2nd ISSMO/AIAA Internet Conference

on Approximations and Fast Reanalysis in Engineering Optimization

May 25 – June 2, 2000

Reduced-Basis Output-Bound Methods for Elliptic Partial Differential Equations[1]

Dimitrios V. Rovas
Department of Mechanical Engineering, Room 3-243,
Massachusetts Institute of Technology, Cambridge, MA,
02139-4307, rovas@mit.edu
Anthony T. Patera
Department of Mechanical Engineering, Room 3-264,
Massachusetts Institute of Technology, Cambridge, MA,
02139-4307, patera@mit.edu

Abstract[2] We present a two-stage off-line/on-line blackbox reduced-basis output bound method for the prediction of *outputs of interest* associated with elliptic partial differential equations with affine parameter dependence. The method is characterized by (i) Galerkin projection onto a reduced-basis space comprising solutions at selected points in parameter space, and (ii) a *rigorous output error bound* based on the dual norm of the resulting residual. The computational complexity of the on-line stage of the procedure scales only with the dimension of the reduced-basis space and the parametric complexity of the partial differential operator. The method is thus both efficient and certain: thanks to the *a posteriori* error bounds, we may safely retain only the minimal number of modes necessary to achieve the prescribed accuracy in the output of interest. The technique is particularly appropriate for applications such as design, optimization, and control, in which repeated and rapid evaluation of the output is required; in the limit of many evaluations, the method can be *several orders of magnitude faster* than standard (finite element) approximation. To illustrate the method, we consider the design of a thermal fin.

1. Motivation

To motivate and illustrate our methods we consider a specific example, a thermal fin. The fin, shown in Figure 1, consists of a central "post" and four horizontal plates which we denote "subfins;" the fin conducts heat from a prescribed flux "source" at the root through the large-surface-area subfins to surrounding flowing air. The fin is characterized by seven design parameters, or "inputs," $\mu \in \mathcal{D} \subset \mathbb{R}^{P=7}$, where $\mu^i = k^i, i = 1, \ldots, 4, \mu^5 = \mathrm{Bi}, \mu^6 = L$, and $\mu^7 = t$. Here k^i is the thermal conductivity of the i^{th} subfin (normalized relative to the post conductivity); Bi is the Biot number, a nondimensional heat transfer coefficient reflecting convective transport to the air at the fin surfaces; and L and t are the length and thickness of the subfins (normalized relative to the post width). The performance metric, or "output," $s \in \mathbb{R}$, is chosen to be the average temperature of the fin root normalized by the prescribed heat flux into the fin root, Γ_{root}.

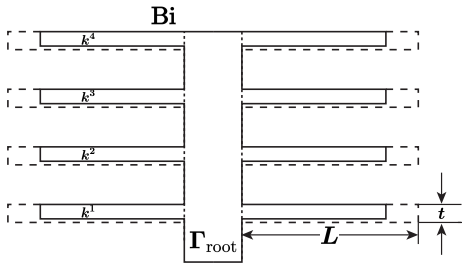

Figure 1

We can express our input-output relationship as $s = \ell^O(u(\mu))$, where $\ell^O(v)$ is a (continuous) linear functional — $\ell^O(v) = \int_{\Gamma_{\text{root}}} v$ — and $u(\mu)$ is the temperature distribution within the fin. (The

[1]The material presented in this article is an expository version of work performed in collaboration with Dr. Luc Machiels of Lawrence Livermore National Laboratory and Professor Yvon Maday of University of Paris VI and reported in greater detail in references [1, 2, 3, 4]. We also thank Professor Jaime Peraire of MIT, Professor Einar Rønquist of Norwegian University of Science and Technology, Mr. Roland Von Kaenel of EPFL, and Ms. Shidrati Ali of National University of Singapore–Singapore-MIT Alliance for helpful comments. The work is supported by AFOSR, NASA Langley Research Center, and the Singapore-MIT Alliance.

[2]This article will appear in the next issue of the SIAM SIAG/OPT Views-and-News (ed. N. Alexandrov).

temperature field is of course a function of the spatial coordinate, $\mathbf{x}$; we explicitly indicate this dependence only as needed.) The temperature distribution $u(\mu) \in Y$ satisfies the weak form of the elliptic partial differential equation describing heat conduction in the fin,

$$a(u, v; \mu) = \ell(v), \forall v \in Y; \tag{1}$$

$a(u, v; \mu)$ is the weak representation of the Laplacian, and $\ell(v)$ reflects the prescribed heat flux at the root. Here Y is the appropriate Hilbert space with associated inner product $(\cdot, \cdot)_Y$ and induced norm $\|\cdot\|_Y$ [3]. The bilinear form $a(\cdot, \cdot; \mu)$ is symmetric, $a(w, v; \mu) = a(v, w; \mu), \forall w, v \in Y^2, \forall \mu \in \mathcal{D}$; uniformly continuous, $|a(w, v; \mu)| \leq \gamma \|w\|_Y \|v\|_Y, \forall w, v \in Y^2, \forall \mu \in \mathcal{D}$; and coercive, $\alpha \|v\|_Y^2 \leq a(v, v; \mu), \forall v \in Y, \forall \mu \in \mathcal{D}$. Here α and γ are positive real constants. Finally, the form $\ell(v)$ is a linear bounded functional; for our choice of scaling and output, $\ell^O(v) = \ell(v)$, which we will exploit to simplify the exposition.

It is readily shown that our form a can be expressed as

$$a(w, v; \mu) = \sum_{q=1}^{Q} \sigma^q(\mu) a^q(w, v), \forall w, v \in Y^2, \forall \mu \in \mathcal{D}, \tag{2}$$

for appropriately chosen functions $\sigma^q : \mathcal{D} \to \mathbb{R}$ and associated μ-independent bilinear forms $a^q : Y \times Y \to \mathbb{R}$, $q = 1, \ldots, Q$: the parameter dependence is thus "affine" or "separable." Note that we pose our problem on a *fixed* fin reference domain Ω in order to ensure that the parametric dependence on geometry — L and t — enters through $a(\cdot, \cdot; \mu)$ and ultimately the $\sigma^q(\mu)$. For our particular problem, $Q = 15$; if we freeze (fix) all parameters except L and t (such that $P_{\text{eff}} = 2$), $Q = 8$; if we freeze only L and t (such that $P_{\text{eff}} = 5$), $Q = 6$.

Our goal is to construct an approximation to u, $\tilde{u}$, and hence approximation to s, $\tilde{s} = \ell^O(\tilde{u})$, which is (i) *certifiably* accurate, and (ii) very efficient in the limit of *many evaluations*. By the latter we mean that, following an initial *fixed* investment, the additional *incremental* cost to evaluate $\tilde{s}(\mu)$ for any new $\mu \in \mathcal{D}$ is much less than the effort required to directly compute $s(\mu) = \ell^O(u(\mu))$ by (say) standard finite element approximation. This definition of efficiency is particularly appropriate in the context of design, optimization, and control, in which we require very rapid response and many output evaluations.

2. Reduced-Basis Approximation

Reduced-basis methods (e.g., [5, 6, 7]) are a now-classical approach that are a special "parameter-space" version of weighted–residual (here Galerkin) approximation. To define the (or a) reduced–basis procedure, we first introduce a sample in parameter space, $S^N = \{\mu_1, \ldots, \mu_N\}$, and associated reduced-basis space $W^N = \text{span}\{\zeta_1 \equiv u(\mu_1), \zeta_2 \equiv u(\mu_2), \ldots, \zeta_N \equiv u(\mu_N)\}$, where $u(\mu_i)$ satisfies (1) for $\mu = \mu_i \in \mathcal{D}$ (note μ^i refers to the i^{th} component of the P–tuple μ, whereas μ_i refers to the i^{th} P–tuple in S^N). We then require our reduced-basis approximation to $u(\mu)$ for any given μ, $u^N(\mu) \in W^N \subset Y$, to satisfy

$$a(u^N(\mu), v; \mu) = \ell(v), \forall v \in W^N; \tag{3}$$

the reduced-basis approximation to $s(\mu)$ can subsequently be evaluated as $s^N(\mu) = \ell^O(u^N(\mu))$.

It is a simple matter to show that

$$\|u(\mu) - u^N(\mu)\|_Y \leq \sqrt{\frac{\gamma}{\alpha}} \min_{w^N \in W^N} \|u(\mu) - w^N\|_Y, \tag{4}$$

which states that our approximation is in some sense optimal in the Y norm. It can also be readily shown for our particular problem that

$$s(\mu) = s^N(\mu) + a(e^N(\mu), e^N(\mu); \mu), \tag{5}$$

where $e^N = u - u^N$. It follows from (4),(5), and the continuity of a that

$$|s(\mu) - s^N(\mu)| \leq \frac{\gamma^2}{\alpha} (\min_{w^N \in W^N} \|u(\mu) - w^N\|_Y)^2; \tag{6}$$

thus our output approximation is also optimal in some sense.

We must, of course, also understand the extent to which the best w^N in W^N can, indeed, approximate the requisite temperature distribution. The essential point is that, although W^N clearly does not have any approximation properties for *general*

[3] Here $Y = H^1(\Omega)$, the space of functions that are square integrable and that have square integrable first (distributional) derivatives over the fin reference domain Ω. The inner product $(w, v)_Y$ may be chosen to be $\int_\Omega \nabla w \cdot \nabla v + wv$.

functions in Y, simple interpolation arguments in parameter space suggest that W^N should approximate well $u(\mu)$ *even for very modest* N; indeed, exponential converge is obtained in N for sufficiently smooth μ-dependence (e.g., [6, 7]). It is for this reason that, even in high-dimensional (large P) parameter spaces, reduced-basis methods continue to perform well — much better than "non-state-space" direct interpolation of $(\mu, s(\mu))$ input-output pairs. In some sense, reduced-basis methods transform extensive parameter-space exploration from a problem into an opportunity.

We now turn to the computational issues. We first express the reduced-basis approximation as

$$u^N(\boldsymbol{x};\mu) = \sum_{j=1}^{N} u_j^N(\mu)\zeta_j(\boldsymbol{x}) = (\underline{u}^N(\mu))^T \underline{\zeta}(\boldsymbol{x}), \quad (7)$$

and choose for test functions $v = \zeta_i(\boldsymbol{x}), i = 1, \ldots, N$. We then insert these representations into (3) to yield the desired algebraic equations for $\underline{u}^N(\mu) \in \mathbb{R}^N$,

$$\sum_{j=1}^{N} a(\zeta_j, \zeta_i; \mu) u_j^N = \ell(\zeta_i), \quad i = 1, \ldots, N. \quad (8)$$

Equation (8) can be written in matrix form as

$$\underline{A}(\mu)\underline{u}^N(\mu) = \underline{L}, \quad (9)$$

where $\underline{A}(\mu) \in \mathbb{R}^{N\times N}$ is the SPD matrix with entries $A_{i,j}(\mu) = a(\zeta_j, \zeta_i; \mu), 1 \le i, j \le N$, and $\underline{L} \in \mathbb{R}^N$ is the "load" vector with entries $L_i = \ell(\zeta_i), 1 \le i \le N$.

We now evoke (2) to note that

$$A_{i,j}(\mu) = a(\zeta_j, \zeta_i; \mu) = \sum_{q=1}^{Q} \sigma^q(\mu) a^q(\zeta_j, \zeta_i) = \sum_{q=1}^{Q} \sigma^q(\mu) A_{i,j}^q, \quad (10)$$

where the matrices $\underline{A}^q \in \mathbb{R}^{N\times N}$ are given by $A_{i,j}^q = a^q(\zeta_j, \zeta_i), 1 \le i, j \le N, q = 1, \ldots, Q$. The off-line/on-line decomposition is now clear. In the *off-line* stage, we construct the $\underline{A}^q, q = 1, \ldots, Q$. In the *on-line* stage, for any given μ, we first form $\underline{A}$ from the $\underline{A}^q$ according to (10); we next invert (9) to find $\underline{u}^N(\mu)$; and we then compute $s^N(\mu) = \ell^O(u^N(\mu)) = \ell(u^N(\mu)) = (\underline{u}^N(\mu))^T \underline{L}$. As we shall see, N will typically be $O(10)$ for our particular problem. Thus, as required, the incremental cost to evaluate $s^N(\mu)$ for any given new μ is very small: $O(N^2Q)$ to form $\underline{A}(\mu)$; $O(N^3)$ to invert (the typically dense) $\underline{A}(\mu)$ system; and $O(N)$ to evaluate $s^N(\mu)$ from $\underline{u}^N(\mu)$.

But all is not well. The above *a priori* results tell us only that we are doing as well as possible; it does not tell us *how* well we are doing. In particular, the error in our output is not known, and hence the minimal number of basis functions required to satisfy the desired error tolerance can not be ascertained. As a result, either too many or too few functions are retained; the former results in computational inefficiency, the latter in uncertainty and unacceptably inaccurate predictions. We thus need *a posteriori* error bounds as well.

3. Output Bounds

To begin, we assume that we may find a function $g(\mu): \mathcal{D} \to \mathbb{R}_+$ and symmetric continuous coercive bilinear form $\hat{a}: Y \times Y \to \mathbb{R}$ such that

$$\underline{c}\|v\|_Y^2 \le g(\mu)\hat{a}(v, v) \le a(v, v; \mu), \forall v \in Y, \forall \mu \in \mathcal{D}, \quad (11)$$

for some real positive constant $\underline{c}$; for our thermal fin problem we can readily find a $g(\mu)$ and $\hat{a}(w, v)$ such that (11) is satisfied. The procedure is then simple: we first compute $\hat{e}(\mu) \in Y$ solution of

$$g(\mu)\hat{a}(\hat{e}(\mu), v) = R(v; \mu), \forall v \in Y, \quad (12)$$

where $R(v; \mu) \equiv \ell(v) - a(u^N, v; \mu)$ is the residual; we then evaluate our bounds as

$$s_-^N(\mu) = s^N(\mu), \quad s_+^N(\mu) = s^N(\mu) + \Delta^N(\mu), \quad (13)$$

where

$$\Delta^N(\mu) = g(\mu)\hat{a}(\hat{e}(\mu), \hat{e}(\mu)) \quad (14)$$

is the bound gap. The notion of output bounds is not restricted to reduced–basis approximations: it can also be applied within the context of standard (adaptive) finite element discretization as well as iterative (Krylov) solution strategies [8, 9].

We can then show that

$$s_-^N(\mu) \le s(\mu) \le s_+^N(\mu), \quad \forall N; \quad (15)$$

we thus have a *certificate of fidelity* for s^N — it is within $\Delta^N(\mu)$ of $s(\mu)$. The proof of (15) is simple. To prove the left inequality we need only appeal to (5) and the coercivity of a. To demonstrate

the right inequality we first note that $R(e^N(\mu);\mu) = \ell(e^N(\mu)) - a(u^N(\mu), e^N(\mu);\mu) = a(e^N(\mu), e^N(\mu);\mu)$, since $\ell(e^N(\mu)) = a(u, e^N(\mu);\mu)$ from (1) for $v = e^N(\mu)$; we next choose $v = e^N(\mu)$ in (12) to obtain $g(\mu)\hat{a}(\hat{e}(\mu), e^N(\mu)) = a(e^N(\mu), e^N(\mu);\mu)$; from this result and the right inequality of (11) we then have

$$\begin{aligned}\Delta^N(\mu) &\equiv g(\mu)\hat{a}(\hat{e},\hat{e}) \\ &= g(\mu)\hat{a}(\hat{e}-e^N,\hat{e}-e^N) + 2a(e^N,e^N) \\ &\quad - g(\mu)\hat{a}(e^N,e^N) \\ &\geq g(\mu)\hat{a}(\hat{e}-e^N,\hat{e}-e^N) + a(e^N,e^N);\end{aligned} \tag{16}$$

from (16) and the left inequality of (11) we thus conclude that $\Delta^N(\mu) \geq a(e^N, e^N)$; a comparison of (5) and (13) then completes the proof.

Our *a posteriori* bound result indicates that, through Δ^N, we can now ascertain the accuracy of our output prediction, which will in turn permit us to adaptively modify our approximation until any prescribed error tolerance is satisfied (see below). However, from the perspective of efficiency, it is also critical that $\Delta^N(\mu)$ be a *good* error estimator; a poor estimator will encourage us to unnecessarily refine an approximation which is, in fact, adequate. To prevent the latter we would like the effectivity $\eta^N(\mu) \equiv \Delta^N(\mu)/|s(\mu) - s^N(\mu)|$ to be order unity. For our problem it is simple to prove that $\eta^N(\mu) \leq \gamma/\underline{c}$, and thus $\eta^N(\mu)$ is certainly bounded independent of μ and N; in practice, effectivities are typically less than 10, which is adequate given the rapid convergence of reduced-basis approximations.

We now turn to the computational issues. We first note that, from (2) and (7), (12) can be rewritten as

$$\hat{a}(\hat{e}(\mu), v) = \frac{1}{g(\mu)}\left(\ell(v) - \sum_{q=1}^{Q}\sum_{j=1}^{N}\sigma^q(\mu)u_j^N(\mu)a^q(\zeta_j, v)\right), \quad \forall v \in Y. \tag{17}$$

We thus see from simple linear superposition that $\hat{e}(\mu)$ can be expressed as

$$\hat{e}(\mu) = \frac{1}{g(\mu)}(\hat{z}_0 + \sum_{q=1}^{Q}\sum_{j=1}^{N}\sigma^q(\mu)u_j^N(\mu)\hat{z}_j^q), \tag{18}$$

where $\hat{z}_0 \in Y$ satisfies $\hat{a}(\hat{z}_0, v) = \ell(v), \forall v \in Y$, and $\hat{z}_j^q \in Y, j = 1,\ldots,N, q = 1,\ldots,Q$, satisfies $\hat{a}(\hat{z}_j^q, v) = -a^q(\zeta_j, v), \forall v \in Y$. It then follows that we can express $\Delta^N(\mu)$ of (14) as

$$\begin{aligned}\Delta^N(\mu) = \frac{1}{g(\mu)}\Bigg[&\underbrace{\hat{a}(\hat{z}_0,\hat{z}_0)}_{c_0} + \\ &2\sum_{q=1}^{Q}\sum_{j=1}^{N}\sigma^q(\mu)u_j^N(\mu)\underbrace{\hat{a}(\hat{z}_0,\hat{z}_j^q)}_{\Lambda_j^q} + \\ &\sum_{q=1}^{Q}\sum_{q'=1}^{Q}\sum_{j=1}^{N}\sum_{j'=1}^{N}\sigma^q(\mu)\sigma^{q'}(\mu)u_j^N(\mu)u_{j'}^N(\mu)\underbrace{\hat{a}(\hat{z}_j^q,\hat{z}_{j'}^{q'})}_{\Gamma_{jj'}^{qq'}}\Bigg];\end{aligned} \tag{19}$$

$s_+^N(\mu)$ then directly follows from (13).

The off-line/on-line decomposition is now clear. In the *off-line* stage we compute $\hat{z}_0$ and $\hat{z}_j^q, j = 1,\ldots,N, q = 1,\ldots,Q$, and then the inner products c_0, Λ_j^q, and $\Gamma_{jj'}^{qq'}$ defined in (19). In the *on-line* stage, for any given new μ, and given $s^N(\mu)$ and $\underline{u}^N(\mu)$ as computed in the on-line stage of the output prediction process (Section 2), we first evaluate $\Delta^N(\mu)$ as

$$\begin{aligned}\Delta^N(\mu) = \frac{1}{g(\mu)}\Bigg[&c_0 + 2\sum_{q=1}^{Q}\sum_{j=1}^{N}\sigma^q(\mu)u_j^N(\mu)\Lambda_j^q + \\ &\sum_{q=1}^{Q}\sum_{q'=1}^{Q}\sum_{j=1}^{N}\sum_{j'=1}^{N}\sigma^q(\mu)\sigma^{q'}(\mu)u_j^N(\mu)u_{j'}^N(\mu)\Gamma_{jj'}^{qq'}\Bigg],\end{aligned} \tag{20}$$

and then evaluate $s_+^N(\mu) = s^N(\mu) + \Delta^N(\mu)$. The incremental cost to evaluate $s_+^N(\mu)$ for any given new μ is very small: $O(N^2Q^2)$.

4. Numerical Algorithm

In the simplest case we take our field and output approximations to be $\tilde{u}(\mu) = u^N(\mu)$ and $\tilde{s}(\mu) = s^N(\mu)$, respectively, for some given N, and then compute $\Delta^N(\mu)$ to assess the error. However, it is very easy to improve upon this recipe. To wit, we take $\tilde{u}(\mu) = u^{\tilde{N}}(\mu)$ and $\tilde{s}(\mu) = s^{\tilde{N}}(\mu)$, where $u^{\tilde{N}}(\mu)$ and $s^{\tilde{N}}(\mu)$ are the reduced-basis approximations associated with a subspace of W^N, $W^{\tilde{N}}$, in which we select only $\tilde{N}$ of our available basis functions. In practice, we include in $W^{\tilde{N}}$ the basis functions corresponding

to sample points μ_i closest to the new μ of interest; we continue to augment our space until $\Delta^{\tilde{N}}(\mu) \leq \varepsilon$ (and hence $|s(\mu) - s^{\tilde{N}}(\mu)| \leq \varepsilon$), where ε is the acceptable error in the output prediction. If we satisfy our criterion for $\tilde{N} \leq N$ the adaptive procedure is entirely contained within the on-line stage of the procedure; and the complexity of this stage is reduced (roughly) from $O(N^2Q + N^3 + N^2Q^2)$ to $O(\tilde{N}^2Q + \tilde{N}^3 + \tilde{N}^2Q^2)$ — often representing considerable savings. Note the critical role that our error bound plays in effecting this economy.

In practice — to ensure that the $\zeta_i, \hat{z}_0, \hat{z}_j^q$ are actually calculable — we replace the infinite–dimensional space Y with a very high dimensional "truth" space Y_T (e.g., a finite–element space associated with a very fine triangulation). It follows that our bounds are not in fact for s, but rather for $s_T = \ell^O(u_T)$, where $u_T \in Y_T$ satisfies $a(u_T, v; \mu) = \ell(v), \forall v \in Y_T$. The essential point is that Y_T may be chosen very conservatively — and hence the difference between s_T and s rendered arbitrarily small — since (i) the on–line work and storage is in fact *independent* of the dimension of Y_T, $\mathcal{N}$, and (ii) the off–line work will remain modest since N will typically be quite small. The uncertainty introduced by the truth approximation is thus minimal.

5. Results and Discussion

We first demonstrate the accuracy of the reduced–basis output prediction and the output bounds by considering the case $P_{\text{eff}} = 5$ in which $L = 2.5$ and $t = 0.25$ are fixed; the remaining parameters are permitted to vary in the domain $k^1, k^2, k^3, k^4, \text{Bi} \in \mathcal{D}_{\text{eff}} \equiv [0.1, 10]^4 \times [0.01, 1]$. The sample points for S^N are chosen randomly (uniformly) over $\mathcal{D}_{\text{eff}}$; the new value of μ to which we apply the reduced–basis approximation is taken to be $k^1 = 0.5, k^2 = 1.0, k^3 = 3.0, k^4 = 9.0, \text{Bi} = 0.1$ (similar results are obtained at other points in $\mathcal{D}_{\text{eff}}$). We present in Table 1 the actual error $|s(\mu) - s^N(\mu)|$; the estimated error $\Delta^N(\mu)$ (our strict upper bound for $|s(\mu) - s^N(\mu)|$); and the effectivity $\eta^N(\mu)$ (the ratio of the estimated and actual errors). We observe the high accuracy and rapid convergence of the reduced–basis prediction, even for this relatively high–dimensional parameter space (10 points corresponds to fewer than two points "in each direction"); and the very good accuracy (low effectivity) of our error bound $\Delta^N(\mu)$. The combination of high accuracy and certifiable fidelity permits us to proceed with an extremely low number of modes, with correspondingly low computational cost.

N	$\lvert s - s^N \rvert$	Δ^N	η^N
10	1.48×10^{-3}	2.34×10^{-2}	15.82
20	2.94×10^{-4}	2.59×10^{-3}	8.81
30	1.80×10^{-5}	3.09×10^{-4}	17.12
40	1.87×10^{-6}	2.45×10^{-5}	13.10
50	1.17×10^{-7}	2.08×10^{-6}	17.98

Table 1

As regards computational cost, in the limit of "infinitely many" evaluations, the calculation of $\tilde{s}(\mu)$ to within 0.1% of s_T is roughly 285 times faster than direct calculation of $s_T = \ell^O(u_T)$; here u_T is our underlying "truth" finite element approximation (see Section 4). The breakeven point — at which the reduced–basis approximation first becomes less expensive than direct evaluation of s_T — is roughly 142 evaluations. In making these comparisons we must of course not bias the conclusion: our "truth" approximation here is *not* overly fine; and our solution strategy for $u_T \in Y_T$ — an ILU–preconditioned sparsity–exploiting conjugate–gradient procedure — is quite efficient. The reduced–basis approach is much faster simply because the dimension of W^N, N, is much smaller than the dimension of Y_T, $\mathcal{N}$ (which more than compensates for the loss of sparsity in $\underline{A}$). For more difficult problems that require larger $\mathcal{N}$ — problems with more spatial structure, or in more complicated geometry, or in three dimensions — or that are not as amenable to fast solution methods on Y_T — problems with less "nice" mathematical structure — the relative computational efficiency of the reduced–basis approach will be even more dramatic.

The obvious advantage of the reduced–basis approach within the design, optimization, and control environment is the very rapid response. However, the "blackbox" nature of the on–line component of the procedure has other advantages. In particular, the on–line (e.g., MATLAB) code is very simple, non–proprietary, transportable, and completely decoupled from the (often very complicated) off–line "truth" code. This is particularly important in the context of multidisciplinary design optimization, in which various models and approximations

must be integrated. The blackbox implementation also suggests new approaches to electronic handbooks — parameter–space exploration through actionable equations that provide rapid and certifiably accurate solutions to complex problems.

We close this section with a more applied example. We now fix all parameters except L and t, so that $P_{\text{eff}} = 2$; (L, t) are permitted to vary within $\mathcal{D}_{\text{eff}} = [2.0, 3.0] \times [0.1, 0.5]$. We choose for our two outputs the volume of the fin (easily calculable of course), $\mathcal{V}$, and the root average temperature (as defined above), s. As our "design exercise" we now construct the achievable set — all those $(\mathcal{V}, s)$ pairs associated with some (L, t) in $\mathcal{D}$; the result, based on many evaluations of $(\mathcal{V}, s_+^{\tilde{N}})$ for different values of $(L, t) \in \mathcal{D}_{\text{eff}}$, is shown in Figure 2. We present the results in terms of $s_+^{\tilde{N}}$ rather than $s^{\tilde{N}}$ to ensure that the actual temperature s_T will always be lower than our predictions (that is, conservative within the context of the design problem); and we choose $\tilde{N}$ (see Section 4) such that $s_+^{\tilde{N}}$ is always within 0.1% of s_T to ensure that the design process is not misled by inaccurate predictions. Note that, given the obvious preferences of lower volume and lower temperature, the designer will be most interested in the lower left boundary of the achievable set — the Pareto efficient frontier; although this boundary can of course be found without constructing the entire achievable set, many evaluations of the outputs will still be required.

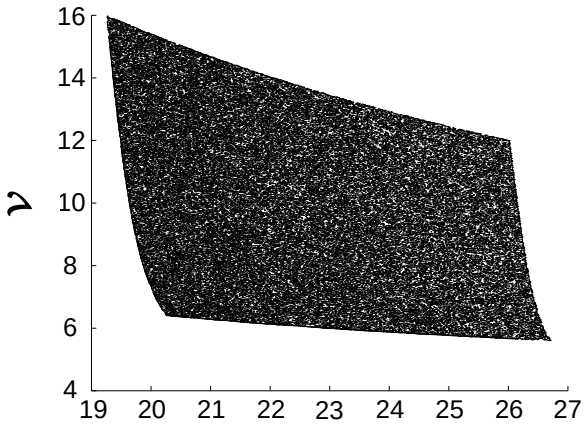

Figure 2

6. Generalizations and Issues

Many (though not all) of the assumptions that we have introduced are assumptions of convenience, not necessity, intended to simplify the exposition. First, the output functional ℓ^O need not be same as the inhomogeneity ℓ; with the introduction of an adjoint (or dual) problem [2], all of our results above extend to the more general case. Second, the function $g(\mu)$ need not be known *a priori*: $g(\mu)$ is related to an eigenvalue problem which can itself be readily approximated by a reduced–basis space constructed as the span of appropriate eigenfunctions (in theory we can now only prove asymptotic bounding properties as $N \to \infty$, however in practice the reduced–basis eigenvalue approximation converges very rapidly, and there is thus little loss of certainty). Third, these same notions extend, with some modification, to noncoercive problems, where $g(\mu)$ is now in fact the inf–sup stability parameter [3, 4]. Finally, nonsymmetric operators are readily treated, as are certain classes of nonlinearity in the state variables (e.g., eigenvalue problems [1] and Burgers equation).

Perhaps the most limiting assumption is (2), affine dependence on the parameter. In some cases (2) may indeed apply, but Q may be rather large. In such cases we can perhaps reduce the $O(Q^2)$ complexity and storage of the off–line and on–line stages to $O(Q)$ by introducing a reduced–basis approximation of the error equation (12) for a suitably chosen "staggered" sample set S_{err}^M and associated reduced–basis space constructed as the span of appropriate error functions. These ideas may also extend to the case in which the parameter dependence can not be expressed (or accurately approximated) as in (2); however we would now need to at least partially abandon the blackbox nature of the on–line stage of computation, allowing evaluation (though not inversion) of the truth–approximation operator, as well as storage of some reduced–basis vectors of size $\mathcal{N}$. These methods are currently under development; the ideas of this final paragraph are at present speculative.

REFERENCES

[1] L. Machiels, Y. Maday, I.B. Oliveira, A.T. Patera, and D.V. Rovas. Output bounds for reduced-basis approximations of symmetric positive definite

eigenvalue problems. *C. R. Acad. Sci. Paris, Sèrie I*, to appear.

[2] Y. Maday, L. Machiels, A.T. Patera, and D.V. Rovas. Blackbox reduced-basis output bound methods for shape optimization. In *Proceedings* 12^{th} *International Domain Decomposition Conference*, Chiba Japan, 2000, to appear.

[3] D.V. Rovas. An overview of blackbox reduced-basis output bound methods for elliptic partial differential equations. In *Proceedings* 16^{th} *IMACS World Congress 2000*, Lausanne Switzerland, 2000, to appear.

[4] Y. Maday, A.T. Patera, and D.V. Rovas. A blackbox reduced-basis output bound method for noncoercive linear problems. *MIT-FML Report 00-2-1, 2000; also in the College de France Series, to appear.*

[5] A.K. Noor and J.M. Peters. Reduced basis technique for nonlinear analysis of structures. *AIAA Journal*, 18(4):455-462, 1980.

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

[7] T.A. Porsching. Estimation of the error in the reduced basis method solution of nonlinear equations. *Mathematics of Computation*, 45(172):487-496, 1985.

[8] Y. Maday, A.T. Patera, and J. Peraire. A general formulation for a posteriori bounds for output functionals of partial differential equations; application to the eigenvalue problem. *C. R. Acad. Sci. Paris, Sèrie I*, 328:823-829, 1999.

[9] A.T. Patera and E.M. Rønquist. A general output bound result: application to discretization and iteration error estimation and control. *Math. Models Methods Appl. Sci.*, to appear.