element space X^h; this is obtained through refinement of the mesh or increasing the polynomial order. As an alternative it is tempting to build the error approximation on a subspace $Z^h \subset X^h$; however, this will lead to the approximation $\hat{e}^h_{\boldsymbol{\mu}} = 0$ because of Galerkin orthogonality. It is still possible to compute the error on a space $Z^h \not\subset X^h$, with Z^h having fewer degrees of freedom than X^h, as long as it is possible to transfer the solution $u^h_{\boldsymbol{\mu}} \in X^h$ onto the new grid. However, if the grid corresponding to Z^h is too coarse, the auxiliary residual $\|\rho^h_{\boldsymbol{\mu}}\|_Y$ will be too large, making the error estimate less effective.

5.1. Offline Algorithm. With an initial $\boldsymbol{\mu}_1 \in \mathcal{D}_{\text{train}}$ we compute the solution $u^h_{\boldsymbol{\mu}_1}$ to the equation

$$a(u^h_{\boldsymbol{\mu}_1}, v^h; \boldsymbol{\mu}_1) = F(v^h; \boldsymbol{\mu}_1), \qquad v^h \in X^h. \tag{5.4}$$

followed by the error approximation via the equation

$$a(\hat{e}^h_{\boldsymbol{\mu}_1}, w^h; \boldsymbol{\mu}_1) = R(w^h, \tilde{\xi}_1; \boldsymbol{\mu}_1) \qquad w^h \in Z^h. \tag{5.5}$$

We then normalize $u^h_{\boldsymbol{\mu}_1}$ and $\hat{e}^h_{\boldsymbol{\mu}_1}$ to have unit X-norm and obtain our first pair of basis functions $\xi_1 \in X^h$ and $\phi_1 \in Z^h$.

Assume then that we we have constructed two orthonormal bases $\{\xi_1, \dots, \xi_N\} \subset X^h$ and $\{\phi_1, \dots, \phi_N\} \subset Z^h$ corresponding to parameters $\{\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_N\}$. For each $\boldsymbol{\mu} \in \mathcal{D}_{\text{train}}$, we compute the solution to the projected problem

$$a(u^N_{\boldsymbol{\mu}}, \xi_i; \boldsymbol{\mu}) = F(\xi_i; \boldsymbol{\mu}), \qquad i = 1, \dots, N, \tag{5.6}$$

and the corresponding projected error from

$$a(\hat{e}^N_{\boldsymbol{\mu}}, \phi_i; \boldsymbol{\mu}) = R(\phi_i, u^N_{\boldsymbol{\mu}}; \boldsymbol{\mu}), \qquad i = 1, \dots N. \tag{5.7}$$

Defining the reduced residual by $r^N_{\boldsymbol{\mu}} = f_{\boldsymbol{\mu}} - \mathcal{L}_{\boldsymbol{\mu}} u^N_{\boldsymbol{\mu}}$, and the corresponding reduced auxiliary residual $\rho^N_{\boldsymbol{\mu}} = r^N_{\boldsymbol{\mu}} - \mathcal{L}_{\boldsymbol{\mu}} \hat{e}^N_{\boldsymbol{\mu}}$, the next parameter value is then selected through

$$\boldsymbol{\mu}_{N+1} = \underset{\boldsymbol{\mu} \in \mathcal{D}_{\text{train}} \setminus \{\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_N\}}{\arg\max} M^N(\boldsymbol{\mu}) \coloneqq \|\hat{e}^N_{\boldsymbol{\mu}}\|_X + \frac{\|\rho^N_{\boldsymbol{\mu}}\|_Y}{\sqrt{\alpha_{\text{LB}}(\boldsymbol{\mu})}}. \tag{5.8}$$

Here, we denote $M^N(\boldsymbol{\mu})$ as error estimate $M^h(\boldsymbol{\mu})$ when restricted to approximations in the N-dimensional subspaces $\text{span}\{\xi_1, \dots, \xi_N\}$ and $\text{span}\{\phi_1, \dots, \phi_N\}$. The basis functions ξ_{N+1} and ϕ_{N+1} are obtained from the full-order solutions $u^h_{\boldsymbol{\mu}_{N+1}}$ and $\hat{e}^h_{\boldsymbol{\mu}_{N+1}}$ by orthonormalizing against the existing basis functions in the X-inner product.

The algorithm terminates whenever

$$\frac{\|\rho^N_{\boldsymbol{\mu}}\|_Y}{\sqrt{\alpha(\boldsymbol{\mu})}\|\hat{e}^N_{\boldsymbol{\mu}}\|_X} \leq \delta, \qquad \forall \boldsymbol{\mu} \in \mathcal{D}_{\text{train}}. \tag{5.9}$$

for a prescribed tolerance $\delta \in (0, 1)$. During the course of the offline algorithm, δ is increased if a full-order error estimate is encountered that exceeds the current value; if the full-order error indicator for a given $\boldsymbol{\mu}$ value is not bounded by δ, then we cannot expect a reduced-order analogue to be bounded by this quantity either. If δ is too large at the end of the algorithm, mesh or polynomial refinement of the space is needed to increase accuracy.

Once the basis functions have been constructed, the reduced basis matrices with entries $a_k(\xi_i, \xi_j)$, $a_k(\phi_i, \phi_j)$, $a_k(\phi_i, \xi_j)$ and reduced basis vectors with entries $F_k(\xi_i)$ and $F_k(\phi_i)$ are assembled. The algorithm for the offline stage is given in Algorithm 5.1.

If the algorithm terminates with $N < N_{\max}$, and $\delta < 1$ then (5.9) and Theorem 2.1 imply

$$\frac{M^N(\boldsymbol{\mu})}{\|u_{\boldsymbol{\mu}} - u^N_{\boldsymbol{\mu}}\|_X} \leq \frac{1+\delta}{1-\delta}, \qquad \forall \boldsymbol{\mu} \in \mathcal{D}_{\text{train}}. \tag{5.10}$$

Thus, we obtain an upper bound for the effectivity ratio in $\mathcal{D}_{\text{train}}$, in addition to the error itself.

Algorithm 5.1 Least Squares Reduced Basis Offline Algorithm

Choose $\boldsymbol{\mu}_1 \in \mathcal{D}_{\text{train}}$
Compute full-order solutions $u^h_{\boldsymbol{\mu}_1}$ and $\hat{e}^h_{\boldsymbol{\mu}_1}$. ▷ (5.4), (5.5)
Normalize to obtain primal basis $\{\xi_1\}$, and error basis $\{\phi_1\}$.
for $n = 1, \ldots, N_{\max}$ **do**
  if $\frac{\|\rho^n_{\boldsymbol{\mu}}\|_Y}{\sqrt{\alpha_{\text{LB}}(\mu)}\|\hat{e}^n_{\boldsymbol{\mu}}\|_X} \leq \delta$ for all $\boldsymbol{\mu} \in \mathcal{D}_{\text{train}}$ **then** ▷ (5.9)
    Break
  end if
  $\boldsymbol{\mu}_{n+1} = \arg\max M^n(\boldsymbol{\mu})$ ▷ (5.8)
  Compute full-order solutions $u^h_{\boldsymbol{\mu}_{n+1}}$ and $\hat{e}^h_{\boldsymbol{\mu}_{n+1}}$.
  If full-order estimate $\frac{\|\rho^h_{\boldsymbol{\mu}_{n+1}}\|_Y}{\sqrt{\alpha_{\text{LB}}(\boldsymbol{\mu}_{n+1})}\|\hat{e}^h_{\boldsymbol{\mu}_{n+1}}\|_X} > \delta$, set $\delta = \frac{\|\rho^h_{\boldsymbol{\mu}_{n+1}}\|_Y}{\sqrt{\alpha_{\text{LB}}(\boldsymbol{\mu}_{n+1})}\|\hat{e}^h_{\boldsymbol{\mu}_{n+1}}\|_X}$.
  Orthonormalize $u^h_{\boldsymbol{\mu}_{n+1}}$ against $\{\xi_1, \ldots, \xi_n\}$ to obtain ξ_{n+1}, and append.
  Orthonormalize $\hat{e}^h_{\boldsymbol{\mu}_{n+1}}$ against $\{\phi_1, \ldots, \phi_n\}$ to obtain ϕ_{n+1}, and append.
end for
Assemble matrices $a_k(\xi_i, \xi_j)$, $a_k(\phi_i, \phi_j)$, and $a_k(\phi_i, \xi_j)$. ▷ (5.3)
Assemble vectors $F_k(\xi_i)$ and $F_k(\phi_i)$. ▷ (5.3)

5.2. Online Algorithm. In the online stage, for a new parameter $\boldsymbol{\mu}$, the corresponding RB approximation and error bound is computed as follows. First, the $N \times N$ projected problem is assembled via

$$\begin{aligned} A_N(\boldsymbol{\mu}) &= \sum_{k=1}^{Q_a} \theta^a_k(\boldsymbol{\mu}) a_k(\xi_j, \xi_i), \\ \boldsymbol{b}_N(\boldsymbol{\mu}) &= \sum_{m=1}^{Q_f} \theta^F_m(\boldsymbol{\mu}) F_m(\xi_i). \end{aligned} \tag{5.11}$$

Since the parameter-independent terms in (5.11) are assembled in the offline stage, assembly in the online stage incurs a cost of $\mathcal{O}(Q_a N^2 + Q_f N)$. The resulting dense system $A_N(\boldsymbol{\mu})\boldsymbol{c}_N = \boldsymbol{b}_N(\boldsymbol{\mu})$ is solved directly, incurring a cost of $\mathcal{O}(N^3)$. Here $\boldsymbol{c}_N \in \mathbb{R}^N$ are the the RB coefficients, that is, $u_N(\boldsymbol{\mu}) = \sum_{n=1}^N [\boldsymbol{c}_N]_n \xi_n$.

To compute the approximation of the error, the $N \times N$ matrix

$$\widehat{A}_N(\boldsymbol{\mu}) = \sum_{k=1}^{Q_a} \theta^a_k(\boldsymbol{\mu}) a_k(\phi_j, \phi_i) \tag{5.12}$$

is assembled, which is once again the linear combination of pre-assembled matrices; the cost is $\mathcal{O}(Q_a N^2)$. To compute the right-hand side of the error equation, the following computations are performed:

$$\begin{aligned} G_k(\phi_i) &= \sum_{j=1}^{N} a_k(\xi_j, \phi_i)[\boldsymbol{c}_N]_j \\ \widehat{\boldsymbol{b}}_N(\boldsymbol{\mu}) &= \sum_{m=1}^{Q_f} \theta^F_m(\boldsymbol{\mu}) F_m(\phi_i) - \sum_{k=1}^{Q_a} \theta^a_k(\boldsymbol{\mu}) G_k(\phi_i). \end{aligned} \tag{5.13}$$

This requires $\mathcal{O}(Q_a N^2)$ operations to perform the matrix vector multiplications for G_k, and $\mathcal{O}\left((Q_a + Q_f)N\right)$ cost to form the necessary linear combination of vectors. Finally, the solution of this system is also $\mathcal{O}(N^3)$, which determines the RB error coefficients $\widehat{\boldsymbol{c}}_N \in \mathbb{R}^N$, that is, $\hat{e}^N = \sum_{n=1}^N [\hat{c}_N]_n \phi_n$.

To form the error bound, the quantities $\|\hat{e}^N_{\boldsymbol{\mu}}\|_X$ and $\|\rho^N_{\boldsymbol{\mu}}\|_Y$ must be computed. Since the error basis functions $\{\phi_j\}_{j=1}^N$ were orthonormalized in the (computable) X-norm, the first quantity is easily computed by

$$\|\hat{e}^N_{\boldsymbol{\mu}}\|_X = \sqrt{\widehat{\boldsymbol{c}}_N^T \widehat{\boldsymbol{c}}_N}. \tag{5.14}$$

To compute the norm of the auxiliary residual, we compute

$$\begin{aligned}
\|\rho^N_{\boldsymbol{\mu}}\|_Y^2 &= \|f_{\boldsymbol{\mu}} - \mathcal{L}_{\boldsymbol{\mu}} u^N_{\boldsymbol{\mu}} - \mathcal{L}_{\boldsymbol{\mu}} \hat{e}^N_{\boldsymbol{\mu}}\|_Y^2 \\
&= (f_{\boldsymbol{\mu}}, f_{\boldsymbol{\mu}})_Y + \left[(\mathcal{L}_{\boldsymbol{\mu}} u^N_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} u^N_{\boldsymbol{\mu}})_Y - (f_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} u^N_{\boldsymbol{\mu}})_Y\right] \\
&\quad + \left[(\mathcal{L}_{\boldsymbol{\mu}} \hat{e}^N_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} \hat{e}^N_{\boldsymbol{\mu}})_Y + (\mathcal{L}_{\boldsymbol{\mu}} u^N_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} \hat{e}^N_{\boldsymbol{\mu}})_Y - (f_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} \hat{e}^N_{\boldsymbol{\mu}})_Y\right] \\
&\quad - (f_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} u^N_{\boldsymbol{\mu}})_Y - \left[(f_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} \hat{e}^N_{\boldsymbol{\mu}})_Y - (\mathcal{L}_{\boldsymbol{\mu}} u^N_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} \hat{e}^N_{\boldsymbol{\mu}})_Y\right] \\
&= (f_{\boldsymbol{\mu}}, f_{\boldsymbol{\mu}})_Y - (f_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} u^N_{\boldsymbol{\mu}})_Y - \left[(f_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} \hat{e}^N_{\boldsymbol{\mu}})_Y - (\mathcal{L}_{\boldsymbol{\mu}} u^N_{\boldsymbol{\mu}}, \mathcal{L}_{\boldsymbol{\mu}} \hat{e}^N_{\boldsymbol{\mu}})_Y\right] \\
&= (f_{\boldsymbol{\mu}}, f_{\boldsymbol{\mu}})_Y - \boldsymbol{b}_N^T \boldsymbol{c}_N - \widehat{\boldsymbol{b}}^T \widehat{\boldsymbol{c}}_N.
\end{aligned} \tag{5.15}$$

Thus, the auxiliary residual is computed through inner products between vectors in $\mathbb{R}^N$ and the computation of $(f_{\boldsymbol{\mu}}, f_{\boldsymbol{\mu}})_Y$. For LSFEM, the Y-norm usually corresponds to the L^2-inner product, the affine parametric dependence of the problem ensures the quantity $(f_{\boldsymbol{\mu}}, f_{\boldsymbol{\mu}})_Y$ is computed at a cost independent of the problem size, either analytically or through quadrature.

If a method such as SCM is used to compute $\alpha_{\text{LB}}(\boldsymbol{\mu})$ at a cost independent of problem size, every necessary computation of the online stage is performed at a cost that is independent of the size of the high-fidelity problem. Finally, the error estimate is computed via

$$M^N(\boldsymbol{\mu}) = \|\hat{e}^N_{\boldsymbol{\mu}}\|_X + \frac{\|\rho^N_{\boldsymbol{\mu}}\|_Y}{\sqrt{\alpha_{\text{LB}}(\boldsymbol{\mu})}} \tag{5.16}$$

These steps are collected in Algorithm 5.2.

5.3. The Coercivity Constant and the Rigor of the Error Estimate. We make use of a lower bound α_{LB} in Algorithm 5.1 to the coercivity constant α. In practice, this lower bound is approximated via the SCM, which computes a lower bound to the discrete coercivity constant $\alpha^h(\boldsymbol{\mu})$. We now examine the implication of this approximation.

The coercivity constant $\alpha(\boldsymbol{\mu})$ is the infimum of a Rayleigh Quotient

$$\alpha(\boldsymbol{\mu}) = \inf_{u \in X} \frac{\|\mathcal{L}_{\boldsymbol{\mu}} u\|_Y^2}{\|u\|_X^2}. \tag{5.17}$$

The discrete coercivity constant $\alpha^h(\boldsymbol{\mu})$ is the infimum of the same Rayleigh Quotient, but u is restricted to be in the finite-dimensional approximation space:

$$\alpha^h(\boldsymbol{\mu}) = \inf_{u^h \in X^h} \frac{\|\mathcal{L}_{\boldsymbol{\mu}} u^h\|_Y^2}{\|u^h\|_X^2}. \tag{5.18}$$

Algorithm 5.2 Least Squares Reduced Basis Online Algorithm

Input: parameter $\boldsymbol{\mu} \in \mathbb{R}^P$
Assemble RB system: $A_N(\boldsymbol{\mu})$ and $\boldsymbol{b}_N(\boldsymbol{\mu})$ ▷ (5.11)
Solve RB system $A_N(\boldsymbol{\mu})\boldsymbol{c}_N = \boldsymbol{b}_N(\boldsymbol{\mu})$
Assemble RB error system: $\widehat{A}_N(\boldsymbol{\mu})$, $G_k(\phi_i)$ and $\widehat{\boldsymbol{b}}_N(\boldsymbol{\mu})$ ▷ (5.12),(5.13)
Solve RB error system $\widehat{A}_N(\boldsymbol{\mu})\widehat{\boldsymbol{c}}_N = \widehat{\boldsymbol{b}}_N(\boldsymbol{\mu})$
Compute error norm $\|\hat{e}^N_{\boldsymbol{\mu}}\|_X = \sqrt{\widehat{\boldsymbol{c}}_N^T \widehat{\boldsymbol{c}}_N}$ ▷ (5.14)
Compute auxiliary residual $\|\rho^N_{\boldsymbol{\mu}}\|^2_Y = (f_{\boldsymbol{\mu}}, f_{\boldsymbol{\mu}})_Y - \boldsymbol{b}_N^T \boldsymbol{c}_N - \widehat{\boldsymbol{b}}^T \widehat{\boldsymbol{c}}_N$ ▷ (5.15)
Compute $\alpha_{\mathrm{LB}}(\boldsymbol{\mu})$ (e.g. through online phase of SCM)
Compute error estimate $M^N(\boldsymbol{\mu}) = \|\hat{e}^N_{\boldsymbol{\mu}}\|_X + \frac{\|\rho^N_{\boldsymbol{\mu}}\|_Y}{\sqrt{\alpha_{\mathrm{LB}}(\boldsymbol{\mu})}}$ ▷ (5.16)

In a conforming discretization, since $X^h \subset X$, it follows that $\alpha(\boldsymbol{\mu}) \leq \alpha^h(\boldsymbol{\mu})$. Thus there is a possibility that

$$\alpha(\boldsymbol{\mu}) < \alpha_{\mathrm{LB}}(\boldsymbol{\mu}) \leq \alpha^h(\boldsymbol{\mu}). \tag{5.19}$$

This leads to a potential underestimation of the error in $M^h(\boldsymbol{\mu})$ or $M^N(\boldsymbol{\mu})$. However, extensive research has been done with respect to the convergence of finite element approximations of eigenvalue problems; see [7] for a thorough overview.

If the problem

$$(\mathcal{L}_{\boldsymbol{\mu}} u, \mathcal{L}_{\boldsymbol{\mu}} v)_Y = (f, v)_X, \quad \forall v \in X, \tag{5.20}$$

has a *compact* solution operator $T_{\boldsymbol{\mu}} : X \to X$, where $T_{\boldsymbol{\mu}} f = u$ is the solution to the variational problem (5.20), then the error $\alpha^h(\boldsymbol{\mu}) - \alpha_(\boldsymbol{\mu})$ is bounded by the *square* of the approximation error of the finite dimensional space X^h. Thus, if a LSFEM with order of convergence r is used, we would expect

$$\alpha^h(\boldsymbol{\mu}) \leq \alpha(\boldsymbol{\mu}) + \mathcal{O}(h^{2r}), \tag{5.21}$$

where h is the mesh size. That is, the error in the approximation of the coercivity constant is much lower than the error in the LSFEM solution. Thus, in the worst case, a non-rigorous lower bound α_{LB} computed by the successive constraint method would still be bounded by $\alpha(\boldsymbol{\mu}) + \mathcal{O}(h^{2r})$ and thus for sufficiently small h,

$$\frac{1}{\sqrt{\alpha(\boldsymbol{\mu})}} \leq \frac{1}{\sqrt{\alpha_{\mathrm{LB}}(\boldsymbol{\mu}) - \mathcal{O}(h^{2r})}} = \frac{1}{\sqrt{\alpha_{\mathrm{LB}}(\boldsymbol{\mu})}}(1 + \mathcal{O}(h^{2r})). \tag{5.22}$$

If this estimate holds, the lower bound built by SCM is asymptotically rigorous, and so is the corresponding error estimate $M^N(\boldsymbol{\mu})$.

Showing the compactness of the solution operator of (5.20) will be the focus of future research, but numerical experiments have shown evidence of the higher order convergence rate. In Figure 2, convergence to the exact coercivity constant of an LSFEM applied to the ordinary differential equation $-u'' = f$ with homogeneous Dirichlet boundary conditions is shown. In this case, the first-order reformulation leads to the operator $\mathcal{L} : H^1(\Omega) \times H^1_0(\Omega) \to L^2(\Omega) \times L^2(\Omega)$ defined as

$$\mathcal{L}\left[(q, u)\right] = \begin{pmatrix} q + u' \\ q' \end{pmatrix}, \tag{5.23}$$

with coercivity constant $\alpha = 1 - (1 + \sqrt{1 + 4\pi^2})/2(1 + \pi^2)$. Using piecewise linear finite elements with 1st order convergence, we see the expected 2nd order convergence of the discrete coercivity constant.

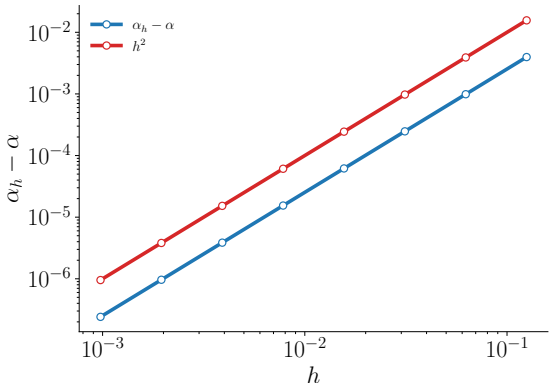


Fig. 2: Convergence of the discrete coercivity constant for the operator defined in (5.23). A piecewise linear finite element space with mesh spacing h is used.

6. Numerical Evidence. In this section we present numerical numerical evidence in support of the LSFEM-RB method introduced in § 5. A single parameter study is given in § 6.1 in order to detail the bounds on the error, while a three-parameter study is discussed in § 6.2. Finally, in § 6.3, the method is applied to an elasticity problem to highlight robustness. The software library Firedrake [46] is used in the following tests. Moreover, it is easy to check that the numerical examples considered below are affinely parametrized in the least-squares setting.

6.1. Thermal Block — 1 Parameter. We first apply the LSFEM-RB framework to a standard test problem in the reduced basis community, the "thermal block" problem [26, 48, 50]. The governing partial differential equation is a variable coefficient Poisson problem:

$$\begin{aligned} -\nabla \cdot \kappa(\boldsymbol{x}) \nabla u &= 0 \qquad \text{in } \Omega, \\ u &= 0 \qquad \text{on } \Gamma_D, \\ \kappa(\boldsymbol{x}) \nabla u \cdot \boldsymbol{n} &= g, \qquad \text{on } \partial\Omega \setminus \Gamma_D. \end{aligned} \tag{6.1}$$

Here, Ω is the unit square, $\Gamma_D = \{(x, y) \in \partial\Omega \mid y = 1\}$, $g(x, y)$ is a function satisfying $g(0, y) = g(1, y) = 0$ and $g(x, 0) = 1$, and $\kappa(\boldsymbol{x})$ is a piecewise constant function taking two different values in subdomains Ω_1, Ω_2; see Figure 3. Specifically,

$$\kappa(\boldsymbol{x}) = \begin{cases} \mu & \boldsymbol{x} \in \Omega_1 \\ 1 & \boldsymbol{x} \in \Omega_2, \end{cases} \tag{6.2}$$

with $\mu \in [10^{-1}, 10^{1}]$.

By introducing a constant lifting function $\boldsymbol{q}_\ell = (0, -1)^T$, and defining the flux variable

$$\boldsymbol{q} = -\kappa \nabla u + \boldsymbol{q}_\ell, \tag{6.3}$$

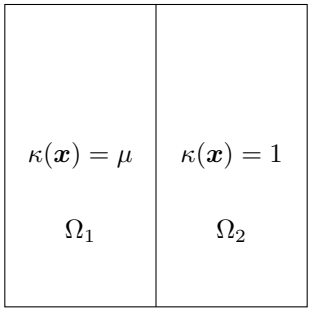


Fig. 3: Variable Poisson problem with conductivity in two subdomains.

the following equivalent first order system is obtained:

$$\begin{aligned} \kappa^{-1/2}\boldsymbol{q} + \kappa^{1/2}\nabla u &= \kappa^{-1/2}\boldsymbol{q}_\ell && \text{in } \Omega, \\ \nabla \cdot \boldsymbol{q} &= 0 && \text{in } \Omega, \\ u &= 0 && \text{on } \Gamma_D, \\ \boldsymbol{q} \cdot \boldsymbol{n} &= 0 && \text{on } \partial\Omega \setminus \Gamma_D. \end{aligned} \tag{6.4}$$

The first order system (6.4) defines an operator $\mathcal{L}_{\boldsymbol{\mu}}$ with domain $X \subset H(\mathrm{div}) \times H^1(\Omega)$ and range $Y = (L^2(\Omega))^2 \times L^2(\Omega)$; here X is the subspace of functions that satisfy the homogeneous boundary conditions. It is shown in [6, 42] that the resulting operator $\mathcal{L}_{\boldsymbol{\mu}}$ satisfies (2.3) — i.e., is continuous and has a bounded inverse — with respect to the $H(\mathrm{div}) \times H^1$ norm. Using this norm on X leads to a well-posed problem and the applicability of the error estimate (2.13).

We compute an approximation using the subspace $X^h = (\mathrm{RT}_0) \times P_1$, approximating $\boldsymbol{q}$ by the lowest order Raviart-Thomas space [47] and u by piecewise linear polynomials. The approximations $\boldsymbol{q}^h$ and u^h are computed on a mesh corresponding to 1,016 degrees of freedom.

The reduced basis is constructed using a sample of 50 logarithmically spaced samples $\mu \in [0.1, 10.]$. The auxiliary error equation is solved on the same mesh using $(\mathrm{RT}_1) \times P_2$ elements, corresponding to 3,556 degrees of freedom.

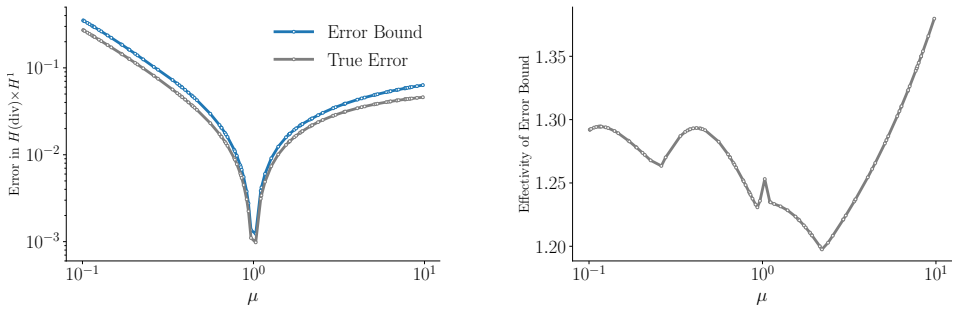
The offline algorithm consists of two stages: the offline SCM portion and the construction of the reduced basis. The SCM requires the solution of 11 generalized eigenvalue problems of size $1,016 \times 1,016$. For the construction of the reduced basis, the algorithm terminates after computing only $N = 3$ full-order solutions are required. The final tolerance upon termination is approximately $\delta \approx 0.3984$. Thus, the error estimate is guaranteed to satisfy the effectivity bound

$$\frac{M^N(\mu)}{\|e_\mu^N\|_X} \leq \frac{1+\delta}{1-\delta} \approx 2.3244. \tag{6.5}$$

Thus, our error bound overestimates the true error by at worst a factor of approximately 2.3244.

To test the reduced basis, we generate 100 randomly sampled parameter values $\mu \in [0.1, 10.0]$ and compute a reference solution using $(\mathrm{RT}_2) \times P_3$ elements after performing two uniform mesh refinements. This corresponds to 121,920 degrees of freedom. We then compute the reduced basis approximation for these parameter

values and the corresponding RB error estimate. For each parameter value, the lower bound for the coercivity constant is computed through the online SCM algorithm; the resulting linear program has 3 variables and 16 inequality constraints. The reduced basis solution and approximate error requires the solution of two 3×3 linear systems. The true error and the corresponding error estimate are shown for the test set in Figure 4a. The error bound is rigorous and resolves the difference in error throughout the parameter domain.



(a) (b)

Fig. 4: (a) The error between the RB solution and the reference solution, along with the corresponding RB error bound over the testing set of parameter values of μ. (b) The effectivity of the RB error estimate in (6.5) over the testing set of parameter values.

In Figure 4b, we plot the effectivity of the error estimate over the same testing set of data. The error bound overestimates the error by a small factor, less than 1.40, which outperforms the effectivity bound in (6.5).

Finally, we plot the run-time for each new parameter value encountered in the online phase for both the reduced basis method, and solutions computed using only full-order solutions in Figure 5.[2] The reduced basis method incurs an offline cost of approximately 4 seconds; thus, for 16 parameter values or fewer, the reduced basis approach is more expensive. However, because the size of the linear system is only 3×3, total run-time grows extremely slowly for additional parameters. Thus after 17 parameter values, the reduced basis approach becomes more computationally efficient, with rapidly increasing computational gains as the number of online parameters grows.

6.2. Thermal Block — 3 Parameters. We repeat the same variable coefficient Poisson problem, now with four subdomains, and consequently, three different parameter values $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)^T$.

The flux reaches a singularity $(1/2, 1/2)$, where all four subdomains meet, which necessitates a finer grid and requires computing the auxiliary error equation by performing a mesh refinement in addition to the increase in polynomial order. Our approximation is again computed on $X^h = (\mathrm{RT}_0) \times P_1$, with 1,456 degrees of freedom. The auxiliary error equation is computed after one mesh refinement using $(\mathrm{RT}_1) \times P_2$ elements, which corresponds to 20,384 degrees of freedom.

[2]The compute node uses two Intel Xeon E5-2630 processors.

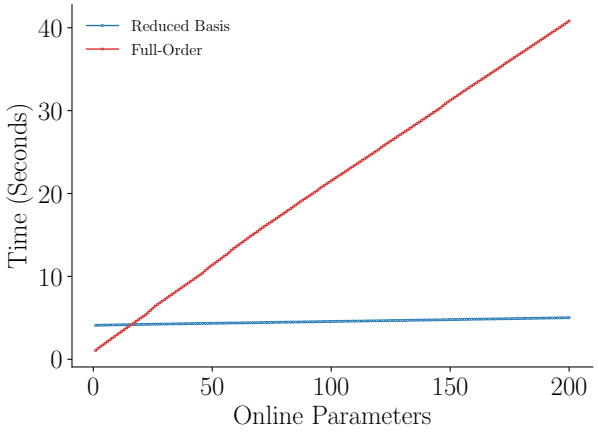


Fig. 5: The total run-time (offline and online) of the reduced basis method for each new parameter value encountered in the online stage compared to the use of full-order models. The reduced basis method incurs an offline cost of approximately 4 seconds, but quickly becomes an extremely efficient alternative over the use of full-order models.

$\kappa(\boldsymbol{x}) = \mu_3$	$\kappa(\boldsymbol{x}) = 1$
$\kappa(\boldsymbol{x}) = \mu_1$	$\kappa(\boldsymbol{x}) = \mu_2$

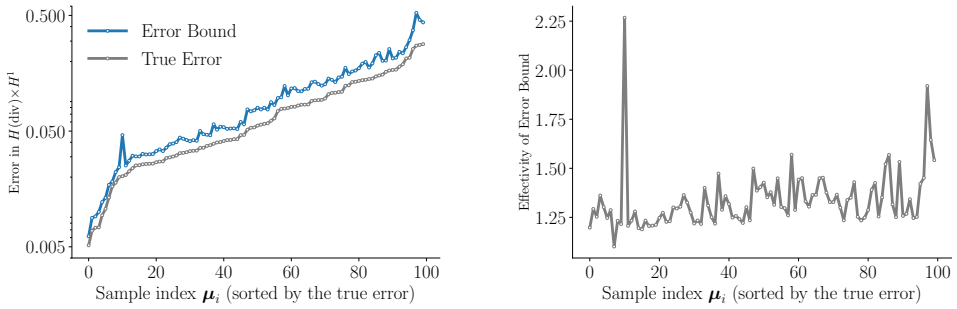
Fig. 6: Conductivity for the variable coefficient Poisson equation with four subdomains. The parameter μ takes values in the interval $[5^{-1}, 5^1]$.

The reduced basis is constructed using a sample of 75 randomly generated samples $\boldsymbol{\mu} \in [0.2, 5.]^3$ using Latin hypercube sampling. We also include the vertices of the parameter domain cube. The SCM method requires the solution of 45 generalized eigenvalue problems of size $1,456 \times 1,456$, and the reduced basis construction terminates after the computation of $N = 13$ basis functions with a final tolerance of $\delta \approx 0.7557$. Thus, the error estimate is guaranteed to satisfy the effectivity bound

$$\frac{M^N(\mu)}{\|e_\mu^N\|_X} \leq \frac{1+\delta}{1-\delta} \approx 7.1877. \tag{6.6}$$

To test the reduced basis, we generate 100 randomly sampled parameter values $\mu \in [0.2, 5.0]^3$ using Latin hypercube sampling and compute a reference solution using $(\mathrm{RT}_2) \times P_3$ elements after performing two uniform mesh refinements. This corresponds to 174,720 degrees of freedom. We then compute the reduced basis approximation

for these parameter values and the corresponding RB error estimate. For each parameter, $\alpha_{\text{LB}}(\boldsymbol{\mu})$ is computed through the online SCM algorithm; the resulting linear program has 7 variables and 39 inequality constraints. The reduced basis solution and approximate error requires the solution of two 13×13 linear systems. The true error and the corresponding error estimate are shown for the test set in Figure 7a. Once again, we see a rigorous upper bound of the error over the testing set.



(a) (b)

Fig. 7: (a) The error between the RB solution and the reference solution, along with the corresponding RB error bound over the testing set of parameter values of μ for the three parameter thermal block. (b) The effectivity of the RB error estimate in (6.5) over the testing set of parameter values in the three parameter thermal block.

Plotting the effectivity of the parameter set in Figure 7b, we see that the effectivity is bounded by 1.5 over much of the testing set, and the error bound overestimates the true error by no more than a factor < 2.4. In this case the guaranteed effectivity bound (6.6) is a slightly pessimistic prediction on the tightness of the error bound.

We also examine the convergence of the reduced solution as the dimension of the basis increases. We choose the parameter value in the testing set with the largest effectivity, i.e. the one corresponding to the largest over-estimation of the error via the reduced basis method with $N = 13$. In this case, $\boldsymbol{\mu} \approx (.223, .244, .746)$ and the error is overestimated by a factor of approximately 2.26. The true error and corresponding error estimate is shown in Figure 8 as a function of basis dimension. The error estimate remains rigorous and reliable for all sizes of basis; the worst effectivity is approximately 3.76 corresponding to $N = 3$.

In terms of run-time, the online algorithm takes considerably longer than the 1 parameter thermal block problem; approximately 43.5 seconds. However, the average run-time per parameter for the RB solution and error estimate in the online stage is only around 5.4×10^{-3} seconds. In comparison, the full-order solution and error estimate takes 1.4 seconds per parameter on average. The speed-up is over 250 times faster and the RB approach is more efficient for 29 or more parameters in the online stage.

6.3. Linear Elasticity. For this experiment we consider linear elasticity and the model problem originating from [45]. The setup consists of a two-dimensional plate with a circular whole at the center. Given the symmetry of the problem we consider only the upper right quarter for Ω as in Figure 9a.

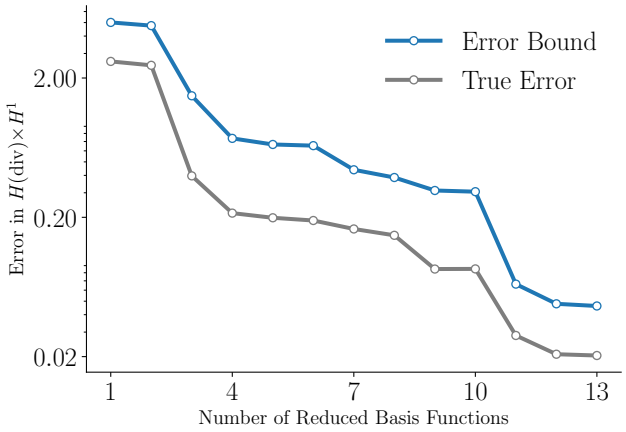
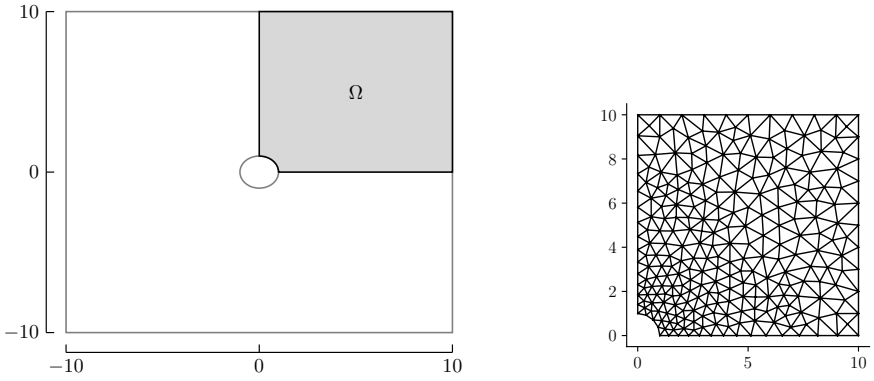


Fig. 8: Convergence of the error bound and true error for increasing RB basis dimension N. The parameter value is $\boldsymbol{\mu} \approx (.223, .244, .746)$, corresponding to the largest over-estimation of the error.



(a) Domain Ω.

(b) Mesh with 466 cells.

Fig. 9: Domain and mesh for the elasticity problem.

We denote the material properties as E (Young's Modulus) and ν (Poisson's ratio), which are related through the Lamé constants

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}, \qquad \mu = \frac{E}{2(1+\nu)}. \tag{6.7}$$

Then let $\boldsymbol{u} = [u_1, u_2]^T$ be the displacement of the plate, and let $\boldsymbol{\sigma}$ be the 2×2 stress tensor. Using a substitution $\boldsymbol{\sigma} \leftarrow \frac{1}{\mu}\boldsymbol{\sigma}$, leads to a change in units so that $\mu = 1$. With

this we arrive at the following first-order system of PDEs, following [14]:

$$\begin{aligned} \mathcal{A}\boldsymbol{\sigma} - \boldsymbol{\epsilon}(\boldsymbol{u}) &= \mathbf{0}, \\ \nabla \cdot \boldsymbol{\sigma} &= \mathbf{0}. \end{aligned} \tag{6.8}$$

Here, we assumed no body forces, the divergence of a tensor is taken row-wise, and the operators $\mathcal{A}$ and $\boldsymbol{\epsilon}$ are defined as:

$$\begin{aligned} \mathcal{A}\boldsymbol{\sigma} &= \frac{1}{2}\left(\boldsymbol{\sigma} - \frac{\lambda}{2(\lambda+1)}(\mathrm{tr}\boldsymbol{\sigma})\boldsymbol{I}\right) = \frac{1}{2}\left(\boldsymbol{\sigma} - \nu(\mathrm{tr}\boldsymbol{\sigma})\boldsymbol{I}\right) \\ \epsilon(\boldsymbol{u}) &= \frac{1}{2}\left(\nabla\boldsymbol{u} + \nabla\boldsymbol{u}^T\right). \end{aligned} \tag{6.9}$$

We apply a (scaled by μ) traction force via the boundary condition $\boldsymbol{\sigma}\boldsymbol{n} = K\boldsymbol{n}$, along the top boundary $y = 10$. To enforce this inhomogenous boundary condition, a lifting function $\boldsymbol{\sigma}_\ell$ is introduced that satisfies this condition.

The parameters for this problem are now of the form $\boldsymbol{\mu} = [\mu_1, \mu_2]^T = [\nu, K]^T$. We restrict Poisson's ratio ν to values in $[0.1, 0.5]$, since 0.5 corresponds to an incompressible material. In addition, we limit the scaled traction coefficient K to the interval $[-0.25, 0.25]$.

Each row of $\boldsymbol{\sigma}$ is viewed as a two-dimensional vector, and we define an operator $\mathcal{L}_{\boldsymbol{\mu}} = \mathcal{L}_\nu$ that maps $\boldsymbol{U} = [\boldsymbol{\sigma}, \boldsymbol{u}] \in X \subset \left[H(\mathrm{div}; \Omega)\right]^2 \times \left[H^1(\Omega)\right]^2$ into $Y = \left[L^2(\Omega)\right]^{2\times 2} \times \left[L^2(\Omega)\right]^2$:

$$\mathcal{L}_\nu \boldsymbol{U} = \begin{pmatrix} \mathcal{A} & -\epsilon \\ \nabla\cdot & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\sigma} \\ \boldsymbol{u} \end{pmatrix} = \begin{pmatrix} -\mathcal{A}\boldsymbol{\sigma}_\ell \\ -\nabla\cdot\boldsymbol{\sigma}_\ell \end{pmatrix} \tag{6.10}$$

Here, X is the subspace of functions that satisfy the corresponding homogeneous boundary conditions. This form of $\mathcal{L}_\nu$ satisfies (2.3) (see [14]) with respect to the norm

$$\|(\boldsymbol{\tau}, \boldsymbol{v})\|_X^2 = \|\boldsymbol{\epsilon}(\boldsymbol{v})\|_0^2 + \|\boldsymbol{\tau}\|_0^2 + \|\nabla\cdot\boldsymbol{\tau}\|_0^2, \tag{6.11}$$

where $\|\cdot\|_0$ is the $L^2(\Omega)$ norm for vector or tensor valued functions, depending on context.

We compute a discrete approximation using the subspace $X^h = (\mathrm{RT}_0)^2 \times (\mathrm{P}_1)^2$ — i.e., approximate the rows of the stress tensor $\boldsymbol{\sigma}$ by functions in the lowest-order Raviart-Thomas space [47], and the components of the displacement $\boldsymbol{u}$ by piecewise linear polynomials. For the mesh in Figure 9b, this corresponds to 1,970 degrees of freedom.

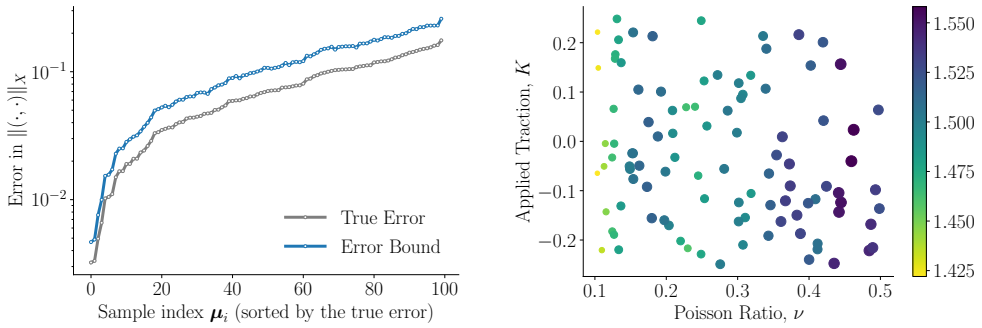
For the reference solutions $\boldsymbol{\sigma}_{\boldsymbol{\mu}}$ and $\boldsymbol{u}_{\boldsymbol{\mu}}$, we perform one refinement on the original mesh, and approximate the solution by functions in $(\mathrm{RT}_2)^2 \times (\mathrm{P}_3)^2$, which corresponds to 56,546 degrees of freedom. We compute the lower bound to the coercivity constant via the SCM method, with a tolerance of 30%. Since the operator $\mathcal{L}_\nu$ only depends on Poisson's ratio, the SCM method is performed using 50 uniformly sample values of $\nu \in [0.1, 0.5]$.

For the basis construction, $\mathcal{D}_{\mathrm{train}}$ consists of a 10×10 uniform grid sampling of $(\nu, K) \in [0.1, 0.5] \times [-0.25, 0.25]$. Algorithm 5.1 terminates after computing $N = 5$ basis functions with a final tolerance of $\delta \approx 0.6445$. That is, all reduced-order solutions $\boldsymbol{\sigma}_{\boldsymbol{\mu}}^N$, $\boldsymbol{u}_{\boldsymbol{\mu}}^N$ corresponding to parameters in the sampled grid satisfy

$$\frac{\|\rho_{\boldsymbol{\mu}}^N\|_0}{\sqrt{\alpha_{\mathrm{LB}}(\boldsymbol{\mu})}\|\hat{e}_{\boldsymbol{\mu}}^N\|_X} \leq \delta \approx 0.6445$$
$$\frac{M^N(\boldsymbol{\mu})}{\|e_{\boldsymbol{\mu}}^N\|_X} \leq \frac{1+\delta}{1-\delta} \approx 4.6266 \tag{6.12}$$

so that our error bound overestimates the true error by at most a factor of 4.6266×.

To test the reduced basis, we generate 100 randomly sampled (ν, K) pairs in $[0.1, 0.5]\times[-0.25, 0.25]$ that were not involved in the basis construction. In Figure 10a, we see that the error bound generated by the reduced basis approximation is a rigorous bound for all parameters in the testing set.



(a) (b)

Fig. 10: (a) Error over 100 samples of (ν, K), computed with respect to a high-order representation of smooth solution, labeled "True Error". The error bounds generated by reduced basis solution is labeled as "Error Bound". Note: The x-axis corresponds to the indices of the test parameters, which are ordered by magnitude of the true error. (b) The effectivity ratios $\frac{M^N(\boldsymbol{\mu})}{\|e_{\boldsymbol{\mu}}^N\|_X}$ over the test parameter set.

We next examine the bound on the effectivity ratio in (6.12), which is again pessimistic. Indeed, the mean effectivity over the test set is approximately 1.501 and no error bound has an effectivity larger than 1.558, as shown in Figure 10b.

7. Conclusions and Future directions. In this paper we have introduced a reduced basis method for parametrized elliptic partial differential equations using least-squares finite element methods. We demonstrated that the first-order system formulation provides an opportunity to construct a rigorous error bound on the *exact* solution by solving an auxiliary error problem. This is in contrast to standard RB approaches that estimate the error with respect to a solution from a fixed finite-dimensional subspace. Rigorous bounds on the effectivity of this estimate have also been established when the auxiliary equation properly resolves the error.

The least-squares finite element reduced basis method is also applicable to bases constructed via POD. In the offline stage, the effectivity of the error bound no longer guides the sampling of the parameter domain since the POD algorithm relies on the decay of the eigenvalues to form the basis. However, the decay of the eigenvalues does not give a quantitative bound on the actual error of reduced basis approximations.

The error estimate and the bound on the effectivity developed in this article may be used after the POD basis is formed to give an indication of whether the basis was truncated too soon. This in turn should provide guidance in determining the number of basis functions needed to produce sufficiently accurate reduced basis solutions.

From the numerical experiments, we see that the bound on the effectivity, while not sharp, is still accurate. Since the error of the RB solution is estimated with respect to the true solution, there may be regions of the parameter domain that require much finer mesh resolution or polynomial orders. Using this reduced basis method as a guide to partitioning the parameter domain into separate reduced order models has the potential to increase accuracy and develop sharper effectivity bounds.

In many cases, an output or quantity of interest $Q(u_{\boldsymbol{\mu}})$ is of more interest than the solution itself. Future work should extend the least-squares finite element reduced basis method to these situations by developing computable bounds on the error $|Q(u_{\boldsymbol{\mu}}) - Q(u^N_{\boldsymbol{\mu}})|$.

Finally, least-squares finite element methods are not the only variational method that re-formulates a PDE into a first-order system. An investigation of other such methods, e.g., mixed Galerkin finite element methods, would make the results more broadly applicable.

REFERENCES

[1] J. Adler and P. S. Vassilevski, *Error analysis for constrained first-order system least-squares finite-element methods*, SIAM Journal on Scientific Computing, 36 (2014), pp. A1071–A1088.

[2] T. Atherton and J. Adler, *Competition of elasticity and flexoelectricity for bistable alignment of nematic liquid crystals on patterned substrates*, Physical Review E, 86 (2012), p. 040701.

[3] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera, *An 'empirical interpolation'method: application to efficient reduced-basis discretization of partial differential equations*, Comptes Rendus Mathematique, 339 (2004), pp. 667–672.

[4] P. Bochev and J. Choi, *Improved least-squares error estimates for scalar hyperbolic problems*, Computational Methods in Applied Mathematics, 1 (2001), pp. 115–124.

[5] P. Bochev, T. A. Manteuffel, and S. F. McCormick, *Analysis of velocity-flux least-squares principles for the navier–stokes equations: Part ii*, SIAM Journal on Numerical Analysis, 36 (1999), pp. 1125–1144.

[6] P. B. Bochev and M. D. Gunzburger, *Least-squares finite element methods*, vol. 166, Springer Science & Business Media, 2009.

[7] D. Boffi, *Finite element approximation of eigenvalue problems.*, Acta Numer., 19 (2010), pp. 1–120.

[8] D. Boffi, F. Brezzi, M. Fortin, et al., *Mixed finite element methods and applications*, vol. 44, Springer, 2013.

[9] S. D. Bond, J. H. Chaudhry, E. C. Cyr, and L. N. Olson, *A first-order system least-squares finite element method for the poisson-boltzmann equation*, Journal of computational chemistry, 31 (2010), pp. 1625–1635.

[10] S. Boyaval, *Reduced-basis approach for homogenization beyond the periodic setting*, Multiscale Modeling & Simulation, 7 (2008), pp. 466–494.

[11] D. Braess, *Finite elements: Theory, fast solvers, and applications in solid mechanics*, Cambridge University Press, 2007.

[12] J. Brannick, C. Ketelsen, T. Manteuffel, and S. McCormick, *Least-squares finite element methods for quantum electrodynamics*, SIAM Journal on Scientific Computing, 32 (2010), pp. 398–417.

[13] T. Bui-Thanh, K. Willcox, and O. Ghattas, *Model reduction for large-scale systems with high-dimensional parametric input space*, SIAM Journal on Scientific Computing, 30 (2008), pp. 3270–3288.

[14] Z. Cai and G. Starke, *Least-squares methods for linear elasticity*, SIAM Journal on Numerical Analysis, 42 (2004), pp. 826–842.

[15] K. Carlberg, M. Barone, and H. Antil, *Galerkin v. least-squares petrov–galerkin projection*

in nonlinear model reduction, Journal of Computational Physics, 330 (2017), pp. 693–734.
[16] K. Carlberg, C. Bou-Mosleh, and C. Farhat, *Efficient non-linear model reduction via a least-squares petrov–galerkin projection and compressive tensor approximations*, International Journal for numerical methods in engineering, 86 (2011), pp. 155–181.
[17] J. H. Chaudhry, N. Burch, and D. Estep, *Efficient distribution estimation and uncertainty quantification for elliptic problems on domains with stochastic boundaries*, SIAM/ASA Journal on Uncertainty Quantification, 6 (2018), pp. 1127–1150.
[18] Y. Chen, J. S. Hesthaven, Y. Maday, and J. Rodríguez, *A monotonic evaluation of lower bounds for inf-sup stability constants in the frame of reduced basis approximations*, Comptes Rendus Mathematique, 346 (2008), pp. 1295–1300.
[19] Y. Chen, J. S. Hesthaven, Y. Maday, and J. Rodríguez, *Improved successive constraint method based a posteriori error estimate for reduced basis approximation of 2d maxwell's problem*, ESAIM: Mathematical Modelling and Numerical Analysis, 43 (2009), pp. 1099–1116.
[20] H. De Sterck, T. A. Manteuffel, S. F. McCormick, and L. Olson, *Least-squares finite element methods and algebraic multigrid solvers for linear hyperbolic pdes*, SIAM Journal on Scientific Computing, 26 (2004), pp. 31–54.
[21] H. De Sterck, T. A. Manteuffel, S. F. McCormick, and L. Olson, *Numerical conservation properties of h (div)-conforming least-squares finite element methods for the burgers equation*, SIAM Journal on Scientific Computing, 26 (2005), pp. 1573–1597.
[22] S. Deparis and G. Rozza, *Reduced basis method for multi-parameter-dependent steady navier–stokes equations: applications to natural convection in a cavity*, Journal of Computational Physics, 228 (2009), pp. 4359–4378.
[23] M. A. Dihlmann and B. Haasdonk, *Certified pde-constrained parameter optimization using reduced basis surrogate models for evolution problems*, Computational Optimization and Applications, 60 (2015), pp. 753–787.
[24] M. A. Grepl, N. C. Nguyen, K. Veroy, A. T. Patera, and G. R. Liu, *Certified rapid solution of partial differential equations for real-time parameter estimation and optimization*, in Real-time PDE-constrained optimization, SIAM, 2007, pp. 199–216.
[25] M. A. Grepl and A. T. Patera, *A posteriori error bounds for reduced-basis approximations of parametrized parabolic partial differential equations*, ESAIM: Mathematical Modelling and Numerical Analysis, 39 (2005), pp. 157–181.
[26] B. Haasdonk, *Reduced basis methods for parametrized pdes—a tutorial introduction for stationary and instationary problems*, Model reduction and approximation: theory and algorithms, 15 (2017), p. 65.
[27] M. W. Hess and P. Benner, *Fast evaluation of time–harmonic maxwell's equations using the reduced basis method*, IEEE Transactions on Microwave Theory and Techniques, 61 (2013), pp. 2265–2274.
[28] J. S. Hesthaven, B. Stamm, and S. Zhang, *Efficient greedy algorithms for high-dimensional parameter spaces with applications to empirical interpolation and reduced basis methods*, ESAIM: Mathematical Modelling and Numerical Analysis, 48 (2014), pp. 259–283.
[29] J. J. Heys, E. Lee, T. A. Manteuffel, and S. F. McCormick, *An alternative least-squares formulation of the navier–stokes equations with improved mass conservation*, Journal of Computational Physics, 226 (2007), pp. 994–1006.
[30] D. Huynh, D. Knezevic, Y. Chen, J. S. Hesthaven, and A. Patera, *A natural-norm successive constraint method for inf-sup lower bounds*, Computer Methods in Applied Mechanics and Engineering, 199 (2010), pp. 1963–1975.
[31] D. Huynh and A. Patera, *Reduced basis approximation and a posteriori error estimation for stress intensity factors*, International Journal for Numerical Methods in Engineering, 72 (2007), pp. 1219–1259.
[32] B.-n. Jiang, *The Least-Squares Finite Element Method: Theory and Applications in Computational Fluid Dynamics and Electromagnetics*, Springer Science & Business Media, 2013.
[33] M. Kahlbacher and S. Volkwein, *Galerkin proper orthogonal decomposition methods for parameter dependent elliptic systems*, Discussiones Mathematicae, Differential Inclusions, Control and Optimization, 27 (2007), pp. 95–117.
[34] R. Krause, B. Müller, and G. Starke, *An adaptive least-squares mixed finite element method for the signorini problem*, Numerical Methods for Partial Differential Equations, 33 (2017), pp. 276–289.
[35] C. A. Leibs and T. A. Manteuffel, *Nested iteration and first-order systems least squares for a two-fluid electromagnetic darwin model*, SIAM Journal on Scientific Computing, 37 (2015), pp. S314–S333.
[36] Y. Liang, H. Lee, S. Lim, W. Lin, K. Lee, and C. Wu, *Proper orthogonal decomposition and*

its applications—part i: Theory, Journal of Sound and vibration, 252 (2002), pp. 527–544.
[37] Y. Maday, A. T. Patera, and D. V. Rovas, *A blackbox reduced-basis output bound method for noncoercive linear problems*, Studies in Mathematics and its Applications, 31 (2002), pp. 533–569.
[38] T. A. Manteuffel and K. J. Ressel, *Least-squares finite-element solution of the neutron transport equation in diffusive regimes*, SIAM journal on numerical analysis, 35 (1998), pp. 806–835.
[39] A. Manzoni and F. Negri, *Automatic reduction of pdes defined on domains with variable shape*, in MATHICSE Technical Report, EPFL, 2016.
[40] I. Oliveira and A. Patera, *Reduced-basis techniques for rapid reliable optimization of systems described by affinely parametrized coercive elliptic partial differential equations*, Optimization and Engineering, 8 (2007), pp. 43–65.
[41] L. N. Olson, *Multilevel Least-Squares Finite Element Methods for Hyperbolic PDEs*, PhD thesis, University of Colorado at Boulder, Department of Applied Mathematics, 2003.
[42] A. Pehlivanov, G. Carey, and R. Lazarov, *Least-squares mixed finite elements for second-order elliptic problems*, SIAM Journal on Numerical Analysis, 31 (1994), pp. 1368–1377.
[43] C. Prud'Homme, D. V. Rovas, K. Veroy, L. Machiels, 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 (2001), pp. 70–80.
[44] A. Quarteroni, A. Manzoni, and F. Negri, *Reduced basis methods for partial differential equations: an introduction*, vol. 92, Springer, 2016.
[45] E. Ramm, E. Rank, R. Rannacher, K. Schweizerhof, E. Stein, W. Wendland, G. Wittum, P. Wriggers, and W. Wunderlich, *Error-controlled adaptive finite elements in solid mechanics*, John Wiley & Sons, 2003.
[46] F. Rathgeber, D. A. Ham, L. Mitchell, M. Lange, F. Luporini, A. T. T. Mcrae, G.-T. Bercea, G. R. Markall, and P. H. J. Kelly, *Firedrake: Automating the finite element method by composing abstractions*, ACM Trans. Math. Softw., 43 (2016), https://doi.org/10.1145/2998441, https://doi.org/10.1145/2998441.
[47] P.-A. Raviart and J.-M. Thomas, *Primal hybrid finite element methods for 2nd order elliptic equations*, Mathematics of computation, 31 (1977), pp. 391–413.
[48] G. Rozza, D. B. P. Huynh, and A. T. Patera, *Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations*, Archives of Computational Methods in Engineering, 15 (2007), p. 1.
[49] G. Rozza, D. P. Huynh, and A. Manzoni, *Reduced basis approximation and a posteriori error estimation for stokes flows in parametrized geometries: roles of the inf-sup stability constants*, Numerische Mathematik, 125 (2013), pp. 115–152.
[50] A. Schmidt, D. Wittwar, and B. Haasdonk, *Rigorous and effective a-posteriori error bounds for nonlinear problems—application to rb methods*, Advances in Computational Mathematics, 46 (2020), pp. 1–30.
[51] S. Sen, *Reduced-basis approximation and a posteriori error estimation for many-parameter heat conduction problems*, Numerical Heat Transfer, Part B: Fundamentals, 54 (2008), pp. 369–389.
[52] S. Sen, K. Veroy, D. Huynh, S. Deparis, N. C. Nguyen, and A. T. Patera, *"natural norm" a posteriori error estimators for reduced basis approximations*, Journal of Computational Physics, 217 (2006), pp. 37–62.
[53] L. Q. Tang and T. T. Tsang, *A least-squares finite element method for time-dependent incompressible flows with thermal convection*, International Journal for Numerical Methods in Fluids, 17 (1993), pp. 271–289.
[54] K. Veroy, C. Prud'Homme, D. Rovas, and A. Patera, *A posteriori error bounds for reduced-basis approximation of parametrized noncoercive and nonlinear elliptic partial differential equations*, in 16th AIAA Computational Fluid Dynamics Conference, 2003, p. 3847.
[55] S. Volkwein, *Model reduction using proper orthogonal decomposition*, Lecture Notes, Institute of Mathematics and Scientific Computing, University of Graz. see http://www. uni-graz. at/imawww/volkwein/POD. pdf, 1025 (2011).
[56] D.-p. Yang, *Some least-squares galerkin procedures for first-order time-dependent convection–diffusion system*, Computer methods in applied mechanics and engineering, 180 (1999), pp. 81–95.
[57] M. Yano, *A space-time petrov–galerkin certified reduced basis method: Application to the boussinesq equations*, SIAM Journal on Scientific Computing, 36 (2014), pp. A232–A266.