NORGES TEKNISK-NATURVITENSKAPELIGE
UNIVERSITET

An “hp” Certified Reduced Basis Method for Parametrized Elliptic Partial Differential Equations

by

Jens L. Eftang, Anthony T. Patera, and Einar M. Rønquist

PREPRINT
NUMERICS NO. 10/2009

NORWEGIAN UNIVERSITY OF
SCIENCE AND TECHNOLOGY
TRONDHEIM, NORWAY

This report has URL
`http://www.math.ntnu.no/preprint/numerics/2009/N10-2009.pdf`
Address: Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491 Trondheim, Norway.

AN "hp" CERTIFIED REDUCED BASIS METHOD FOR PARAMETRIZED ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS*

JENS L. EFTANG†, ANTHONY T. PATERA‡, AND EINAR M. RØNQUIST§

Abstract. We present a new "hp" parameter multi-domain certified reduced basis method for rapid and reliable online evaluation of functional outputs associated with parametrized elliptic partial differential equations. We propose a new procedure and attendant theoretical foundations for adaptive partition of the parameter domain into parameter subdomains ("h"-refinement); subsequently, we construct individual standard reduced basis approximation spaces for each subdomain ("p"-refinement). Greedy parameter sampling procedures and *a posteriori* error estimation are the main ingredients of the new algorithm. We present illustrative numerical results for a convection-diffusion problem: the new "hp"-approach is considerably faster (respectively, more costly) than the standard "p"-type reduced basis method in the online (respectively, offline) stage.

Key words. reduced basis; *a posteriori* error estimation; Greedy sampling; h-type; p-type; hp convergence; parameter domain decomposition

AMS subject classifications. 35J25, 65M12, 65N15, 65N15, 65N30

1. Introduction.

The certified reduced basis (RB) method provides a computational framework for rapid and reliable computation of functional outputs associated with parametrized partial differential equations. Given any *input* parameter vector—e.g., geometric factors or material property coefficients—the RB field approximation is constructed as a Galerkin-optimal linear combination of pre-computed "truth" finite element (FE) "snapshots" for judiciously chosen parameters; the RB output approximation is then evaluated as a functional of the RB field approximation. The methodology is originally introduced in [1, 17] and then further analyzed in [18, 19]; for a review of both earlier and more recent contributions, see [20].

For problems in which the field variable varies smoothly with the parameters good RB approximations can be obtained with very few snapshots: the RB approximation converges exponentially fast [7, 20]. Furthermore, rigorous *a posteriori* upper bounds for the error in the RB approximation (with respect to the truth discretization) can be readily developed [20]. Finally, under an assumption on "affine" parameter dependence (perhaps only approximate [4, 10]), both the RB output approximation *and* the associated RB output error bound can be computed very efficiently by an offline-online computational procedure [20]. The RB method is especially attractive in important engineering contexts in which low marginal (online) computational cost is advantageous: "real-time"—such as parameter estimation [15] and optimal control—and "many-query"—such as multiscale [5] or stochastic simulation [6].

The RB approximation space is specifically constructed to provide accurate approximations for any parameter value in a predefined parameter domain. Hence, larger parameter domains typically induce larger RB spaces and greater computational cost. In this paper, we propose a new procedure for adaptive partition ("h"-refinement) of the parameter domain into smaller parameter subdomains: a hierarchical splitting of

*This work has been supported by the Norwegian University of Science and Technology and AFOSR Grant number FA 9550-07-1-0425.

†Department of Mathematical Sciences, NTNU (*eftang@math.ntnu.no*)

‡Department of Mechanical Engineering, MIT (*patera@mit.edu*)

§Department of Mathematical Sciences, NTNU (*ronquist@math.ntnu.no*)

Submitted to SIAM Journal on Scientific Computing, Dec. 14th, 2009.

the parameter (sub)domains based on proximity to judiciously chosen parameter *anchor points* within each subdomain. Subsequently, we construct *individual* standard RB approximation spaces ("p"-refinement) over each subdomain. Greedy sampling procedures and rigorous *a posteriori* error estimation play important roles in both the "h"-type and "p"-type stages of the algorithm.

In this new approach, the RB approximation associated with any new parameter value is, as always, constructed as a linear (Galerkin) combination of snapshots from the parameter (sub)domain in which the parameter value resides. However, we expect the online computational cost of the new approach to be greatly reduced relative to the online cost of the standard RB approach due to the smaller parameter (sub)domains and lower dimensional local RB approximation spaces associated with the "hp" approximation. The method should be particularly effective for problems in which the solution has very different structure in different regions of the parameter domain—problems for which a snapshot from one parameter region may be of limited value for the RB approximation in another parameter region.

The notion of parameter domain refinement within the model order reduction framework is considered in several earlier works. In [2, 3], a reduced-order parameter multi-element "interpolation" procedure is introduced for aeroelasticity problems. The approach [2, 3] and our approach here share a similar error-adaptive domain-decomposition foundation. However, the approach of [2, 3] and the approach described in the current paper are quite different in conception: interpolation on a manifold (in [2, 3]) rather than Galerkin projection (here); parameter domain partition based on a Voronoi diagram rather than a hierarchical tree structure decomposition; heuristic error indicators rather than rigorous error bounds; and less strict rather than strict offline-online segregation. However, our own approach cannot yet treat problems of the complexity considered in [2, 3].

In other related work [11, 21], adaptive *train sample* refinement is considered to render the Greedy parameter sampling procedure more efficient: richer samples are consided only as needed in the Greedy iterations [21] and only where needed in the parameter domain [11]. Our approach invokes a similar technique: we include new points in the train sample within each subdomain at each new level of "h"-refinement; we thus effectively adapt the train sample to "difficult" parameter regions.

In §2 we give the general problem statement along with various entities required throughout the paper. In §3 we review the standard ("p"-type) RB method; in §4 we present the new "h"-type RB method and provide an *a priori* convergence theory for a "zeroth order" approximation in the one-parameter case; in §5 we present the new "hp"-type RB method as a combination of the "p"- and "h"-type methods. In §6 we present numerical results for a convection-diffusion model problem and in particular we compare the computational cost of the new "hp"-approach to the standard method. Finally, we conclude in §7 with some final remarks.

2. Problem Statement.

We shall consider linear, elliptic, second-order equations. We denote the physical domain by $\Omega \subset \mathbb{R}^2$, and we introduce the spaces $L^2(\Omega) = \{v : \int_\Omega v^2 \,\mathrm{d}\Omega < \infty\}$, $H^1(\Omega) = \{v \in L^2(\Omega) : |\nabla v| \in L^2(\Omega)\}$, and $H^1_0(\Omega) = \{v \in H^1(\Omega) : v|_{\partial\Omega} = 0\}$. We further define the space associated with the exact solution (hence ^{e}) $X^e \equiv X^e(\Omega)$ such that $H^1_0(\Omega) \subseteq X^e(\Omega) \subset H^1(\Omega)$. We denote the admissible parameter domain by $\mathcal{D} \subset \mathbb{R}^P$; a point in $\mathcal{D}$ shall be denoted $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_P)$.

For each $\boldsymbol{\mu} \in \mathcal{D}$, $a(\cdot, \cdot; \boldsymbol{\mu})$ is an X^e–coercive and X^e–continuous bilinear form and $f(\cdot; \boldsymbol{\mu})$ is an X^e–bounded linear functional. To accomodate an efficient offline–online computational procedure, we assume that a and f admit affine expansions as

$$a(\cdot,\cdot;\boldsymbol{\mu}) = \sum_{q=1}^{Q_a} a^q(\cdot,\cdot)\Theta_a^q(\boldsymbol{\mu}), \qquad f(\cdot;\boldsymbol{\mu}) = \sum_{q=1}^{Q_f} f^q(\cdot)\Theta_f^q(\boldsymbol{\mu}), \tag{2.1}$$

for modest Q_a and Q_f, where the a^q and f^q are $\boldsymbol{\mu}$-independent continuous bilinear forms and linear functionals, respectively, and the Θ_a^q and Θ_f^q are $\boldsymbol{\mu}$-dependent continuous functions. (The assumption (2.1) can be relaxed with the *empirical interpolation* method [4, 10] for the construction of good affine *approximations* to a and f.) For simplicity, we introduce $Q = \max\{Q_a, Q_f\}$.

The exact problem statement reads: Given any $\boldsymbol{\mu} \in \mathcal{D}$, find $u^e(\boldsymbol{\mu}) \in X^e$ such that

$$a(u^e(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = f(v;\boldsymbol{\mu}), \quad \forall v \in X^e. \tag{2.2}$$

The output of interest can then be evaluated as a functional of the field variable, say $s(\boldsymbol{\mu}) = l(u^e(\boldsymbol{\mu});\boldsymbol{\mu})$ for some X^e–bounded linear functional $l(\cdot;\boldsymbol{\mu})$. In this paper, however, for simplicity of exposition, we consider no particular output(s) of interest; our "*hp*" procedure does not depend on the output functional(s) chosen.

We next introduce a "truth" finite element (FE) space $X \equiv X^{\mathcal{N}}(\Omega) \subset X^e(\Omega)$ of finite dimension $\mathcal{N}$. The truth discretization of (2.2) reads: For any $\boldsymbol{\mu} \in \mathcal{D}$, find $u(\boldsymbol{\mu}) \in X$ such that

$$a(u(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = f(v;\boldsymbol{\mu}), \quad \forall v \in X. \tag{2.3}$$

We assume that X is rich enough that the error between the truth and exact solutions is in practice negligible. The reduced basis approximations will be built upon truth snapshots $u(\boldsymbol{\mu}_n) \approx u^e(\boldsymbol{\mu}_n)$, $1 \le n \le N$, for judiciously chosen $\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_N \in \mathcal{D}$ and the reduced basis error shall be measured with respect to the truth FE approximation.

For any $\boldsymbol{\mu} \in \mathcal{D}$, let $a_s(\cdot,\cdot;\boldsymbol{\mu})$ denote the symmetric part of $a(\cdot,\cdot;\boldsymbol{\mu})$—for all $v, w \in X$, $a_s(w,v;\boldsymbol{\mu}) = \frac{1}{2}(a(w,v;\boldsymbol{\mu}) + a(v,w;\boldsymbol{\mu}))$; further, let $\bar{\boldsymbol{\mu}} \in \mathcal{D}$ denote a fixed *reference parameter*. We then define the parameter-independent X–inner-product and corresponding X-norm as

$$(\cdot,\cdot)_X \equiv a_s(\cdot,\cdot;\bar{\boldsymbol{\mu}}), \qquad \|\cdot\|_X = \sqrt{(\cdot,\cdot)_X}, \tag{2.4}$$

respectively. By our assumptions, $\|\cdot\|_X$ is equivalent to the H^1 norm.

Finally, we introduce for all $\boldsymbol{\mu} \in \mathcal{D}$ the coercivity and continuity constants of $a(\cdot,\cdot;\boldsymbol{\mu})$ with respect to the X-norm,

$$\alpha(\boldsymbol{\mu}) \equiv \inf_{w\in X} \frac{a(w,w;\boldsymbol{\mu})}{\|w\|_X^2}, \qquad \gamma(\boldsymbol{\mu}) \equiv \sup_{v\in X}\sup_{w\in X} \frac{a(v,w;\boldsymbol{\mu})}{\|v\|_X\|w\|_X}, \tag{2.5}$$

respectively. For any particular $\boldsymbol{\mu} \in \mathcal{D}$, we further introduce lower and upper bounds,

$$0 < \alpha_{\mathrm{LB}}(\boldsymbol{\mu}) \le \alpha(\boldsymbol{\mu}), \tag{2.6}$$

$$\infty > \gamma_{\mathrm{UB}}(\boldsymbol{\mu}) \ge \gamma(\boldsymbol{\mu}), \tag{2.7}$$

which shall play a role in our computational procedures. We shall also require lower and upper bounds over $\mathcal{D}$,

$$\underline{\alpha} = \inf_{\boldsymbol{\mu}\in\mathcal{D}} \alpha(\boldsymbol{\mu}), \tag{2.8}$$

$$\overline{\gamma} = \sup_{\boldsymbol{\mu}\in\mathcal{D}} \gamma(\boldsymbol{\mu}), \tag{2.9}$$

for the purposes of our theoretical arguments.

We shall later require the following lemma,

LEMMA 2.1. *Let* $\Theta_a^q : \mathcal{D} \to \mathbb{R}$, $1 \le q \le Q_a$, $\Theta_f^q : \mathcal{D} \to \mathbb{R}$, $1 \le q \le Q_f$, *satisfy Lipschitz conditions*

$$|\Theta_a^q(\boldsymbol{\mu}_1) - \Theta_a^q(\boldsymbol{\mu}_2)| \le C_a|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2|, \quad \forall \boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \in \mathcal{D}, \quad 1 \le q \le Q_a, \tag{2.10}$$

$$|\Theta_f^q(\boldsymbol{\mu}_1) - \Theta_f^q(\boldsymbol{\mu}_2)| \le C_f|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2|, \quad \forall \boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \in \mathcal{D}, \quad 1 \le q \le Q_f. \tag{2.11}$$

Then, given any $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \in \mathcal{D}$, *there exists a positive constant* $\tilde{C}$ *such that*

$$\|u(\boldsymbol{\mu}_1) - u(\boldsymbol{\mu}_2)\|_X \le \tilde{C}|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2|. \tag{2.12}$$

Proof. We have

$$a(u(\boldsymbol{\mu}_1), v; \boldsymbol{\mu}_1) = f(v; \boldsymbol{\mu}_1), \quad \forall v \in X, \tag{2.13}$$

$$a(u(\boldsymbol{\mu}_2), v; \boldsymbol{\mu}_2) = f(v; \boldsymbol{\mu}_2), \quad \forall v \in X. \tag{2.14}$$

By bilinearity of a, we thus have for all $v \in X$,

$$\begin{aligned} a(u(\boldsymbol{\mu}_1) - u(\boldsymbol{\mu}_2), v; \boldsymbol{\mu}_1) = f(v; \boldsymbol{\mu}_1) - f(v; \boldsymbol{\mu}_2) \\ + a(u(\boldsymbol{\mu}_2), v; \boldsymbol{\mu}_2) - a(u(\boldsymbol{\mu}_2), v; \boldsymbol{\mu}_1). \end{aligned} \tag{2.15}$$

We first examine the right-hand side of (2.15).

By the triangle inequality and the affine expansions (2.1) for a and f, we have for all $w, v \in X$ and any $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \in \mathcal{D}$,

$$|a(w, v; \boldsymbol{\mu}_1) - a(w, v; \boldsymbol{\mu}_2)| \le \sum_{q=1}^{Q_a} |a^q(w, v)\big(\Theta_a^q(\boldsymbol{\mu}_1) - \Theta_a^q(\boldsymbol{\mu}_2)\big)|, \tag{2.16}$$

and

$$|f(v; \boldsymbol{\mu}_1) - f(v; \boldsymbol{\mu}_2)| \le \sum_{q=1}^{Q_f} |f^q(v)\big(\Theta_f^q(\boldsymbol{\mu}_1) - \Theta_f^q(\boldsymbol{\mu}_2)\big)|, \tag{2.17}$$

respectively. By our hypothesis (2.10) and (2.11) on Θ_a^q, $1 \le q \le Q_a$ and Θ_f^q, $1 \le q \le Q_f$, respectively, and continuity of a^q, $1 \le q \le Q_a$, and f^q, $1 \le q \le Q_f$, there exist constants $\tilde{c}_1$ and $\tilde{c}_2$ (independent of $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$) such that

$$|a(w, v; \boldsymbol{\mu}_1) - a(w, v; \boldsymbol{\mu}_2)| \le \tilde{c}_1\|v\|_X\|w\|_X|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2|, \tag{2.18}$$

and

$$|f(v; \boldsymbol{\mu}_1) - f(v; \boldsymbol{\mu}_2)| \le \tilde{c}_2\|v\|_X|\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2|. \tag{2.19}$$

Recall that Q_a and Q_f are fixed and finite.

We now let $v = u(\boldsymbol{\mu}_1) - u(\boldsymbol{\mu}_2)$ in (2.15) and deduce from the triangle inequality, (2.18), and (2.19) that

$$\begin{aligned} a(u(\boldsymbol{\mu}_1) - u(\boldsymbol{\mu}_2), u(\boldsymbol{\mu}_1) - u(\boldsymbol{\mu}_2); \boldsymbol{\mu}_1) \\ \le \big(\tilde{c}_1\|u(\boldsymbol{\mu}_2)\|_X + \tilde{c}_2\big)\|u(\boldsymbol{\mu}_1) - u(\boldsymbol{\mu}_2)\|_X|\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1|. \end{aligned} \tag{2.20}$$

By coercivity and the bound (2.8), we get

$$\|u(\boldsymbol{\mu}_1) - u(\boldsymbol{\mu}_2)\|_X \leq \frac{1}{\underline{\alpha}}\big(\tilde{c}_1\|u(\boldsymbol{\mu}_2)\|_X + \tilde{c}_2\big)|\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1|. \tag{2.21}$$

Finally, by the Lax-Milgram Lemma,

$$\|u(\boldsymbol{\mu}_2)\|_X \leq \frac{\|f\|_{X'}}{\underline{\alpha}}, \tag{2.22}$$

(here X' denotes the dual space of X) and we thus obtain the desired result with

$$\tilde{C} = \frac{\tilde{c}_1\|f\|_{X'} + \tilde{c}_2\underline{\alpha}}{\underline{\alpha}^2}. \tag{2.23}$$

(We can develop a constant $\tilde{C}$ that is furthermore independent of $\mathcal{N}$ by replacing the truth entities $\underline{\alpha}$ and $\|f\|_{X'}$ in (2.23) by the corresponding exact entities.) □

3. The "*p*"-type Reduced Basis Method. In the standard RB approach, a single approximation space is enriched with new basis functions until the space is considered sufficiently rich; we shall refer to this approach as the "*p*"-type RB method. The new "*h*"-type and "*hp*"-type methods will borrow and adapt several of the ingredients from the standard approach: *a posteriori* error estimation; greedy parameter sampling; and offline–online computational decoupling of the RB approximation and the truth FE discretization through a construction–evaluation decomposition. Below, we summarize the standard RB approximation with particular emphasis on these key ingredients.

3.1. Approximation. The RB approximation space $X_N \equiv X_N(\Omega) \subset X^{\mathcal{N}}(\Omega)$ is defined in terms of a set of parameter vectors $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_N \in \mathcal{D}$ as

$$X_N = \text{span}\{u(\boldsymbol{\mu}_1), \ldots, u(\boldsymbol{\mu}_1)\}. \tag{3.1}$$

(Note that in practice, the $(\cdot,\cdot)_X$–orthonormal basis for X_N is constructed by a Gram-Schmidt procedure.) The RB approximation reads: Given any $\boldsymbol{\mu} \in \mathcal{D}$, find $u_N(\boldsymbol{\mu}) \in X_N$ such that

$$a(u_N(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = f(v; \boldsymbol{\mu}), \quad \forall v \in X_N. \tag{3.2}$$

Under the assumption that $u(\boldsymbol{\mu})$ depends smoothly on the parameters, we expect that N—the dimension of the RB space—can be chosen much smaller than $\mathcal{N}$—the dimension of the truth space X—for comparable numerical accuracy.

We finally *define* the "order" p of the RB approximation as $p \equiv N^{1/P} - 1$.

3.2. *A Posteriori* **Error Estimation.** We develop here an *a posteriori* X-norm bound for the error in the RB field approximation relative to the corresponding truth approximation.

Given any $\boldsymbol{\mu} \in \mathcal{D}$, we obtain the RB approximation, $u_N(\boldsymbol{\mu})$, from (3.2); we then define for all $v \in X$ the RB residual as

$$r_N(v; \boldsymbol{\mu}) \equiv f(v; \boldsymbol{\mu}) - a(u_N(\boldsymbol{\mu}), v; \boldsymbol{\mu}); \tag{3.3}$$

the *Riesz representation* of the residual, $\mathcal{R}_N(\boldsymbol{\mu}) \in X$, satisfies

$$(\mathcal{R}_N(\boldsymbol{\mu}), v)_X = r_N(v; \boldsymbol{\mu}), \quad \forall v \in X. \tag{3.4}$$

We can now state

LEMMA 3.1 (*A Posteriori X*-norm Error Bound). *For any $\boldsymbol{\mu} \in \mathcal{D}$, the RB error bound*

$$\Delta_N(\boldsymbol{\mu}) \equiv \frac{\|\mathcal{R}_N(\boldsymbol{\mu})\|_X}{\alpha_{\mathrm{LB}}(\boldsymbol{\mu})}, \tag{3.5}$$

satisfies

$$\|u(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})\|_X \leq \Delta_N(\boldsymbol{\mu}), \tag{3.6}$$

$$\frac{\Delta_N(\boldsymbol{\mu})}{\|u(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})\|_X} \leq \frac{\gamma_{\mathrm{UB}}(\boldsymbol{\mu})}{\alpha_{\mathrm{LB}}(\boldsymbol{\mu})}, \tag{3.7}$$

for $\alpha_{\mathrm{LB}}(\boldsymbol{\mu})$ and $\gamma_{\mathrm{UB}}(\boldsymbol{\mu})$ given by (2.6) *and* (2.7)*, respectively.*

Proof. The RB error, $e_N(\boldsymbol{\mu}) = u(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})$, satisfies the error-residual equation

$$a(e_N(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = r_N(v; \boldsymbol{\mu}), \quad \forall v \in X. \tag{3.8}$$

To obtain (3.6), we choose $e_N(\boldsymbol{\mu})$ for v in (3.8) and invoke (3.4) and the Cauchy-Schwarz inequality to get

$$a(e_N(\boldsymbol{\mu}), e_N(\boldsymbol{\mu}); \boldsymbol{\mu}) = (\mathcal{R}_N(\boldsymbol{\mu}), e_N(\boldsymbol{\mu}))_X \leq \|\mathcal{R}_N(\boldsymbol{\mu})\|_X \|e_N(\boldsymbol{\mu})\|_X; \tag{3.9}$$

we then invoke coercivity and (2.6) to arrive at

$$\alpha_{\mathrm{LB}} \|e_N(\boldsymbol{\mu})\|_X^2 \leq \|\mathcal{R}_N(\boldsymbol{\mu})\|_X \|e_N(\boldsymbol{\mu})\|_X. \tag{3.10}$$

The result (3.6) now directly follows from the definition (3.5).

To obtain (3.7), we choose $\mathcal{R}_N(\boldsymbol{\mu})$ for v in (3.8) and invoke (3.4), continuity, and (2.7) to get

$$\|\mathcal{R}_N(\boldsymbol{\mu})\|_X^2 = a(e_N(\boldsymbol{\mu}), \mathcal{R}_N(\boldsymbol{\mu}); \boldsymbol{\mu}) \leq \gamma_{\mathrm{UB}}(\boldsymbol{\mu}) \|e_N(\boldsymbol{\mu})\|_X \|\mathcal{R}_N(\boldsymbol{\mu})\|_X; \tag{3.11}$$

hence $\|\mathcal{R}_N(\boldsymbol{\mu})\|_X / \|e_N(\boldsymbol{\mu})\|_X \leq \gamma_{\mathrm{UB}}(\boldsymbol{\mu})$ and the result (3.7) follows from the definition (3.5). □

3.3. Construction–Evaluation Decomposition. Thanks to the assumption (2.1) on affine parameter dependence, the computational procedures for the RB solution and error bound admit *construction-evaluation* decompositions (see also [14, 16]): the construction stage is computationally expensive—the operation count depends on $\mathcal{N}$—but enables the subsequent evaluation stage in which we can rapidly—independently of $\mathcal{N}$—evaluate the RB approximation and RB error bound for any $\boldsymbol{\mu} \in \mathcal{D}$. (In actual practice we would of course also evaluate the RB output and RB output error bound—at negligible additional cost.) The construction-evaluation decomposition in turn permits the full offline–online computational decoupling described in the Introduction; we further discuss this decoupling below.

We first describe the construction-evaluation decomposition for the RB approximation: Let $\{\zeta_1 \in X_N, \ldots, \zeta_N \in X_N\}$ denote an X-orthonormal basis for X_N. In the construction stage, we assemble the matrices $A_N^q \in \mathbb{R}^{N \times N}$, $1 \leq q \leq Q_a$, and the vectors $F_N^q \in \mathbb{R}^N$, $1 \leq q \leq Q_f$, whose elements are defined by

$$A_{N,mn}^q \equiv a^q(\zeta_n, \zeta_m), \qquad F_{N,m}^q \equiv f^q(\zeta_m), \qquad 1 \leq m, n \leq N, \tag{3.12}$$

respectively. In the evaluation stage—given any $\boldsymbol{\mu} \in \mathcal{D}$—we evaluate the functions $\Theta_a^q(\boldsymbol{\mu})$, $1 \le q \le Q_a$, and $\Theta_f^q(\boldsymbol{\mu})$, $1 \le q \le Q_f$, in $\mathcal{O}(Q)$ operations; we then construct the RB stiffness matrix and load vector as

$$A_N(\boldsymbol{\mu}) = \sum_{q=1}^{Q_a} \Theta_a^q(\boldsymbol{\mu}) A_N^q, \qquad F_N(\boldsymbol{\mu}) = \sum_{q=1}^{Q_f} \Theta_f^q(\boldsymbol{\mu}) F_N^q, \tag{3.13}$$

respectively, in $\mathcal{O}(Q_a N^2 + Q_f N) = \mathcal{O}(QN^2)$ operations; finally, we solve the associated system of equations

$$A_N(\boldsymbol{\mu}) \underline{u}_N(\boldsymbol{\mu}) = F_N(\boldsymbol{\mu}) \tag{3.14}$$

for the RB basis coefficients $\underline{u}_N(\boldsymbol{\mu}) \equiv [u_{N,1}(\boldsymbol{\mu}), \dots, u_{N,N}(\boldsymbol{\mu})]^{\mathrm{T}}$ in $\mathcal{O}(N^3)$ operations (we must anticipate that $A_N(\boldsymbol{\mu})$ is dense).

We next describe the construction-evaluation decomposition for the dual norm of the residual. By linearity, we can write (3.4) as

$$\begin{aligned}(\mathcal{R}_N(\boldsymbol{\mu}), v)_X &= \sum_{q=1}^{Q_f} \Theta_f^q(\boldsymbol{\mu}) f^q(v) - \sum_{q=1}^{Q_a} \sum_{n=1}^{N} \Theta_a^q(\boldsymbol{\mu}) u_{N,n}(\boldsymbol{\mu}) a^q(\zeta_n, v) &(3.15)\\ &\equiv \sum_{n=1}^{\tilde{N}} \Gamma_n(\boldsymbol{\mu}) \mathcal{L}_n(v), &(3.16)\end{aligned}$$

where $\tilde{N} = Q_f + NQ_a$. By linear superposition, we can thus write

$$\mathcal{R}_N(\boldsymbol{\mu}) = \sum_{n=1}^{\tilde{N}} \Gamma_n(\boldsymbol{\mu}) \mathcal{G}_n, \tag{3.17}$$

where, for $1 \le n \le \tilde{N}$,

$$(\mathcal{G}_n, v)_X = \mathcal{L}_n(v), \quad \forall v \in X. \tag{3.18}$$

We thus have

$$\begin{aligned}\|\mathcal{R}_N(\boldsymbol{\mu})\|_X^2 &= (\mathcal{R}_N(\boldsymbol{\mu}), \mathcal{R}_N(\boldsymbol{\mu}))_X &(3.19)\\ &= \sum_{m=1}^{\tilde{N}} \sum_{n=1}^{\tilde{N}} \Gamma_m(\boldsymbol{\mu}) \Gamma_n(\boldsymbol{\mu}) G_{mn}, &(3.20)\end{aligned}$$

where the G_{mn} are defined as

$$G_{mn} \equiv (\mathcal{G}_m, \mathcal{G}_n)_X, \quad 1 \le m, n \le \tilde{N}. \tag{3.21}$$

In the construction stage we first perform the truth solves (3.18) for $\mathcal{G}_n$, $1 \le n \le \tilde{N}$; we then compute and store the inner products G_{mn}, $1 \le m, n \le \tilde{N}$. In the evaluation stage, we evaluate the functions $\Gamma_n(\boldsymbol{\mu})$, $1 \le n \le \tilde{N}$, in $\mathcal{O}(Q_f + NQ_a) = \mathcal{O}(NQ)$ operations and then perform the summation (3.19) in $\mathcal{O}(Q_f^2 + N^2 Q_a^2) = \mathcal{O}(N^2 Q^2)$ operations.

In general, the coercivity lower bound $\alpha_{\mathrm{LB}}(\boldsymbol{\mu})$ will not be known analytically and must be computed. An efficient construction-evaluation decomposition for the coercivity lower bound—the *successive constraint method*—can be found in [12, 20]; the evaluation complexity is independent of $\mathcal{N}$. We do not discuss this component further here in particular because for our particular numerical example of §6 an analytical lower bound $\alpha_{\mathrm{LB}}(\boldsymbol{\mu})$ is in fact available.

Algorithm 1: Greedy(Ξ, $\boldsymbol{\mu}_1$, ϵ_{tol}, $N_{\max}$)

```
initialize: $N \leftarrow 0$, $\epsilon_0 \leftarrow \infty$, $X_0 \leftarrow \{0\}$
while $\epsilon_N > \epsilon_{\mathrm{tol}}$ and $N < N_{\max}$ do
    $N \leftarrow N + 1$
    $X_N \leftarrow X_{N-1} \oplus \mathrm{span}\{u(\boldsymbol{\mu}_N)\}$
    $\epsilon_N \leftarrow \max_{\boldsymbol{\mu} \in \Xi} \Delta_N(\boldsymbol{\mu})$
    $\boldsymbol{\mu}_{N+1} \leftarrow \arg\max_{\boldsymbol{\mu} \in \Xi} \Delta_N(\boldsymbol{\mu})$
end
$N_{\max} \leftarrow N$
```

3.4. Greedy Parameter Sampling. We now discuss the construction of the hierarchical RB approximation spaces $X_N = \mathrm{span}\{u(\boldsymbol{\mu}_n)\}_{n=1}^N$, $1 \leq N \leq N_{\max}$ (see also [20, 22]). We first introduce a finite *train sample* $\Xi \subset \mathcal{D}$; a (random, say) initial parameter vector $\boldsymbol{\mu}_1 \in \mathcal{D}$; an error tolerance ϵ_{tol}; and a maximum RB dimension $N_{\max}$. We then perform Algorithm 1. The outputs of the algorithm are *nested* RB spaces $X_1 \subset X_2 \subset \ldots \subset X_{N_{\max}}$ $\big(\subset X\big)$. Note that the construction-evaluation decomposition allows us to use a dense train sample: each evaluation of the error bound in the arg max is very inexpensive; the truth is invoked only for the "winning" candidates, $\boldsymbol{\mu}_N$, $1 \leq N \leq N_{\max}$.

3.5. Offline-Online Computational Decoupling. We now describe the full offline-online decoupling procedure for the "p"-type RB approximation: the offline stage—performed only once as pre-processing—may be very expensive ($\mathcal{N}$-dependent) but enables the subsequent very fast ($\mathcal{N}$-independent) online stage—performed many times for the computation of the RB solution (and output) and RB error bound (and output error bound).

The offline stage is essentially the Greedy algorithm (Algorithm 1). The parameter independent entities $A_N^q \in \mathbb{R}^{N \times N}$, $1 \leq q \leq Q_a$, $F_N^q \in \mathbb{R}^N$, $1 \leq q \leq Q_f$ and $(\mathcal{G}_m, \mathcal{G}_n)_X$, $1 \leq m, n \leq \tilde{N}$ are retained from the construction stage of the last iteration. The permanent *online* storage requirement is thus $\mathcal{O}(Q_a N^2 + Q_f N) = \mathcal{O}(QN^2)$ for the A_N^q and F_N^q, and $\mathcal{O}(Q_a^2 N^2 + Q_f^2) = \mathcal{O}(Q^2 N^2)$ for the $(\mathcal{G}_m, \mathcal{G}_n)_X$. We note that since the RB spaces are nested, we can extract subarrays from the stored entities in order to construct RB approximations of any order $1 \leq N \leq N_{\max}$ (online adaptivity).

The online stage is for the "p"-type method equivalent to the evaluation stage: given any $\boldsymbol{\mu} \in \mathcal{D}$, we assemble the RB system in $\mathcal{O}(Q_f N + Q_a N^2) = \mathcal{O}(QN^2)$ operations, compute the RB solution in $\mathcal{O}(N^3)$ operations, and finally evaluate the RB error bound in $\mathcal{O}((Q_f + NQ_a)^2) = \mathcal{O}(N^2 Q^2)$ operations.

4. The "h"-type Reduced Basis Method. In this section we formulate the "h"-type reduced basis method. We first provide preliminaries required throughout this section; we next present the "h"-type approximation algorithm; we then consider *a posteriori* error estimation; we subsequently describe the offline-online computational decomposition; finally, we develop a new *a priori* convergence theory for the "zeroth order" approximation in the case of one parameter.

4.1. Preliminaries. We first introduce a set of Boolean vectors of length L,

$$\mathcal{B}_L \equiv \{1\} \times \{0,1\}^{L-1}; \tag{4.1}$$

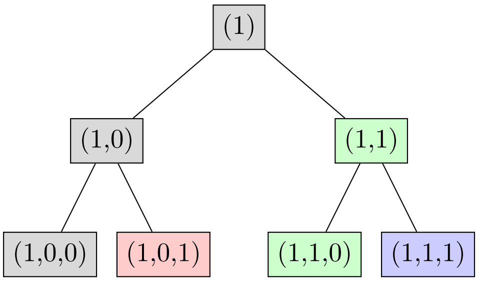

FIGURE 4.1. *A binary tree and associated Boolean vectors corresponding to the parameter domain partition in Figure 4.2.*

we denote a particular member of $\mathcal{B}_L$ as

$$B_L = (1, i_2 \dots, i_L) \in \mathcal{B}_L. \tag{4.2}$$

We can associate to $\mathcal{B}_L$ a perfect binary tree with L levels and at most $K = 2^{L-1}$ leaf nodes—as shown in Figure 4.1 for the particular case $L = 3$; we can identify to each B_L a node in the tree. Appending a ‘0’ to a vector B_L corresponds to a left bend and appending a ‘1’ to a vector B_L corresponds to a right bend. We define the concatenation

$$(B_L, i) \equiv (1, i_2 \dots, i_L, i), \qquad i \in \{0, 1\}; \tag{4.3}$$

we say that B_L is the parent of the children (B_L, i), $i \in \{0, 1\}$.

Given an initial parameter domain $\mathcal{D}$, we shall perform the “h”-refinement by recursive splitting of $\mathcal{D}$ into smaller parameter subdomains. The subdomains are defined hierarchically; thus for some $L \geq 1$, we can organize $K = 2^{L-1}$ subdomains in a perfect binary tree. We denote the subdomains as

$$\mathcal{V}_{B_l} \subset \mathcal{D}, \quad B_l \in \mathcal{B}_l, \quad 1 \leq l \leq L, \tag{4.4}$$

and we require the parent-child hierarchy

$$\mathcal{V}_{(B_l,0)} \subset \mathcal{V}_{B_l}. \tag{4.5}$$

$$\mathcal{V}_{(B_l,1)} \subset \mathcal{V}_{B_l}. \tag{4.6}$$

We associate to each subdomain $\mathcal{V}_{B_l}$ a set of $\bar{N}$ parameter values denoted by

$$\mathcal{M}_{\bar{N},B_l} = \{\boldsymbol{\mu}_{1,B_l}, \dots, \boldsymbol{\mu}_{\bar{N},B_l}\}, \quad 1 \leq l \leq L, \tag{4.7}$$

in which $\boldsymbol{\mu}_{1,B_l}, \dots, \boldsymbol{\mu}_{\bar{N},B_l} \in \mathcal{V}_{B_l}$; we may then define the RB approximation spaces (of dimension $\bar{N}$) associated with the subdomains as

$$X_{\bar{N},B_l} = \operatorname{span}\{u(\boldsymbol{\mu}_{1,B_l}), \dots, u(\boldsymbol{\mu}_{\bar{N},B_l})\}, \quad 1 \leq l \leq L. \tag{4.8}$$

(The actual bases are, as always, orthonormalized.)

To each “model” $\mathcal{M}_{\bar{N},B_l}$ and corresponding subdomain we associate a parameter *anchor point*, $\hat{\boldsymbol{\mu}}_{B_l}$, defined as

$$\hat{\boldsymbol{\mu}}_{B_l} \equiv \boldsymbol{\mu}_{1,B_l}. \tag{4.9}$$

We shall further require (by construction) that, for $2 \leq l \leq L-1$,

$$\hat{\boldsymbol{\mu}}_{(B_l,0)} = \hat{\boldsymbol{\mu}}_{B_l}, \tag{4.10}$$

$$\hat{\boldsymbol{\mu}}_{(B_l,1)} \neq \hat{\boldsymbol{\mu}}_{B_l}; \tag{4.11}$$

the anchor point is thus inherited only by the "left" child. The partition of $\mathcal{D}$ into subdomains is inferred from proximity to the anchor points.

To this end, we introduce for any Boolean vector $B_l \in \mathcal{B}_l$, $1 \leq l \leq L$, a "proximity function" $d_{B_l} : \mathcal{D} \to \mathbb{R}^+$. For example, we can choose the Euclidian distance between two points,

$$d_{B_l}(\boldsymbol{\mu}) = \|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}_{B_l}\|_2. \tag{4.12}$$

To determine for any new $\boldsymbol{\mu} \in \mathcal{D}$ which subdomain $\mathcal{V}_{(1,i_2^*,\ldots,i_L^*)} \subset \mathcal{D}$ contains $\boldsymbol{\mu}$, we successively evaluate the proximity function,

$$\begin{aligned}
i_2^* &= \arg\min_{i\in\{0,1\}} d_{(1,i)}(\boldsymbol{\mu}),\\
i_3^* &= \arg\min_{i\in\{0,1\}} d_{(1,i_2^*,i)}(\boldsymbol{\mu}),\\
&\ \vdots\\
i_L^* &= \arg\min_{i\in\{0,1\}} d_{(1,i_2^*,\ldots,i_{L-1}^*,i)}(\boldsymbol{\mu}).
\end{aligned} \tag{4.13}$$

We discuss the computational complexity shortly.

In general, the partition will not have the same number of refinement levels along every branch of the associated binary tree; hence the tree is not necessarily perfect. In this case, L shall denote the maximum number of levels in the tree—the tree depth; we will then have "empty models": $M_{B_l} = \emptyset$ for some $B_l \in \mathcal{B}_l$, $1 \leq l \leq L$. For any such B_l associated with an empty model we adopt the convention $d_{B_l} \equiv \infty$; we then terminate the search (4.13) whenever $d_{(1,i_2^*,\ldots,i)}(\boldsymbol{\mu}) = \infty$ (for $i = 0,1$). In order to measure the uniformity of the tree associated with the partition of $\mathcal{D}$ into K subdomains, we introduce a *relative tree depth*

$$\eta_{\text{depth}} = \frac{\text{tree depth}}{\log_2 K + 1}; \tag{4.14}$$

note in particular that $\eta_{\text{depth}} \geq 1$.

In what follows we shall need Algorithm 2, which is simply a restatement of the Greedy (Algorithm 1) restricted to a particular subdomain and with one additional output: the evaluation of the *a posteriori* error bound $\Delta_N^{h\text{RB}}$ (defined shortly) is performed over $\Xi_{B_L} \subset \mathcal{V}_{B_L}$; the outputs of the algorithm are nested RB spaces X_{N,B_L} *and* associated models $\mathcal{M}_{N,B_L}$, $1 \leq N \leq N_{\max,B_L}$. Note that even for $N_{\max,B_L} = 1$ we perform one pass of the whole loop and hence identify (and retain) $\boldsymbol{\mu}_{2,B_L}$; however, in general, we only compute at most $N_{\max,B_L}$ snapshots. For the pure "h"-type RB approximation we shall require $N_{\max,B_L} \equiv \bar{N}$ for all B_L.

4.2. Approximation. We now introduce the "h"-type RB approximation algorithm. We start from the original parameter domain $\mathcal{V}_{(1)} = \mathcal{D}$ ($L = 1$, $B_L = (1)$); we introduce a finite train sample $\Xi_{(1)} \subset \mathcal{V}_{(1)}$; we choose an initial parameter anchor point $\hat{\boldsymbol{\mu}}_{(1)} \in \mathcal{D}$; we choose the error tolerance ϵ_{tol}^1; we set the desired maximum RB space dimension $\bar{N} \geq 1$. The partition is then determined as follows.

Algorithm 2: Greedy$^2(\Xi_{B_L},\ \boldsymbol{\mu}_{1,B_L},\ \epsilon_{\text{tol}},\ N_{\max,B_L})$

```
initialize: $N \leftarrow 0$, $\epsilon_{0,B_L} \leftarrow \infty$, $X_{0,B_L} \leftarrow \{0\}$, $\mathcal{M}_{0,B_L} \leftarrow \emptyset$
while $\epsilon_N > \epsilon_{\text{tol}}$ and $N < N_{\max}$ do
    $N \leftarrow N + 1$
    $X_{N,B_L} \leftarrow X_{N-1,B_L} \oplus \operatorname{span}\{u(\boldsymbol{\mu}_{N,B_L})\}$
    $\mathcal{M}_{N,B_L} \leftarrow \mathcal{M}_{N-1,B_L} \cup \{\boldsymbol{\mu}_{N,B_L}\}$
    $\epsilon_{N,B_L} \leftarrow \max_{\boldsymbol{\mu} \in \Xi_{B_L}} \Delta_N^{h\mathrm{RB}}(\boldsymbol{\mu})$
    $\boldsymbol{\mu}_{N+1,B_L} \leftarrow \arg \max_{\boldsymbol{\mu} \in \Xi_{B_L}} \Delta_N^{h\mathrm{RB}}(\boldsymbol{\mu})$
end
$N_{\max,B_L} \leftarrow N$
```

1. Construct a model with $\bar{N}$ parameter values for the current subdomain $\mathcal{V}_{B_L}$ with the standard Greedy algorithm (Algorithm 2). The RB space $X_{\bar{N},B_L}$ and the model $\mathcal{M}_{\bar{N},B_L}$ are outputs from Greedy$^2(\Xi_{B_L},\ \hat{\boldsymbol{\mu}}_{(1)}, 0, \bar{N})$. (Note that we set the argument $\epsilon_{\text{tol}} = 0$ to enforce a RB space of dimension $\bar{N}$.)

2. Compute the maximum *a posteriori* error bound (defined shortly) over the train sample over the current subdomain

$$\epsilon_{\bar{N},B_L} = \max_{\boldsymbol{\mu} \in \Xi_{B_L}} \Delta_{\bar{N}}^{h\mathrm{RB}}(\boldsymbol{\mu}). \tag{4.15}$$

3. If $\epsilon_{\bar{N},B_L} < \epsilon_1^{\text{tol}}$: The refinement is sufficiently good; for all $\bar{N}$ set

$$\mathcal{M}_{\bar{N},(1,i_2,\dots,i_L,0)} = \emptyset, \tag{4.16}$$

$$\mathcal{M}_{\bar{N},(1,i_2,\dots,i_L,1)} = \emptyset; \tag{4.17}$$

we thus terminate the branch of the associated binary tree.

4. If $\epsilon_{\bar{N},B_L} \geq \epsilon_1^{\text{tol}}$:

(i) Define anchor points for two new models $\mathcal{M}_{(B_L,0)}$ and $\mathcal{M}_{(B_L,1)}$, $\hat{\boldsymbol{\mu}}_{(B_L,0)}$ ($= \hat{\boldsymbol{\mu}}_{B_L}$) and $\hat{\boldsymbol{\mu}}_{(B_L,1)}$, respectively; the model $\mathcal{M}_{(B_L,0)}$ inherits the anchor point from its "parent," while the model $\mathcal{M}_{(B_L,1)}$ takes as anchor point the first parameter value chosen by the Greedy algorithm—in the sense of the *a posteriori* error estimator, these two points are maximally different and hence good places to "anchor" the new models. (Note the remaining $\bar{N} - 1$ snapshots of $\mathcal{M}_{B_L}$ are discarded.)

(ii) Define a new and denser train sample $\tilde{\Xi}_{B_L} \subset \mathcal{V}_{B_L}$ of size $|\tilde{\Xi}_{B_L}| = 2|\Xi_{(1)}|$. (The temporary sample $\tilde{\Xi}_{B_L}$ is thus twice as large as the initial train sample.)

(iii) Construct $\Xi_{(B_L,0)} \subset \mathcal{V}_{(B_L,0)}$ and $\Xi_{(B_L,1)} \subset \mathcal{V}_{(B_L,1)}$ from $\tilde{\Xi}_{B_L}$ based on proximity to $\hat{\boldsymbol{\mu}}_{(B_L,0)}$ and $\hat{\boldsymbol{\mu}}_{(B_L,1)}$, respectively: a point $\boldsymbol{\mu} \in \tilde{\Xi}_{B_L}$ belongs to $\Xi_{(B_L,0)}$ if and only if $d_{(B_L,0)}(\boldsymbol{\mu}) \leq d_{(B_L,1)}(\boldsymbol{\mu})$; otherwise $\boldsymbol{\mu}$ belongs to $\Xi_{(B_L,1)}$.

5. Split the current branch in two new branches: set $B_L^{\text{left}} = (B_L, 0)$ and $B_L^{\text{right}} = (B_L, 1)$; proceed to step 1 first for $B_L = B_L^{\text{left}}$ and then for $B_L = B_L^{\text{right}}$.

The procedure may be more precisely defined by hRB$(\Xi_{(1)}, \hat{\boldsymbol{\mu}}_{(1)}, \bar{N}, \epsilon_{\text{tol}}^1)$, where hRB is the recursive function defined in Algorithm 3.

Remark 1 (Train Sample Refinement). In step 4(ii) in the algorithm above additional points are added to the train sample such that the number of points in the two *new* train samples will be roughly the same as in the old train sample, and in particular always much larger than $\bar{N}$. As a result, the "global" train sample over

Algorithm 3: $h\mathrm{RB}(\Xi_{B_L}, \hat{\boldsymbol{\mu}}_{B_L}, \bar{N}, \epsilon^1_{\mathrm{tol}})$

```
Get $X_{\bar{N},B_L}$ and $\mathcal{M}_{\bar{N},B_L}$ from Greedy$^2(\hat{\boldsymbol{\mu}}_{B_L}, B_L, \infty, \bar{N})$
$\epsilon_{\bar{N},B_L} \leftarrow \max_{\boldsymbol{\mu}\in\Xi_{B_L}} \Delta^{h\mathrm{RB}}_{\bar{N}}(\boldsymbol{\mu})$;
if $\epsilon_{\bar{N},B_L} < \epsilon^1_{\mathrm{tol}}$ then
    Terminate branch: $\mathcal{M}_{(B_L,0)} = \emptyset$ and $\mathcal{M}_{(B_L,1)} = \emptyset$.
else
    Define $\hat{\boldsymbol{\mu}}_{(B_L,0)} = \boldsymbol{\mu}_{1,B_L}$ and $\hat{\boldsymbol{\mu}}_{(B_L,1)} = \boldsymbol{\mu}_{2,B_L}$
    Construct $\Xi_{(B_L,0)} \subset \mathcal{V}_{(B_L,0)}$ and $\Xi_{(B_L,1)} \subset \mathcal{V}_{(B_L,1)}$
    $h\mathrm{RB}(\Xi_{(B_L,0)}, \hat{\boldsymbol{\mu}}_{(B_L,0)}, \bar{N}, \epsilon^1_{\mathrm{tol}})$
    $h\mathrm{RB}(\Xi_{(B_L,1)}, \hat{\boldsymbol{\mu}}_{(B_L,1)}, \bar{N}, \epsilon^1_{\mathrm{tol}})$
end
```

$\mathcal{D}$—the union of all the points in the train samples over all parameter subdomains—is adaptively refined as the "h"-type RB approximation becomes more accurate: the train sample is denser in regions of $\mathcal{D}$ with smaller subdomains; hence we add more train points where the solution varies rapidly with the parameters.

In our current implementation, the train sample refinement is performed by sampling of uniformly distributed random points from $\mathcal{D}$; we then use the search (4.13) to determine whether a point belongs to the current subdomain and thus can be included as a new point in the current train sample. In the case that the proximity function is Euclidian distance (as in (4.12)), we need in fact not sample from the entire parameter domain $\mathcal{D}$: we first compute the bounding box of the old train sample; we then sample the new points from a box that contains the bounding box with some specified margin—the assumption is that the box from which we sample contains the entire subdomain. In the case in which the proximity function is the error bound (as we describe shortly), we sample from the entire domain $\mathcal{D}$ since we have no *a priori* knowledge of the shape of the subdomains, and in particular the subdomains might not be connected. $\diamondsuit$

Remark 2 (Offline Speedup). The greedy algorithm—in particular in the case of a low order (small $\bar{N}$) approximation—is likely to choose parameter values close to the boundaries of the parameter subdomains. As a result, two or more models may comprise some identical (or nearly identical) parameter values, and thus some of the offline truth solves are in some sense redundant. One way to reduce this snapshot redundancy is to share basis functions between approximation spaces if the associated greedily selected parameter values are sufficiently close. The development of an efficient algorithm for automatic sharing of basis functions is the subject of future work. $\diamondsuit$

In Figure 4.2 we illustrate the first two levels of "h"-refinement together with the associated binary tree for a "h"-type approximation with $\bar{N} = 1$. The first anchor point $\hat{\boldsymbol{\mu}}_{(1)}$ is chosen as the upper right corner of the parameter domain, and by definition $\mathcal{V}_{(1)} = \mathcal{D}$. The method greedily chooses the point $\hat{\boldsymbol{\mu}}_{(1,1)}$ near the lower left corner of $\mathcal{V}_{(1)}$; the initial anchor point is then re-labelled as $\hat{\boldsymbol{\mu}}_{(1,0)} = \hat{\boldsymbol{\mu}}_{(1)}$. We now have two new models $\mathcal{M}_{1,(1,0)} = \{\hat{\boldsymbol{\mu}}_{(1,0)}\}$ and $\mathcal{M}_{1,(1,1)} = \{\hat{\boldsymbol{\mu}}_{(1,1)}\}$, whose associated subdomains $\mathcal{V}_{(1,0)}$ and $\mathcal{V}_{(1,1)}$ are determined from proximity—here Euclidian distance—to the two anchor points. Next, $\mathcal{V}_{(1,1)}$ and $\mathcal{V}_{(1,0)}$ are partitioned in the same fashion (we have here assumed that the tolerance is satisfied within $\mathcal{V}_{(1,0,0)}$ and $\mathcal{V}_{(1,0,1)}$).

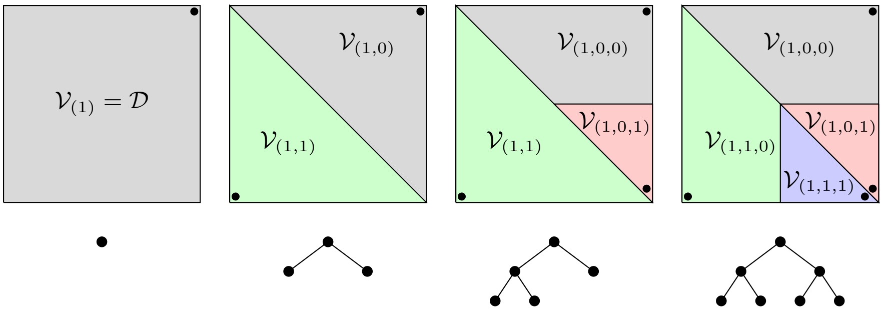

FIGURE 4.2. *Two levels of “h”-refinement and associated binary tree.*

Finally, we may now define the “h”-type RB approximation: given any $\boldsymbol{\mu} \in \mathcal{D}$ we first determine the subdomain $\mathcal{V}_{B_L^*}$ containing $\boldsymbol{\mu}$ from the search (4.13); we then find $u_{\bar{N}}^{h\mathrm{RB}}(\boldsymbol{\mu}) \in X_{\bar{N},B_L^*}$ such that

$$a(u_{\bar{N}}^{h\mathrm{RB}}(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = f(v; \boldsymbol{\mu}), \quad \forall v \in X_{\bar{N},B_L^*}. \tag{4.18}$$

(Note that B_L^* depend on $\boldsymbol{\mu}$.) We discuss computational complexity shortly. We *define* the “order” of the “h”-type approximation as $p \equiv \bar{N}^{1/P} - 1$.

4.3. *A Posteriori* **Error Estimation.** We can apply the same *a posteriori* bound developed for the “p”-type RB approximation in §3.2 to the “h”-type (and below “hp”-type) RB approximations. However, we shall require some new notation for the “h”-type error bound.

Given any $\boldsymbol{\mu} \in \mathcal{D}$ and a partition of $\mathcal{D}$ into subdomains, we determine B_L^* from the binary search (4.13) and compute the RB solution $u_{\bar{N}}^{h\mathrm{RB}}(\boldsymbol{\mu})$ from (4.18). The RB residual is

$$r_{\bar{N}}^{h\mathrm{RB}}(v; \boldsymbol{\mu}) = f(v; \boldsymbol{\mu}) - a(u_{\bar{N}}^{h\mathrm{RB}}(\boldsymbol{\mu}), v; \boldsymbol{\mu}), \quad \forall v \in X. \tag{4.19}$$

We denote the Riesz representation of the residual by $\mathcal{R}_{\bar{N}}^{h\mathrm{RB}}$; as an upper bound for the X-norm error $\|u(\boldsymbol{\mu}) - u_{\bar{N}}^{h\mathrm{RB}}(\boldsymbol{\mu})\|_X$, we define

$$\Delta_{\bar{N}}^{h\mathrm{RB}}(\boldsymbol{\mu}) \equiv \frac{\|\mathcal{R}_{\bar{N}}^{h\mathrm{RB}}(\boldsymbol{\mu})\|_X}{\alpha_{\mathrm{LB}}(\boldsymbol{\mu})}. \tag{4.20}$$

Lemma 3.1 now directly applies with an appropriate change of notation.

Remark 3 (The Error Bound as Proximity Function). For any $B_l \in \mathcal{B}_l$, $1 \le l \le L$, (associated with a non-empty model) and any $\boldsymbol{\mu} \in \mathcal{D}$, we can derive the RB error bound associated with the RB approximation to $u(\boldsymbol{\mu})$ in the space $X_{\bar{N},B_l}$; we denote this error bound by $\Delta_{\bar{N},B_l}(\boldsymbol{\mu})$. As an alternative to the proximity function introduced in (4.12), we can use

$$d_{B_l}(\boldsymbol{\mu}) = \Delta_{B_l}(\boldsymbol{\mu}), \tag{4.21}$$

to measure the “distance” between the points $\hat{\boldsymbol{\mu}}_{B_l}$ and $\boldsymbol{\mu}$. In §6, we provide results with the proximity function defined both as in (4.12) and as in (4.21). ◊

Remark 4 (Multiple Inner-Products). The “h”-type RB approximation offers a natural way of introducing multiple inner-products—multiple reference parameters—in the computation of the dual norm of the residual for the *a posteriori* error bounds.

For example, we could choose the anchor point in any subdomain to be the reference parameter associated with that subdomain. With this approach, we would expect sharper error bounds and thus a better parameter domain partition (as well as, ultimately, greater online efficiency).

To compute the dual norm of the residual we must (in the construction stage) solve a number of problems on the form (3.4) with different right-hand sides. If we solve the discrete system directly, we must invert one operator for each inner-product; hence there is a computational advantage associated with only a single inner-product. If we use an iterative solver, however, the solves are in any event performed independently and we can introduce individual inner-products within each subdomain at no computational penalty. In this paper, however, we have not pursued a multiple inner-product approach for our numerical examples. $\diamondsuit$

4.4. Offline-Online Decomposition. In the *offline* stage, we determine the partition of the parameter domain and construct the corresponding RB models and spaces: we perform $h\mathrm{RB}(\Xi_{(1)}, \hat{\boldsymbol{\mu}}_{(1)}, \bar{N}, \epsilon^1_{\mathrm{tol}})$. For our purposes here, we assume a perfect binary tree; note that a perfect binary tree with K leaf nodes has $K-1$ additional nodes associated with intermediate models (at earlier levels in the tree). We also assume that the cardinality of the train sample over each of the subdomains is equal to n_{train}.

The offline stage computational cost derives from several components:

1. Snapshot Truth Solves. During the partition procedure, we must compute $\bar{N}K$ snapshots associated with the final approximation spaces. In addition, we must compute $(\bar{N}-1)(K-1)$ snapshots associated with intermediate models required to form the partition. (Since the anchor point—and thus the first basis function—at the next level is inherited from the "parent", only $\bar{N}-1$ *new* basis functions are required for each model.)

2. Reduced Basis Preprocessing. We must compute $K(Q_a\bar{N}^2 + Q_f\bar{N})$ truth inner products to form the parameter-independent "stiffness" matrices and loads (e.g., as in (3.12)) for the final models, and an additional $(K-1)(Q_a\bar{N}^2 + Q_f\bar{N} - Q_a - Q_f)$ truth inner products to form the corresponding quantities for the intermediate models.

3. Error Bound Preprocessing. We must compute $K\bar{N}Q_a + Q_f$ truth Poisson solves of the form (3.18) for the final models and an additional $(K-1)(\bar{N}-1)Q_a$ truth Poisson solves for the intermediate models. We must also compute the $K(\bar{N}Q_a + Q_f)^2$ truth inner-products of the form (3.21) in order to evaluate the dual norm of the residual associated with the final models, and an additional $(K-1)((\bar{N}Q_a + Q_f)^2 - (Q_a + Q_f)^2)$ truth inner-products in order to evaluate the dual norm of the residual associated with the intermediate models.

4. Error Bound Evaluations. We must solve $n_{\mathrm{train}}(\bar{N}(K-1) + \bar{N}K)$ RB systems to obtain the residual coefficients and evaluate $n_{\mathrm{train}}(\bar{N}(K-1) + \bar{N}K)$ RB error bounds during the Greedy sampling including both the final and intermediate models. This results in $n_{\mathrm{train}}\bar{N}(2K-1)(\bar{N}^3 + \bar{N}^2Q^2)$ operations in total (to leading order).

The combined offline cost is thus approximately $2\bar{N}K + 2\bar{N}K + Q_f$ truth solves, $2K(\bar{N}Q_a + Q_f)^2 + 2K(Q_a\bar{N}^2 + Q_f\bar{N})$ truth inner-products, and $n_{\mathrm{train}}2\bar{N}K(N^3 + N^2Q^2)$ operations to evaluate the error bounds. Note that the additional cost associated with the intermediate models is not onerous—a factor of two.

The link between the offline and online stages is the parameter-independent data *constructed* in the offline stage and stored (permanently) for *evaluation* in the online stage. If we retain only the data associated with the final models the online storage for the "h"-type RB approximation is Q_aK matrices of size $\bar{N}\times\bar{N}$ and Q_fK vectors of size $\bar{N}$; the online storage associated with the RB error bounds is $K(\bar{N}Q_a + Q_f)^2/2$. If we retain intermediate models for purposes of online adaptivity clearly the online storage will increase; we do not consider this case further since in actual practice online adaptivity is typically pursued through the "*hp*"-approach.

In the *online* stage, given any $\boldsymbol{\mu} \in \mathcal{D}$, we first determine the subdomain which contains $\boldsymbol{\mu}$ via the binary search (4.13) in $\mathcal{O}(\log_2 K)$ operations. Thanks to the construction-evaluation decomposition, we can then assemble and solve the corresponding system of algebraic equations in $\mathcal{O}(Q\bar{N}^2)$ and $\mathcal{O}(\bar{N}^3)$ operations, respectively, and compute the associated *a posteriori* error bound in $\mathcal{O}(\bar{N}^2Q^2)$ operations. Note that the search (4.13) is an $\mathcal{O}(\log_2 K)$ operation only under the hypothesis that the depth of the tree associated with the partition of $\mathcal{D}$ is proportional to $\log_2 K$; we provide numerical results to support this hypothesis in §6. We also emphasize that the efficient $\mathcal{O}(\log_2 K)$ search is a particular property of our hierarchical partition construction; if we were to partition the parameter domain based on (say) a Voronoi diagram, determination of the subdomain which contains $\boldsymbol{\mu} \in \mathcal{D}$ would be less efficient.

4.5. *A Priori* **Theory:** $\bar{N} = 1$**,** $P = 1$**.** In this section we develop *a priori* convergence theory for a "*h*"–type RB approximation of "zeroth order" ($\bar{N} = 1$) in the one-parameter case ($P = 1$) when the Euclidian distance is used as the proximity function. We focus on $\bar{N} = 1$ since in fact $\bar{N} = 1$ is crucial to the "*hp*"-approach of §5; the theory developed here demonstrates that an $\bar{N} = 1$ greedy approach can indeed generate a reasonably efficient partition. We consider $P = 1$ for simplicity; at the conclusion of this section we provide a remark addressing the $\bar{N} > 1$ (higher "order") and $P > 1$ cases.

For our purposes here, we do not need the Boolean indexing of the anchor points and subdomains: we assume that we have partitioned $\mathcal{D}$ into K subdomains; we relabel the K anchor points as $\hat{\boldsymbol{\mu}}'_1, \hat{\boldsymbol{\mu}}'_2, \ldots, \hat{\boldsymbol{\mu}}'_K$ (numbered in the order in which they are chosen by Algorithm 3). When Algorithm 3 adds a new anchor point, the parameter domain partition changes; we introduce mappings $I_{\tilde{K}} : \mathcal{D} \to \{1, \ldots, \tilde{K}\}$, $1 \le \tilde{K} \le K$ such that with $\tilde{K}$ anchor points, for any $\boldsymbol{\mu} \in \mathcal{D}$,

$$\hat{\boldsymbol{\mu}}'_{I_{\tilde{K}}(\boldsymbol{\mu})} = \hat{\boldsymbol{\mu}}_{B^*(\boldsymbol{\mu})}, \tag{4.22}$$

where $B^*(\boldsymbol{\mu})$ is the Boolean index of the particular subdomain containing $\boldsymbol{\mu}$. Below, we omit the $'$ for brevity.

For the purpose of this section, given $\tilde{K}$ anchor points and corresponding subdomains, we denote by $u_{\tilde{K}}(\boldsymbol{\mu})$ the "zeroth order" ($\bar{N} = 1$) "*h*"-type RB approximation for any $\boldsymbol{\mu} \in \mathcal{D}$. With the implicit mapping above, we have

$$u_{\tilde{K}}(\boldsymbol{\mu}) = \omega_{\tilde{K}}(\boldsymbol{\mu}) u(\hat{\boldsymbol{\mu}}_{I_{\tilde{K}}(\boldsymbol{\mu})}), \tag{4.23}$$

where the coefficient $\omega_{\tilde{K}}(\boldsymbol{\mu})$ is given by the Galerkin projection as

$$\omega_{\tilde{K}}(\boldsymbol{\mu}) = \frac{f\big(u(\hat{\boldsymbol{\mu}}_{I_{\tilde{K}}(\boldsymbol{\mu})}); \boldsymbol{\mu}\big)}{a\big(u(\hat{\boldsymbol{\mu}}_{I_{\tilde{K}}(\boldsymbol{\mu})}), u(\hat{\boldsymbol{\mu}}_{I_{\tilde{K}}(\boldsymbol{\mu})}); \boldsymbol{\mu}\big)}. \tag{4.24}$$

We denote by $r_{\tilde{K}}(v; \boldsymbol{\mu}) = f(v; \boldsymbol{\mu}) - a(u_{\tilde{K}}(\boldsymbol{\mu}), v; \boldsymbol{\mu})$ the RB residual, and let $\mathcal{R}_{\tilde{K}}(\boldsymbol{\mu}) \in X$ satisfy $(\mathcal{R}_{\tilde{K}}(\boldsymbol{\mu}), v)_X = r_{\tilde{K}}(v; \boldsymbol{\mu})$ for all $v \in X$. Our X-norm error upper bound is then written

$$\Delta_{\tilde{K}}(\boldsymbol{\mu}) = \frac{\|\mathcal{R}_{\tilde{K}}(\boldsymbol{\mu})\|_X}{\alpha_{\mathrm{LB}}(\boldsymbol{\mu})}, \tag{4.25}$$

which is simply a specialization of (4.20).

We need two further preliminary results. First, it is clear from Cea's Lemma (with respect to the X-norm), (2.9), and (2.8) that for any $\tilde{K}$, $1 \leq \tilde{K} \leq K$, and any $\boldsymbol{\mu} \in \mathcal{D}$,

$$\|u(\boldsymbol{\mu}) - u_{\tilde{K}}(\boldsymbol{\mu})\|_X \leq \frac{\overline{\gamma}}{\underline{\alpha}} \|u(\boldsymbol{\mu}) - u(\hat{\boldsymbol{\mu}}_{I_{\tilde{K}}(\boldsymbol{\mu})})\|_X, \tag{4.26}$$

since $u(\hat{\boldsymbol{\mu}}_{I_{\tilde{K}}(\boldsymbol{\mu})})$ is a particular member of the (one-dimensional) reduced basis space. Second, from (3.7) of Lemma 3.1, we get for any $\tilde{K}$, $1 \leq \tilde{K} \leq K$, and any $\boldsymbol{\mu} \in \mathcal{D}$,

$$\Delta_{\tilde{K}}(\boldsymbol{\mu}) \leq \frac{\overline{\gamma}}{\underline{\alpha}} \|u(\boldsymbol{\mu}) - u_{\tilde{K}}(\boldsymbol{\mu})\|_X, \quad 1 \leq \tilde{K} \leq K. \tag{4.27}$$

We can now state

Proposition 4.1 (Convergence in the case $\bar{N} = 1$, $P = 1$). *The "h"-type RB approximation is convergent for finite $K(\epsilon^1_{\rm tol}) \leq K_{\max}(\epsilon^1_{\rm tol})$ subdomains. Further, the convergence is first order in the sense that*

$$K_{\max}(\epsilon^1_{\rm tol}) = \max\left\{1, \frac{C}{\epsilon^1_{\rm tol}}\right\} \tag{4.28}$$

where the constant C is given by

$$C = \frac{2\overline{\gamma}^2 \tilde{C} |\mathcal{D}|}{\underline{\alpha}^2}, \tag{4.29}$$

where $\tilde{C} = (\tilde{c}_1 \|f\|_{X'} + \underline{\alpha}\tilde{c}_2)/\underline{\alpha}^2$ is the constant developed in Lemma 2.1 and $|\mathcal{D}|$ is the length of $\mathcal{D} \subset \mathbb{R}$.

Proof. Algorithm 3 provides a sequence of anchor points $\hat{\boldsymbol{\mu}}_1, \ldots, \hat{\boldsymbol{\mu}}_K$ for $K \geq 1$. We have by construction of our algorithm *either* $K = 1$ *or* $K > 1$ and

$$\epsilon^1_{\rm tol} < \Delta_{\tilde{K}}(\hat{\boldsymbol{\mu}}_{\tilde{K}+1}), \quad 1 \leq \tilde{K} \leq K - 1. \tag{4.30}$$

In the former case the proof is complete; we henceforth consider the latter case.

We deduce from (4.27), (4.26), and Lemma 2.1,

$$\Delta_{\tilde{K}}(\hat{\boldsymbol{\mu}}_{\tilde{K}+1}) \leq \frac{\overline{\gamma}}{\underline{\alpha}} \|u(\hat{\boldsymbol{\mu}}_{\tilde{K}+1}) - u_{\tilde{K}}(\hat{\boldsymbol{\mu}}_{\tilde{K}+1})\|_X \tag{4.31}$$

$$\leq \frac{\overline{\gamma}^2}{\underline{\alpha}^2} \|u(\hat{\boldsymbol{\mu}}_{\tilde{K}+1}) - u(\hat{\boldsymbol{\mu}}_{I_{\tilde{K}}(\hat{\boldsymbol{\mu}}_{\tilde{K}+1})})\|_X \tag{4.32}$$

$$\leq \frac{\overline{\gamma}^2}{\underline{\alpha}^2} \tilde{C} |\hat{\boldsymbol{\mu}}_{\tilde{K}+1} - \hat{\boldsymbol{\mu}}_{I_{\tilde{K}}(\hat{\boldsymbol{\mu}}_{\tilde{K}+1})}|, \tag{4.33}$$

respectively, for $1 \leq \tilde{K} \leq K - 1$; hence from (4.30)

$$|\hat{\boldsymbol{\mu}}_{\tilde{K}+1} - \hat{\boldsymbol{\mu}}_{I_{\tilde{K}}(\hat{\boldsymbol{\mu}}_{\tilde{K}+1})}| > \frac{\underline{\alpha}^2 \epsilon^1_{\rm tol}}{\overline{\gamma}^2 \tilde{C}}, \tag{4.34}$$

for $1 \leq \tilde{K} \leq K - 1$.

Let δ^K_k denote the length of the subdomain associated with anchor point $\hat{\boldsymbol{\mu}}_k$, $1 \leq k \leq K$. Given $\tilde{K}$, $1 \leq \tilde{K} \leq K - 1$, the algorithm selects the next anchor point $\hat{\boldsymbol{\mu}}_{\tilde{K}+1}$

and the intermediate subdomain associated with anchor point number $I_{\tilde{K}}(\hat{\boldsymbol{\mu}}_{\tilde{K}+1})$ is divided into two new subdomains. It is clear that the length of each of the two new subdomains is at least as large as half the distance between the new anchor point $\hat{\boldsymbol{\mu}}_{\tilde{K}+1}$ and anchor point $\hat{\boldsymbol{\mu}}_{I_{\tilde{K}}(\hat{\boldsymbol{\mu}}_{\tilde{K}+1})}$, namely $|\hat{\boldsymbol{\mu}}_{\tilde{K}+1} - \hat{\boldsymbol{\mu}}_{I_{\tilde{K}}(\hat{\boldsymbol{\mu}}_{\tilde{K}+1})}|/2$. Furthermore, each of the K *final* subdomains results from the splitting of the intermediate subdomain associated with anchor point $\hat{\boldsymbol{\mu}}_{I_{\tilde{K}}(\hat{\boldsymbol{\mu}}_{\tilde{K}+1})}$ for *some* $\tilde{K} \in \{1, \ldots, K-1\}$; hence for $1 \le k \le K$, there exists a $\tilde{K} \in \{1, \ldots, K-1\}$ such that

$$\delta_k^K \ge |\hat{\boldsymbol{\mu}}_{\tilde{K}+1} - \hat{\boldsymbol{\mu}}_{I_{\tilde{K}}(\hat{\boldsymbol{\mu}}_{\tilde{K}+1})}|/2, \tag{4.35}$$

and thus by (4.34)

$$\underline{\delta}^K \equiv \min_{1 \le k \le K} \delta_k^K > \frac{\underline{\alpha}^2 \epsilon_{\text{tol}}^1}{2\overline{\gamma}^2 \tilde{C}}. \tag{4.36}$$

Note that $\underline{\delta}_K$ is not the smallest distance between two anchor points: rather, it is the smallest length of any of the K subdomains.

Let $|\mathcal{D}|$ denote the length of $\mathcal{D}$. With K subdomains, it is clear that $K\underline{\delta}^K \le |\mathcal{D}|$. We now assume $K > K_{\max}$. From (4.28) and (4.36) it then follows that

$$K\underline{\delta}^K > \frac{C}{\epsilon_{\text{tol}}^1}\underline{\delta}^K \ge \left(\frac{2\overline{\gamma}^2\tilde{C}|\mathcal{D}|}{\underline{\alpha}^2\epsilon_{\text{tol}}^1}\right)\left(\frac{\underline{\alpha}^2\epsilon_{\text{tol}}^1}{2\overline{\gamma}^2\tilde{C}}\right) = |\mathcal{D}|. \tag{4.37}$$

We have thus reached a contradiction: the "*h*"-type RB approximation can not generate a sequence of anchor points $\hat{\boldsymbol{\mu}}_1, \ldots, \hat{\boldsymbol{\mu}}_K$ for $K > K_{\max}$; thus the algorithm must be convergent for $1 \le K \le K_{\max}$ subdomains. □

Remark 5 (Convergence in the case $\bar{N} \ge 1$, $P \ge 1$). We first recall a polynomial approximation result. Consider piecewise polynomial interpolation of order p of a sufficiently smooth function on a bounded domain in $\mathbb{R}^P$. We expect the convergence to be of order $p+1$: with K subdomains we expect the error to decrease as $(1/K)^{(p+1)/P}$. Further, with each subdomain we can associate $\bar{N} = (p+1)^P$ degrees of freedom; we can thus expect the error to decrease as $(1/K)^{(\bar{N}^{1/P})/P}$.

In the "zeroth order" multi-parameter case ($\bar{N} = 1$, $P > 1$) we anticipate that our method converges for

$$K < \max\left\{1, \frac{C}{(\epsilon_{\text{tol}}^1)^P}\right\} \tag{4.38}$$

subdomains for some positive constant C. This poor convergence for $P \gg 1$ suggests the advantage of "*p*"–convergence [21] or "*hp*"–convergence rather than solely "*h*"-convergence. Next, in the higher order, one-parameter case ($\bar{N} > 1$, $P = 1$), we might expect convergence of $\bar{N}$th-order in the sense that

$$K < \max\left\{1, \frac{C}{(\epsilon_{\text{tol}}^1)^{\frac{1}{\bar{N}}}}\right\} \tag{4.39}$$

for some positive constant C. Finally, in the *general* case $\bar{N} \ge 1$, $P \ge 1$, we might expect convergence of order $p+1$, or

$$K < \max\left\{1, \frac{C}{(\epsilon_{\text{tol}}^1)^{\frac{P}{\bar{N}^{1/P}}}}\right\}. \tag{4.40}$$

We shall consider these heuristic arguments again in the context of numerical results.

Note that our bound (4.28) and estimators (4.39) and (4.40) should capture the correct order but of course the constant will be very pessimistic: by design, the Greedy should adapt the sample to best accomodate local variations. $\diamondsuit$

5. The "*hp*"-type Reduced Basis Method. With the "*hp*"-type RB method, we combine the "*h*"- and "*p*"-type methods: we first construct a partition of the parameter domain with "*h*"-refinement; we then compute independent approximation spaces restricted to each parameter subdomain with "*p*"-refinement—in general, the approximation spaces will have different dimensions.

5.1. Approximation. The parameter domain partition is first constructed by an $\bar{N} = 1$ "*h*"-type approximation until the error bound tolerance $\epsilon^1_{\rm tol}$ is satisfied. We first construct the initial train sample $\Xi_{(1)} \subset \mathcal{D}$, choose an initial parameter anchor $\hat{\boldsymbol{\mu}}_{(1)} \in \mathcal{D}$, and specify $\epsilon^1_{\rm tol}$; we then execute Algorithm 3, $h{\rm RB}(\Xi_{(1)}, \hat{\boldsymbol{\mu}}_{(1)}, 1, \epsilon^1_{\rm tol})$.

The output from $h{\rm RB}(\Xi_{(1)}, \hat{\boldsymbol{\mu}}_{(1)}, 1, \epsilon^1_{\rm tol})$ is K subdomains with associated one-parameter models and one-dimensional approximation spaces; we denote by $B^1, \ldots, B^K$ the K associated Boolean indices. We also store the train sample over each of the final subdomains. As an additional step we now append additional basis functions to each approximation space with a standard "*p*"-type procedure over each train sample: we specify the maximum RB space dimension $N_{\max,B^k} = N_{\max} > \bar{N}$, $1 \le k \le K$; we specify a new error bound tolerance $\epsilon^2_{\rm tol} < \epsilon^1_{\rm tol}$; for $1 \le k \le K$, we then execute Algorithm 2: ${\rm Greedy}^2(\Xi_{B^k}, \boldsymbol{\mu}_{1,B^k}, \epsilon^2_{\rm tol}, N_{\max,B^k})$. The final output is thus K sets of nested RB approximation spaces and associated models, $X_{1,B^k} \subset, \ldots, \subset X_{N_{\max},B^k}$ and $\mathcal{M}_{1,B^k} \subset \mathcal{M}_{N_{\max},B^k}$, $1 \le k \le K$, respectively. Note that the dimension of the spaces is in general different since the error bound tolerance $\epsilon^2_{\rm tol}$ might be satisfied by the different approximation spaces over the different train samples with a different number of basis functions.

Finally, we may now define the "*hp*"-type RB approximation: given any $\boldsymbol{\mu} \in \mathcal{D}$, we first determine the subdomain $\mathcal{V}_{B^*_L}$ containing $\boldsymbol{\mu}$ from the search (4.13); given any $1 \le N \le N_{\max}$, we then find $u^{hp{\rm RB}}_N(\boldsymbol{\mu}) \in X_{\hat{N},B^*_L}$ such that

$$a(u^{hp{\rm RB}}_N(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = f(v; \boldsymbol{\mu}), \quad \forall v \in X_{\hat{N},B^*_L}, \tag{5.1}$$

where $\hat{N} \equiv \min\{N, N_{\max,B^*_L}\}$. (Note that B^*_L and thus $\hat{N}$ depend on $\boldsymbol{\mu}$.)

5.2. ***A Posteriori* Error Estimation.** We shall require some new notation for the "*hp*"-type *a posteriori* error bound.

Given any $\boldsymbol{\mu} \in \mathcal{D}$ and a partition of $\mathcal{D}$ into subdomains, we determine B^*_L from the binary search (4.13) and compute the RB solution $u^{hp{\rm RB}}_N(\boldsymbol{\mu})$ from (5.1). The RB residual is

$$r^{hp{\rm RB}}_N(v; \boldsymbol{\mu}) = f(v; \boldsymbol{\mu}) - a(u^{hp{\rm RB}}_N(\boldsymbol{\mu}), v; \boldsymbol{\mu}), \quad \forall v \in X. \tag{5.2}$$

We denote the Riesz representation of the residual by $\mathcal{R}^{hp{\rm RB}}_N$; as an upper bound for the X-norm error $\|u(\boldsymbol{\mu}) - u^{hp{\rm RB}}_N(\boldsymbol{\mu})\|_X$, we define

$$\Delta^{hp{\rm RB}}_N(\boldsymbol{\mu}) \equiv \frac{\|\mathcal{R}^{hp{\rm RB}}_N(\boldsymbol{\mu})\|_X}{\alpha_{\rm LB}(\boldsymbol{\mu})}. \tag{5.3}$$

Lemma 3.1 now directly applies with an appropriate change of notation.

5.3. Offline-Online Decomposition. In the *offline* stage, we determine the partition of the parameter domain and construct the corresponding RB models and spaces as discussed above. For our purposes here, we assume a perfect binary tree. We also assume that the cardinality of the train sample over each of the subdomains is equal to n_{train}.

The offline cost derives from several components; it is crucial to note that since the initial "h"-refinement is performed for $\bar{N} = 1$, there is no additional cost associated with the intermediate models:

1. Snapshot Truth Solves. During the partition procedure, we must compute (at most) $N_{\max}K$ snapshots associated with the final *and* intermediate approximation spaces.

2. Reduced Basis Preprocessing. We must compute (at most) $K(Q_a N_{\max}^2 + Q_f N_{\max})$ truth inner products to form the parameter-independent "stiffness" matrices and loads (e.g., as in (3.12)) for the final and intermediate models.

3. Error Bound Preprocessing. We must compute (at most) $KN_{\max}Q_a + Q_f$ truth Poisson solves of the form (3.18) for the final and intermediate models. We must also compute (at most) $K(N_{\max}Q_a + Q_f)^2$ truth inner-products of the form (3.21) in order to evaluate the dual norm of the residual associated with the final and intermediate models.

4. Error Bound Evaulations. We must solve (at most) $n_{\text{train}}((K-1) + N_{\max})$ RB systems to obtain the residual coefficients and evaluate (at most) $n_{\text{train}}((K-1) + N_{\max}K)$ RB error bounds during the Greedy sampling for both the final and intermediate models. This results in (at most) $n_{\text{train}}((K-1) + N_{\max}K)(N_{\max}^3 + N_{\max}^2 Q^2)$ operations in total (to leading order).

The combined offline cost is thus $N_{\max}K(1 + Q_a) + Q_f$ truth solves, $K(N_{\max}Q_a + Q_f)^2 + K(Q_a N_{\max}^2 + Q_f N_{\max})$ truth inner-products and $n_{\text{train}}((K-1)+N_{\max}K)(N_{\max}^3 + N_{\max}^2 Q^2)$ operations to evaluate the error bounds.

For each model, we must construct and retain the parameter-independent data necessary to accomodate the efficient evaluation stage for the RB approximation and the associated *a posteriori* error bound, as discussed in §3.3 for the standard RB method. The online (permanent) storage requirement is $Q_a K$ matrices of maximum size $N_{\max} \times N_{\max}$ and Q_f vectors of maximum size $N_{\max}$; the storage associated with the RB error bounds is $K(N_{\max}Q_a + Q_f)^2/2$.

In the *online* stage, given any $\boldsymbol{\mu} \in \mathcal{D}$, we first determine the subdomain containing $\boldsymbol{\mu}$ via the binary search (4.13) in $\mathcal{O}(\log_2 K)$ operations. Thanks to the construction-evaluation decomposition, we can then, given $1 \le N \le N_{\max}$, assemble and solve the corresponding system of algebraic equations in $\mathcal{O}(QN^2)$ and $\mathcal{O}(N^3)$ operations, respectively, and compute the associated RB error bound in $\mathcal{O}(N^2Q^2)$ operations.

6. A Convection-Diffusion Model Problem.

6.1. Formulation and Truth Discretization. We now apply the "p"-, "h"- and "hp"-type RB methods to a steady convection-diffusion model problem parametrized by the angle and magnitute of the prescribed velocity field: Let $\boldsymbol{\mu} = (\mu_1, \mu_2)$ and define $\mathbf{V}(\boldsymbol{\mu}) = [\mu_2 \cos\mu_1, \mu_2 \sin\mu_1]^{\mathrm{T}}$. The governing equations for the exact field variable $u^e(\boldsymbol{\mu})$ are

$$-\nabla^2 u^e(\boldsymbol{\mu}) + \mathbf{V}(\boldsymbol{\mu}) \cdot \nabla u^e(\boldsymbol{\mu}) = 10 \quad \text{in } \Omega, \tag{6.1}$$

$$u^e(\boldsymbol{\mu}) = 0 \quad \text{on } \partial\Omega. \tag{6.2}$$

The physical domain is the circle $\Omega = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le 2\}$.

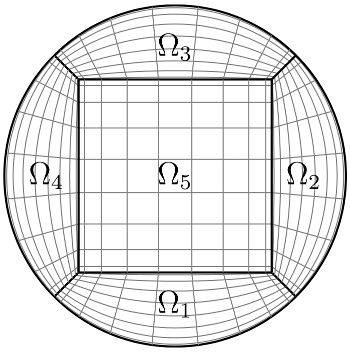

FIGURE 6.1. *The circular physical domain partitioned into five spectral elements.*

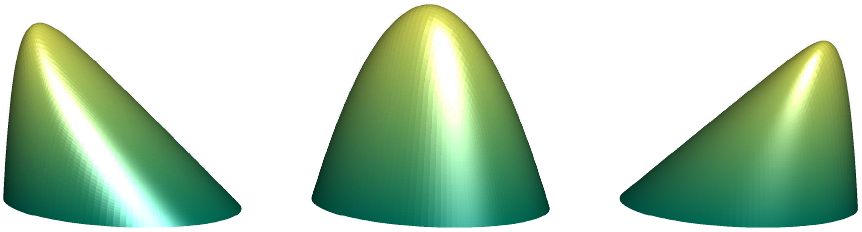

FIGURE 6.2. *Solutions to* (6.7) *for different parameter values* $\boldsymbol{\mu} = (\pi, 10)$ *(left),* $\boldsymbol{\mu} = (0,0)$ *(middle), and* $\boldsymbol{\mu} = (0,10)$ *(right).*

We next define for all $w, v \in X^e \equiv X^e(\Omega) \equiv H_0^1(\Omega)$ the parametrized bilinear form

$$\begin{aligned} a(w,v;\boldsymbol{\mu}) &\equiv \int_\Omega \nabla w \cdot \nabla v \,\mathrm{d}\Omega + \int_\Omega \big(\mathbf{V}(\boldsymbol{\mu}) \cdot \nabla w\big) v \,\mathrm{d}\Omega \\ &\equiv \int_\Omega \nabla w \cdot \nabla v \,\mathrm{d}\Omega + \mu_2 \cos\mu_1 \int_\Omega \frac{\partial w}{\partial x} v \,\mathrm{d}\Omega + \mu_2 \sin\mu_1 \int_\Omega \frac{\partial w}{\partial y} v \,\mathrm{d}\Omega, \end{aligned} \tag{6.3}$$

and the linear functional

$$f(v) \equiv f(v;\boldsymbol{\mu}) \equiv 10 \int_\Omega v \,\mathrm{d}\Omega; \tag{6.4}$$

thus (2.1) obtains for $Q_a = 3$ and $Q_f = 1$. We can then state the exact problem in the standard variational form: Given any $\boldsymbol{\mu} \in \mathcal{D}$, find $u^e \in X^e$ such that

$$a(u^e(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = f(v), \quad \forall v \in X^e. \tag{6.5}$$

Note that for this particular problem, $a_s(w,v;\boldsymbol{\mu}) = \int_\Omega \nabla w \cdot \nabla v \,\mathrm{d}\Omega$ is parameter-independent; thus $a(v,v;\boldsymbol{\mu}) = \|v\|_X^2$ for all $v \in X^e$ and we may choose $\alpha_{\mathrm{LB}} \equiv 1$ as the coercivity lower bound.

Next, we introduce a truth spectral element space $X \equiv X^{\mathcal{N}}(\Omega) \subset X^e(\Omega)$ of dimension $\mathcal{N} = 481$ based on five spectral elements of order ten: we introduce a computational domain $\hat{\Omega} = (-1,1)^2$ and standard transfinite mappings $\mathcal{F}_i : \hat{\Omega} \to \Omega_i$, $1 \le i \le 5$, [9]; we then define

$$X \equiv X^{\mathcal{N}}(\Omega) = \{v \in H_0^1(\Omega) : v|_{\Omega_i} \circ \mathcal{F}_i \in \mathbb{P}^{10}(\hat{\Omega}), 1 \le i \le 5\}, \tag{6.6}$$

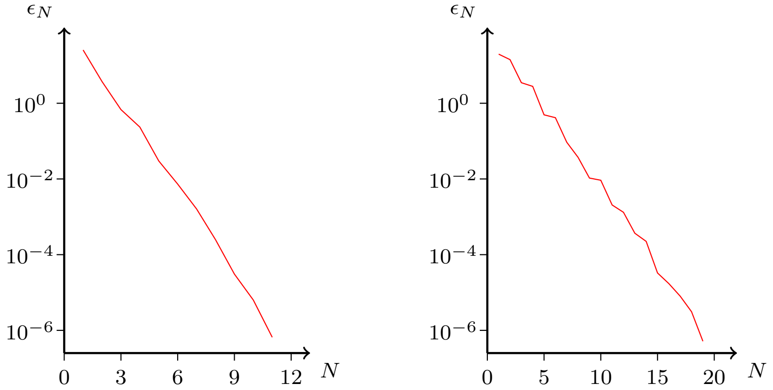

Figure 6.3. *Standard RB (“p”-type) convergence results: ϵ_N as a function of N for the one-parameter cases $\mathcal{D} = \mathcal{D}_{\mathrm{I}}$ (left) and $\mathcal{D} = \mathcal{D}_{\mathrm{II}}$ (right).*

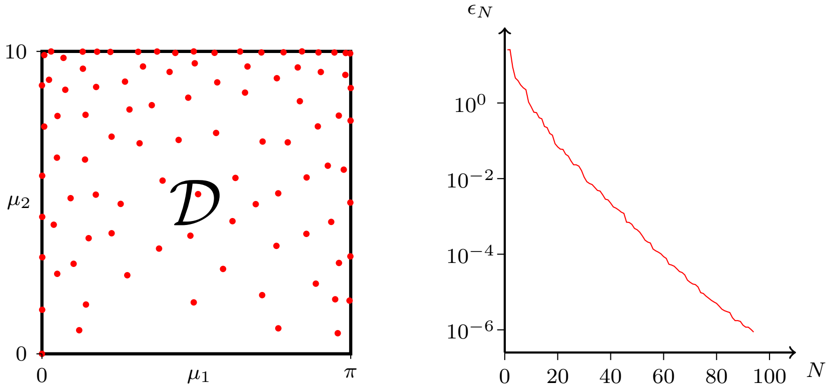

Figure 6.4. *Greedy parameter choices (left) and associated standard RB (“p”-type) convergence results (right) for the two-parameter case $\mathcal{D} = \mathcal{D}_{\mathrm{III}}$.*

where $\mathbb{P}^{10}(\hat{\Omega})$ denotes the space of polynomials of degree 10 (in each spatial direction) over $\hat{\Omega}$. The truth discretization of (6.5) reads: Given any $\boldsymbol{\mu} \in \mathcal{D}$, find $u(\boldsymbol{\mu}) \in X$ such that

$$a(u(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = f(v), \quad \forall v \in X. \tag{6.7}$$

In Figure 6.2, we plot the solution of (6.7) for three different parameter values. Clearly, the three solutions have a very different structure—this particular problem is thus a good candidate for “hp” treatment.

We define three parameter domains:

$$\mathcal{D}_{\mathrm{I}} \equiv \{0\} \times [0, 10], \quad \mathcal{D}_{\mathrm{II}} \equiv [0, \pi] \times \{10\}, \quad \mathcal{D}_{\mathrm{III}} \equiv [0, \pi] \times [0, 10]; \tag{6.8}$$

we shall thus consider $P = 1$ ($\mathcal{D}_{\mathrm{I}}$ or $\mathcal{D}_{\mathrm{II}}$) or $P = 2$ ($\mathcal{D}_{\mathrm{III}}$) parameters.

6.2. “p”-type RB Approximation Results. In this section, we present the standard (“p”-type) RB convergence results for our model problem.

We introduce uniformly distributed random train samples $\Xi_{\mathrm{I}} \subset \mathcal{D}_{\mathrm{I}}$, $\Xi_{\mathrm{II}} \subset \mathcal{D}_{\mathrm{II}}$, and $\Xi_{\mathrm{III}} \subset \mathcal{D}_{\mathrm{III}}$ of size 10^3, 10^3 and 10^4, respectively. We recall that $\epsilon_N = \max_{\boldsymbol{\mu} \in \Xi} \Delta_N(\boldsymbol{\mu})$

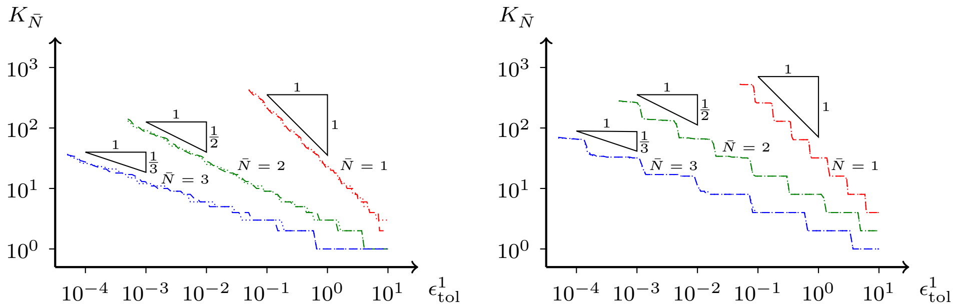

FIGURE 6.5. *"h"-type RB convergence results: $K_{\bar{N}}(\epsilon^1_{\rm tol})$ for $\bar{N} = 1$, $\bar{N} = 2$ and $\bar{N} = 3$ for the one-parameter cases $\mathcal{D} = \mathcal{D}_{\rm I}$ (left) and $\mathcal{D} = \mathcal{D}_{\rm II}$ (right). Both Euclidian distance (dotted lines) and the* a posteriori *error bound (dashed lines) are considered for the proximity function.*

is the maximum X-norm error bound over the train sample associated with the space X_N. In Figure 6.3, we plot ϵ_N as a function of N for the two one-parameter cases $\mathcal{D} = \mathcal{D}_{\rm I}$ and $\mathcal{D} = \mathcal{D}_{\rm II}$: we note that N can be quite small even for $\epsilon_N \approx 10^{-6}$. In Figure 6.4 (right), we plot ϵ_N for the two-parameter case $\mathcal{D} = \mathcal{D}_{\rm III}$. The quite poor convergence of the "p"-type RB is not surprising given the very different solution structures obtained for different parameter values; variations in μ_1 are particularly difficult to resolve—as indicated in Figure 6.3—due to the effect of the location of the boundary layer. In Figure 6.4 (left) we present the parameters chosen by the greedy algorithm: the points are clearly denser for larger velocities—which yield thinner boundary layers.

6.3. "h"-type RB Approximation Results. We now present convergence results for a pure "h"-type RB approximation; the dimension of the approximation spaces is thus fixed. The convergence results are obtained by first specifying the desired tolerance $\epsilon^1_{\rm tol}$ as well as the RB space dimension $\bar{N}$, the initial train sample $\Xi_{(1)}$ and the initial anchor point $\hat{\boldsymbol{\mu}}_{(1)}$; we then perform $h\mathrm{RB}(\Xi_{(1)}, \hat{\boldsymbol{\mu}}_{(1)}, \bar{N}, \epsilon^1_{\rm tol})$. Given $\bar{N}$, we let $K_{\bar{N}}(\epsilon^1_{\rm tol})$ denote the number of subdomains in the partition for specified $\epsilon^1_{\rm tol}$.

We start with the one-parameter cases $\mathcal{D} = \mathcal{D}_{\rm I}$ and $\mathcal{D} = \mathcal{D}_{\rm II}$. In both cases, the initial train samples consist of 100 random points; the initial anchor point is $\hat{\boldsymbol{\mu}}_{(1)} = (0, 0)$. In Figure 6.5 we present $K_{\bar{N}}(\epsilon^1_{\rm tol})$ for $\bar{N} = 1, 2, 3$ for each of the two cases. The proximity function is $d_{B_l}(\boldsymbol{\mu}) = \|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}_{B_l}\|_2$ (dotted) and $d_{B_l}(\boldsymbol{\mu}) = \Delta_{B_l}(\boldsymbol{\mu})$ (dashed lines): we observe the choice of the proximity function has little impact on the results. We indicate the slopes for first, second and third order convergence: for the $\bar{N} = 1$ approximation, the convergence rates are in good agreement with the theoretical result (4.28); for the $\bar{N} > 1$ approximations, the convergence is approximately $\bar{N}$th order and hence in agreement with our conjecture (4.38).

We next consider the two-parameter case $\mathcal{D} = \mathcal{D}_{\rm III}$. The initial train sample $\Xi_{(1)}$ consist of 10^3 random points; the initial anchor point is $\hat{\boldsymbol{\mu}}_{(1)} = (0, 0)$. In Figure 6.6 we present $K_{\bar{N}}(\epsilon^1_{\rm tol})$ for $\bar{N} = 1$ and $\bar{N} = 4$. The proximity function is $d_{B_l}(\boldsymbol{\mu}) = \|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}_{B_l}\|_2$ (solid lines) and $d_{B_l}(\boldsymbol{\mu}) = \Delta_{B_l}(\boldsymbol{\mu})$ (dashed lines): now the choice of the proximity function has some, but very slight, impact on the results but only for the first order approximation. We indicate the slopes for 1/2 and first order convergence; we achieve roughly $K_1 \sim (\epsilon^1_{\rm tol})^{-2}$ and $K_4 \sim (\epsilon^1_{\rm tol})^{-1}$: these results support our conjectures (4.39) and (4.40).

Finally, we empirically examine the depth of the associated binary trees. Ideally,

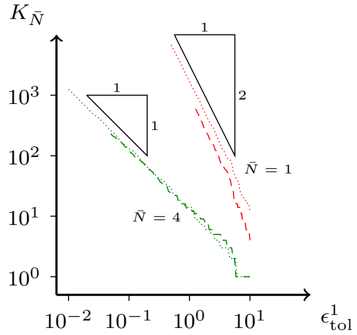

FIGURE 6.6. *"h"-type RB convergence results: $K_{\bar{N}}(\epsilon^1_{\rm tol})$ for $\bar{N}=1$ and $\bar{N}=4$ for the two-parameter case $\mathcal{D}=\mathcal{D}_{\rm III}$. Both Euclidian distance (dotted lines) and the* a posteriori *error bound (dashed lines) are considered for the proximity function.*

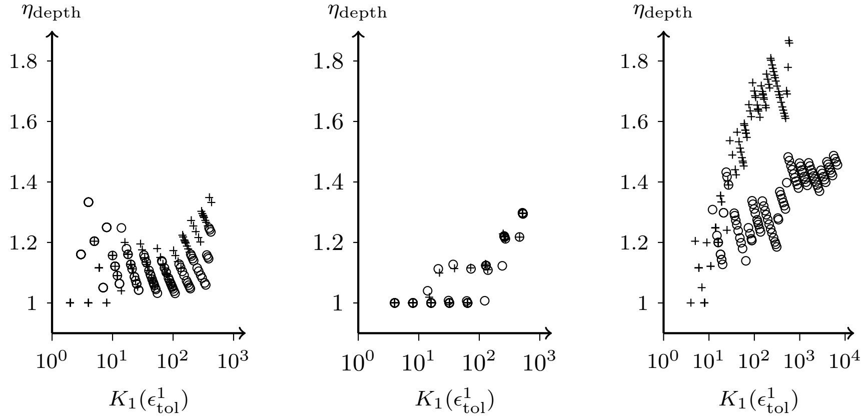

FIGURE 6.7. *Relative tree depths $\eta_{\rm depth}$ as functions of the number of subdomains (leaf nodes) $K_{\bar{N}}(\epsilon^1_{\rm tol})$ for 'the 'h"-type approximation with $\bar{N}=1$ for each parametrization $\mathcal{D}=\mathcal{D}_{\rm I}$ (left), $\mathcal{D}=\mathcal{D}_{\rm II}$ (middle), and $\mathcal{D}=\mathcal{D}_{\rm III}$ (right). Both Eucledian distance (○) and the* a posteriori *error bound (+) are considered for the proximity function.*

we would like the relative tree depth (4.14) to be a constant close to unity; the search (4.13) in this case is an efficient $\log_2 K$ operations binary search. In Figure 6.7, we plot the relative tree depth against the number of subdomains for the approximation with $\bar{N}=1$ for each of our three parametrizations. (Note the scatter in the plots is induced by the range of $\epsilon^1_{\rm tol}$ considered.) Although from these results it is difficult to reach general conclusions, the relative tree depths are all fairly close to unity and increase with increasing K only very modestly even for $1 \leq K \leq 10^4$.

6.4. "*hp*"-type Approximation Results. We now present convergence results for an "*hp*"-type RB approximation. For a partition with K subdomains, let $\bar{\Xi}$ denote the union of the associated K train samples; we then define $\epsilon_N^{hp\rm RB} \equiv \max_{\boldsymbol{\mu}\in\bar{\Xi}} \Delta_N^{hp\rm RB}(\boldsymbol{\mu})$.

We start with the one-parameter cases $\mathcal{D}=\mathcal{D}_{\rm I}$ and $\mathcal{D}=\mathcal{D}_{\rm II}$. For the case $\mathcal{D}=\mathcal{D}_{\rm I}$, we specify $\epsilon^1_{\rm tol}=5$ and $\epsilon^1_{\rm tol}=0.1$, which requires $K=4$ and $K=211$ subdomains, respectively; for the case $\mathcal{D}=\mathcal{D}_{\rm II}$, we specify $\epsilon^1_{\rm tol}=5$ and $\epsilon^1_{\rm tol}=0.1$, which requires

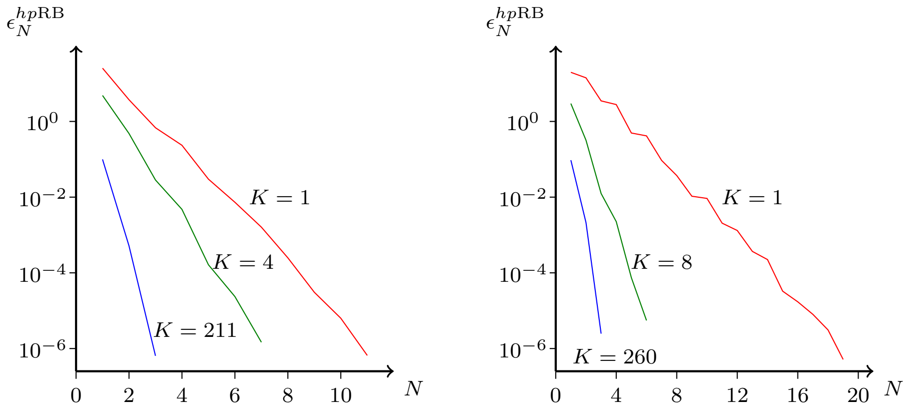

FIGURE 6.8. *"hp"-type RB convergence results: $\epsilon_N^{hp\mathrm{RB}}$ as a function of N for the one-parameter cases $\mathcal{D} = \mathcal{D}_{\mathrm{I}}$ (left) and $\mathcal{D} = \mathcal{D}_{\mathrm{II}}$ (right).*

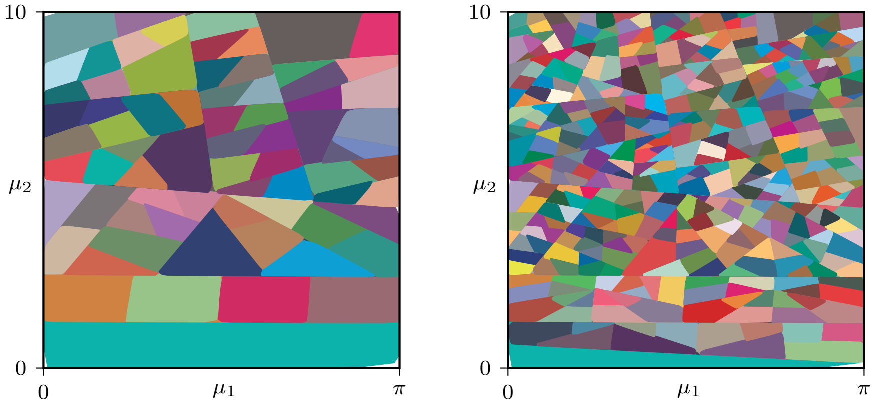

FIGURE 6.9. *Parameter domain partitions for the case $\mathcal{D} = \mathcal{D}_{\mathrm{III}}$. The numbers of subdomains are $K(\epsilon_{\mathrm{tol}}^1) = 72$ for $\epsilon_{\mathrm{tol}}^1 = 5$ (left) and $K(\epsilon_{\mathrm{tol}}^1) = 417$ for $\epsilon_{\mathrm{tol}}^1 = 2$ (right).*

$K = 8$ and $K = 260$ subdomains, respectively. Here, we use $d_{B_l} = \|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}_{B_l}\|_2$ as the proximity function; the initial train sample consist of 100 random points; the initial anchor point is $\hat{\boldsymbol{\mu}}_{(1)} = (0, 0)$. In Figure 6.8 we plot $\epsilon_N^{hp\mathrm{RB}}$ as functions of N for each of the two parametrizations. *Given* any error bound tolerance, we note a significant reduction in the required approximation space dimension (in any subdomain) when compared to a standard RB ($K = 1$) approximation. Of course, the *total* number of snapshots NK (for any given tolerance) will *increase* with K: the greater suitability of local snapshots does not compensate for lower order in terms of global approximation properties.

We next consider the two-parameter case $\mathcal{D} = \mathcal{D}_{\mathrm{III}}$. We use $d_{B_l}(\boldsymbol{\mu}) = \|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}_{B_l}\|_2$ as the proximity function; the initial train sample consist of 10^3 random points; the initial anchor point is $\hat{\boldsymbol{\mu}}_{(1)} = (0, 0)$. In Figure 6.9 we show partitions of the parameter domain for specified $\epsilon_{\mathrm{tol}}^1 = 5$ and $\epsilon_{\mathrm{tol}}^1 = 2$: we obtain $K = 72$ and $K = 417$ subdomains, respectively. We note—similarly to the "p"-type greedy parameter choices in Figure 6.3 (left)—that the subdomains are smaller for larger velocities. In Figure 6.10, we

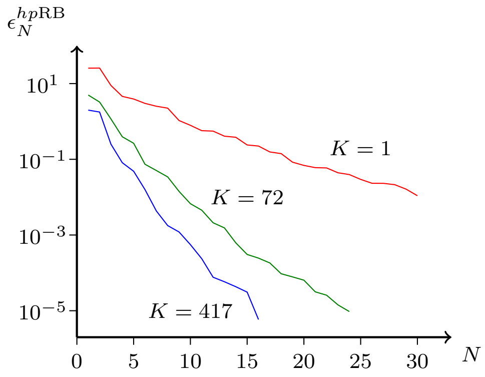

FIGURE 6.10. *Convergence of "hp"-type RB for the two-parameter case $\mathcal{D} = \mathcal{D}_{\mathrm{III}}$.*

TABLE 6.1

Operation count and storage requirement for the "hp"-type RB with $K = 72$ relative to that of the standard RB ($K = 1$) for the two-parameter case $\mathcal{D} = \mathcal{D}_{\mathrm{III}}$.

	$\epsilon^2_{\mathrm{tol}} = 10^{-2}$	$\epsilon^2_{\mathrm{tol}} = 10^{-3}$	$\epsilon^2_{\mathrm{tol}} = 10^{-4}$
Offline Rel. Cost	16.6	15.7	15.9
Online Rel. Cost	0.069	0.053	0.047
Online Rel. Storage	3.96	3.51	3.58

plot for each of the two partitions in Figure 6.9 the maximum error bound $\epsilon_N^{hp\mathrm{RB}}$ as a function of N; we include the results for the standard RB approximation ("*p*"-type or "*hp*"-type with $K = 1$) as well. Again, the order reduction is significant.

In Table 6.1 and Table 6.2 we summarize for $K = 72$ and $K = 417$ subdomains, respectively, the offline and online performance of the "*hp*" approach *relative* to that of the standard RB method. For given tolerances $\epsilon^2_{\mathrm{tol}}$, we report in the three rows of the tables the relative number of truth solves, the relative number of operations for online evaluation of the RB approximation and RB error bound, and relative online storage, respectively. The reported values are based on the theoretical operation count and storage, which we recall here. For N basis functions and $K \geq 1$ subdomains the offline number of truth solves is $KN(1 + Q_a) + Q_f$.[1] The online operation count is roughly $2N^3/3$ for the RB solution and $2(Q_aN + Q_f)^2$ operations for the RB error bound; we neglect the $\mathcal{O}(QN^2)$ cost of forming the RB system and the $\mathcal{O}(\log_2 K)$ cost of finding the correct subdomain via the binary search. The online (permanent) storage requirement is dominated by $\mathcal{O}(KQ^2N^2)$.

Admittedly, the "*hp*" approach requires more truth solves—a larger offline cost—than the standard method. However, the online computational savings are significant: the online cost is about 6-8 percent of that of the standard RB method in our example with $K = 72$, and only about 2-4 percent in our example with $K = 417$. The online storage requirement is somewhat larger with the "*hp*" approach, though in general the storage requirements are quite modest—the N^2 scaling moderates the growth due to K.

[1] We assume that the truth solves constitute the most expensive part of the total offline cost. In fact, this assumption *favors* the standard "*p*"-type RB; the error bound sampling is superlinear in N and thus scales more advantageously for the "*hp*"-approach. We thus expect in particular for n_{train} large that the Offline Rel. Cost. will be lower than reported in the tables.

Table 6.2

Operation count and storage requirement for the "hp"-type RB with $K = 417$ relative to that of the standard RB ($K = 1$) for the two-parameter case $\mathcal{D} = \mathcal{D}_{\rm III}$.

	$\epsilon^2_{\rm tol} = 10^{-2}$	$\epsilon^2_{\rm tol} = 10^{-3}$	$\epsilon^2_{\rm tol} = 10^{-4}$
Offline Rel. Cost	65.8	61.9	63.3
Online Rel. Cost	0.032	0.025	0.019
Online Rel. Storage	10.8	9.39	9.81

7. Concluding Remarks.

The "*hp*"-type RB method has been shown to significantly reduce the online computational cost. On the other hand, the new approach is more expensive than the standard ("*p*"-type) RB method in the offline stage; hence we must trade offline cost for online performance. However, the online effort is often our main concern in the real-time or many-query contexts.

We expect the new approach to be particularly beneficial for problems for which the solution structure is very different in different parts of the parameter domain. While our model problem was specifically constructed to exhibit this property, there are many realistic problems which exhibit similar behavior. As an example, we mention an application of RB to the solution of the Fokker-Planck equation [14]; here, the solution is required for many different parameters, but the required ("*p*"-type) RB spaces are rather large. Also of interest are problems which exhibit non-smooth parameter dependence—the "*hp*"-approach should automatically refine the parameter domain around singularities and hence work better than the standard approach;

There are several opportunities for future work. First, we can improve the algorithm: at present the parameter domain partition is rather sensitive to the choice of $\epsilon^1_{\rm tol}$; we can also reduce the number of truth solves—the offline effort—if we exploit the fact that the greedy algorithm often chooses anchor points near subdomain boundaries—hence some parameter values could be shared between models (c.f. Remark 2). Second, we can generalize our approach to POD-Greedy sampling [11] for parabolic problems [8]: the critical new ingredient is proper balance between additional POD modes and additional Greedy parameter values in the initial subdivision process. Third, we can extend the approach to quadratically nonlinear problems such as the incompressible Navier-Stokes equations [13]—in this case the "*hp*"-approach is particularly advantageous since the (online) computation of the error bound requires $\mathcal{O}(N^4)$ operations for N basis functions and hence the "smaller N for larger K trade" is particularly favorable. Finally, we mention that the offline stage of the "*hp*" approach is readily parallellizable—we can subdivide the parameter domain along each branch of the associated binary tree independently.

Acknowledgements. We acknowledge helpful discussions with D. Knezevic, N. C. Nguyen, S. Boyaval, and B. Haasdonk during this work.

REFERENCES

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

[2] D. Amsallem, J. Cortial, and C. Farhat, *On-Demand CFD-Based Aeroelastic Predictions Using a Database of Reduced-Order Bases and Models*, in 47th AIAA Aerospace Sciences Meeting Including The New Horizons Forum and Aerospace Exposition, January 2009.

[3] D. Amsallem and C. Farhat, *Interpolation Method for Adapting Reduced-Order Models and Application to Aeroelasticity*, AIAA Journal, 46 (2008), pp. 1803–1813.

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

[5] S. Boyaval, *Reduced-basis approach for homogenization beyond the periodic setting*, Multiscale Model. Simul., 7 (2008), pp. 466–494.

[6] S. Boyaval, C. Le Bris, Y. Maday, N. C. Nguyen, and A. T Patera, *A reduced basis approach for variational problems with stochastic parameters: Application to heat conduction with variable robin coefficient*, Computer Methods in Applied Mechanics and Engineering, 198 (2009), pp. 3187 – 3206.

[7] A. Buffa, Y. Maday, A. T. Patera, C. Prud'homme, and G. Turinici, *A Priori Convergence of the Greedy Algorithm for the Parametrized Reduced Basis.* M2AN, submitted 2009.

[8] J. L. Eftang, A. T. Patera, and E. M. Rønquist, *An "hp" Certified Reduced Basis Method for Parametrized Parabolic Partial Differential Equations*, in ICOSAHOM Proceedings, 2009.

[9] William J. Gordon and Charles A. Hall, *Construction of curvilinear co-ordinate systems and applications to mesh generation*, Internat. J. Numer. Methods Engrg., 7 (1973), pp. 461–477.

[10] M. A. Grepl, Y. Maday, N. C. Nguyen, and A. T. Patera, *Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations*, M2AN Math. Model. Numer. Anal., 41 (2007), pp. 575–605.

[11] B. Haasdonk and M. Ohlberger, *Reduced basis method for finite volume approximations of parametrized linear evolution equations*, M2AN Math. Model. Numer. Anal., 42 (2008), pp. 277–302.

[12] D. B. P. Huynh, G. Rozza, S. Sen, and A. T. Patera, *A successive constraint linear optimization method for lower bounds of parametric coercivity and inf-sup stability constants*, C. R. Math. Acad. Sci. Paris, 345 (2007), pp. 473–478.

[13] D. J. Knezevic, N. C. Nguyen, and A. T. Patera, *Reduced Basis Approximation and* a posteriori *Error Estimation for the Parametrized Unsteady Boussinesq Equations.* M3AS, Submitted 2009.

[14] D. J. Knezevic and A. T. Patera, *A Certified Reduced Basis Method for the Fokker-Planck Equation of Dilute Polymeric Fluids: FENE Dumbbells in Extensional Flow.*, SIAM Journal of Scientific Computing, ((submitted 2009)).

[15] N. C. Nguyen, G. Rozza, D. B. P. Huynh, and A. T. Patera, *Reduced Basis Approximation and* a posteriori *Error Estimation for Parametrized Parabolic PDEs; Application to Real-Time Bayesian Parameter Estimation*, in Computational Methods for Large Scale Inverse Problems and Uncertainty Quantifications, Biegler, Biro, Ghattas, Heinkenschloss, Keyes, Mallick, Tenorio, van Bloemen Waanders, and Willcox, eds., John Wiley & and Sons, UK, Submitted 2009.

[16] N. C. Nguyen, G. Rozza, and A. T. Patera, *Reduced Basis Approximation and* a posteriori *Error Estimation for the Time-Dependent Viscous Burgers' Equation*, Calcolo, 46 (2009), pp. 157–185. Calcolo, 2008 (submitted).

[17] A. K. Noor and J. M. Peters, *Reduced basis technique for nonlinear analysis of structures*, AIAA Journal, 18 (1980), pp. 455–462.

[18] T. A. Porsching, *Estimation of the error in the reduced basis method solution of nonlinear equations*, Math. Comp., 45 (1985), pp. 487–496.

[19] W. C. Rheinboldt, *On the theory and error estimation of the reduced basis method for multi-parameter problems*, Nonlinear Anal. Theory Methods Appl., 21 (1993), pp. 849–858.

[20] 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 (2008), pp. 229–275.

[21] 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.

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