12th International Conference on Domain Decomposition Methods
Editors: Tony Chan, Takashi Kako, Hideo Kawarada, Olivier Pironneau,

45. A blackbox reduced-basis output bound method for shape optimization

L. Machiels[1], Y. Maday[2], A. T. Patera[3], D. V. Rovas[4]

Introduction

We present a two-stage off-line/on-line blackbox reduced-basis output bound method for the prediction of outputs of coercive partial differential equations with affine parameter dependence. 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 both efficient and certain: thanks to rigorous *a posteriori* error bounds, we may 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 and optimization, in which repeated and rapid evaluation of the output is required.

Reduced-basis methods [ASB78, Nag79, NP80] — projection onto low-order approximation spaces comprising solutions of the problem of interest at selected points in the parameter/design space — are efficient techniques for the prediction of linear functional outputs. These methods enjoy an optimality property which ensures rapid convergence even in high-dimensional parameter spaces; good accuracy is obtained even for very few modes (basis functions), and thus the computational cost is typically very small.

It is often the case that the parameter enters affinely in the differential operator. This allows us to separate the computational steps into two stages: (i) the *off-line* stage, in which the reduced-basis space is constructed; and (ii) the *on-line*/real time stage, in which for each new parameter value the reduced-basis approximation for the output of interest is calculated. The on-line stage is "blackbox" in the sense that there is no longer any reference to the original problem formulation: the computational complexity of this stage scales only with the dimension of the reduced-basis space and the parametric complexity of the partial differential operator.

Although *a priori* theory [FR83, Por85] suggests the optimality of the reduced-basis space approximation, for a particular choice of the reduced-basis space the error in the output of interest is typically 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 basis functions are retained; the former results in computational inefficiency, the latter in uncertainty and unacceptably inaccurate

[1]Department of Mechanical Engineering, Massachusetts Institute of Technology, Room 3-264, 77 Massachusetts Avenue, Cambridge, MA 01239, USA
[2]ASCI, UPR 9029, bâtiment 506, Université Paris-Sud, 91405 Orsay cedex, France; et Analyse numérique, Université Paris VI, 4, place Jussieu, 75252 Paris cedex 05, France
[3]Department of Mechanical Enginering, Massachusetts Institute of Technology, Room 3-264, 77 Massachusetts Avenue, Cambridge, MA 01239, USA
[4]Department of Mechanical Engineering, Massachusetts Institute of Technology, Room 3-243, 77 Massachusetts Avenue, Cambridge, MA 01239, USA

predictions. In this paper we develop blackbox *a posteriori* methods that address these shortcomings. We consider here equilibrium solutions of coercive problems within the context of shape optimization; see also [MPR00] for treatment of noncoercive equilibrium problems and [MMO+00] for symmetric eigenvalue problems.

Numerical Method

Preliminaries

Let Y be a Hilbert space with an associated inner product $(\cdot,\cdot)_Y$ and an induced norm $\|\cdot\|_Y$. We define our parameter space to be $\mathcal{D} \subset \mathbb{R}$; a point in that space is denoted μ. Our problem is then to find $u \in Y$ such that

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

and subsequently the output of interest $s(u) = \ell^0(u)$; $\ell(\cdot)$ and $\ell^0(\cdot)$ are both in Y', the dual space of Y. The bilinear form a is assumed to be continuous; symmetric, $a(w, v; \mu) = a(v, w; \mu), \ \forall w, v \in Y$; and coercive, $a(v, v; \mu) \geq c\|v\|_Y^2 > 0, \ \forall v \in Y, \forall \mu \in \mathcal{D}$, where c is a strictly positive real constant. Associated with the above primal problem we define the dual problem for $\psi \in Y: \ a(v, \psi; \mu) = -\ell^0(v), \ \forall v \in Y$. The need for this problem will become clear in the error estimation discussion.

We next introduce a symmetric positive-definite form $\hat{a}(w, v)$, and define $\lambda_a^1(\mu)$ to be the minimum eigenvalue of $a(\varphi, v; \mu) = \lambda(\mu)\hat{a}(\varphi, v), \ \forall v \in Y$. A lower bound for this eigenvalue is required by the output bound procedure: we assume that a $g(\mu)$ is known such that

$$a(v, v; \mu) \geq g(\mu)\hat{a}(v, v) > 0, \ \forall v \in Y \text{ and } \forall \mu \in \mathcal{D}. \tag{2}$$

It is also possible to include approximation of $\lambda_a^1(\mu)$ as part of the reduced basis approximation [MPR00].

Finally, for the blackbox method, we shall assume that, for some finite integer Q, there exists a decomposition of $a(w, v; \mu)$ of the form

$$a(w, v; \mu) = \sum_{q=1}^{Q} \sigma^q(\mu) a^q(w, v), \forall w, v \in Y \text{and } \forall \mu \in \mathcal{D}, \tag{3}$$

where we make no assumptions on the a^q other than continuity and bilinearity.

Reduced-Basis Approximation

We choose $N/2$ points in our parameter space $\mathcal{D}$, and form the sample set $S_N = \{\mu_1, \ldots, \mu_{N/2}\}$. The reduced-basis spaces associated with the primal and dual problems are then given by $W_N^{pr} = \text{span}\{u(\mu_1), \ldots, u(\mu_{N/2})\}$ and $W_N^{du} = \text{span}\{\psi(\mu_1), \ldots, \psi(\mu_{N/2})\}$ respectively; we can now form

$$W_N = \text{span}\{u(\mu_1), \psi(\mu_1), \ldots, u(\mu_{N/2}), \psi(\mu_{N/2})\} \equiv \text{span}\{\zeta_1, \ldots, \zeta_N\}. \tag{4}$$

The space W_N defined this way has good approximation properties both for the primal and the dual problems.

For each new desired $\mu \in \mathcal{D}$, we now apply a standard Galerkin procedure over W_N to obtain $u_N(\mu)$ and $\psi_N(\mu)$ according to $a(u_N(\mu), v; \mu) = \ell(v),\ \forall v \in W_N$, and $a(v, \psi_N(\mu); \mu) = -\ell^0(v),\ \forall v \in W_N$. The output can then be calculated as $s_N(\mu) = \ell^0(u_N(\mu))$.

Bounds Evaluation

We start by defining the residuals associated with the primal and dual reduced-basis approximations, $R^{pr}(v; \mu) = \ell(v) - a(u_N(\mu), v; \mu),\ \forall v \in Y$, and $R^{du}(v; \mu) = -\ell^0(v) - a(v, \psi_N(\mu); \mu),\ \forall v \in Y$, respectively. The Riesz representations $\hat{e}^{pr}(\mu)$ and $\hat{e}^{du}(\mu)$ of the primal and dual residuals can then be defined as $\hat{a}(\hat{e}^{pr}(\mu), v) = R^{pr}(v; \mu),\ \forall v \in Y,\ \hat{a}(\hat{e}^{du}(\mu), v) = R^{du}(v; \mu), \forall v \in Y$.

We then define, as in [MMO+00, MPR00],

$$\begin{aligned} \bar{s}_N(\mu) &= s_N(\mu) - \frac{1}{2g(\mu)} \hat{a}(\hat{e}^{pr}(\mu), \hat{e}^{du}(\mu)), \\ \Delta_N(\mu) &= \frac{1}{2g(\mu)} \hat{a}^{1/2}(\hat{e}^{pr}(\mu), \hat{e}^{pr}(\mu))\ \hat{a}^{1/2}(\hat{e}^{du}(\mu), \hat{e}^{du}(\mu)), \end{aligned} \tag{5}$$

and compute lower and upper estimators $s_N^{\pm} = \bar{s}_N \pm \Delta_N$.

It can be shown [MMO+00, MPR00] that s_N^+ (respectively s_N^-) will be an upper (respectively lower) bound for s provided that $g(\mu)$ is a lower bound for the eigenvalue $\lambda_{\hat{a}}^1(\mu)$ (or equivalently satisfies (2)). Note that in the general case, where an $\hat{a}$ and $g(\mu)$ which satisfy (2) may not be readily available, the reduced-basis space must be augmented with eigenmodes corresponding to the minimum eigenvalue of the problem $a(\varphi, v; \mu) = \lambda(\mu)\hat{a}(\varphi, v),\ \forall v \in Y$ [MPR00].

Also of interest is the quality of the bounds — how well they approximate the actual error. We measure the quality of the bounds by the effectivity $\eta_N(\mu)$, defined as the ratio of the bound gap Δ_N to $|s - \overline{s}_N|$. From the bound result we know that $\eta_N(\mu) \geq 1$. We can further prove [MPR00] that $\eta_N(\mu)$ is bounded independent of N; in practice, $\eta_N(\mu)$ is typically $O(1)$, as desired.

Blackbox Method

The parametric dependence assumed in (3) permits us to decouple the computation into two stages: the *off-line* stage, in which (i) the reduced basis is constructed and, (ii) the necessary error-estimation preprocessing is performed; and the *on-line* stage, in which for each new desired value of μ, μ_d, we compute $s_N(\mu_d)$ and the associated bounds. The essential "enabler" is the absence of μ dependence in $\hat{a}$, which allows us to precompute (and later assemble) all the "pieces" of $\hat{e}^{pr}(\mu_d)$, and $\hat{e}^{du}(\mu_d)$ by linear superposition. The details of the blackbox technique follow. For convenience we define $\mathcal{N}$ as the set $\{1, \ldots, N\}$, and $\mathcal{Q}$ as the set $\{1, \ldots, Q\}$.

***Off-line* Stage**

1. Calculate $u(\mu_i)$ and $\psi(\mu_i), i = 1, \ldots, N/2$, to form W_N as in (4).
2. Compute $\underline{A}^q \in \mathbb{R}^{N\times N}$ as $A^q_{i,j} = a^q(\zeta_i, \zeta_j), \forall i, j \in \mathcal{N}^2$ and $\forall q \in \mathcal{Q}$.
3. Solve for $\hat{z}^{0,pr} \in Y$ and $\hat{z}^{0,du} \in Y$ from $\hat{a}(\hat{z}^{0,pr}, v) = \ell(v),\ \forall v \in Y$, and $\hat{a}(\hat{z}^{0,du}, v) = -\ell^0(v),\ \forall v \in Y$, respectively. Also, compute $\hat{z}^q_j \in Y$ from $\hat{a}(\hat{z}^q_j, v) = -a^q(\zeta_j, v),\ \forall v \in Y,\ \forall j \in \mathcal{N}$ and $\forall q \in \mathcal{Q}$.
4. Calculate and store $c^{pr}_0 = \hat{a}(\hat{z}^{0,pr}, \hat{z}^{0,pr})$; $c^{du}_0 = \hat{a}(\hat{z}^{0,du}, \hat{z}^{0,du})$; $c^{pr,du}_0 = \hat{a}(\hat{z}^{0,pr}, \hat{z}^{0,du})$; $F^{pr}_{N,j} = \ell(\zeta_j)$ and $F^{du}_{N,j} = \ell^0(\zeta_j),\ \forall j \in \mathcal{N}$; $\Lambda^{q,pr}_j = \hat{a}(\hat{z}^{0,pr}, \hat{z}^q_j)$ and $\Lambda^{q,du}_j = \hat{a}(\hat{z}^{0,du}, \hat{z}^q_j),\ \forall j \in \mathcal{N}$ and $\forall q \in \mathcal{Q}$; $\Gamma^{pq}_{ij} = \hat{a}(\hat{z}^p_i, \hat{z}^q_j),\ \forall i, j \in \mathcal{N}^2$ and $\forall p, q \in \mathcal{Q}^2$.

This stage requires $(NQ + N + 2)$ Y-linear system solves; $(N^2Q^2 + 2NQ + 3)$ $\hat{a}$-inner products; and $2N$ evaluations of linear functionals.

***On-line* Stage**

For each new desired design point μ_d we then compute the reduced-basis prediction and error bound based on the quantities computed in the off-line stage.

1. Form $\underline{A}_N = \sum_{q=1}^{Q} \sigma^q(\mu_d)\underline{A}^q$ and solve for $\underline{u}_N \equiv \underline{u}_N(\mu_d) \in \mathbb{R}^N$ and $\underline{\psi}_N \equiv \underline{\psi}_N(\mu_d) \in \mathbb{R}^N$ from $\underline{A}_N\ \underline{u}_N = \underline{F}^{pr}_N$ and $\underline{A}_N\ \underline{\psi}_N = -\underline{F}^{du}_N$, respectively.
2. Evaluate the bound average and bound gap as

$$\bar{s}_N = (\underline{F}^{du}_N)^T \underline{u}_N - \frac{1}{2g(\mu_d)}\Big(\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{p=1}^{Q}\sum_{q=1}^{Q} u_{N,i}\psi_{N,j}\sigma^p(\mu_d)\sigma^q(\mu_d)\Gamma^{pq}_{ij} + \sum_{j=1}^{N}\sum_{q=1}^{Q}\psi_{N,j}\sigma^q(\mu_d)\Lambda^{q,pr}_j + \sum_{j=1}^{N}\sum_{q=1}^{Q} u_{N,j}\sigma^q(\mu_d)\Lambda^{q,du}_j + c^{pr,du}_0\Big),$$

and

$$\Delta_N(\mu_d) = \frac{1}{2\ g(\mu_d)} \times \Big(\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{p=1}^{Q}\sum_{q=1}^{Q} u_{N,i}u_{N,j}\sigma^p(\mu_d)\sigma^q(\mu_d)\Gamma^{pq}_{ij} + 2\sum_{j=1}^{N}\sum_{q=1}^{Q} u_{N,j}\sigma^q(\mu_d)\Lambda^{q,pr}_j + c^{pr}_0\Big)^{\frac{1}{2}} \times \Big(\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{p=1}^{Q}\sum_{q=1}^{Q}\psi_{N,i}\psi_{N,j}\sigma^p(\mu_d)\sigma^q(\mu_d)\Gamma^{pq}_{ij} + 2\sum_{j=1}^{N}\sum_{q=1}^{Q}\psi_{N,j}\sigma^q(\mu_d)\Lambda^{q,du}_j + c^{du}_0\Big)^{\frac{1}{2}}.$$

respectively.

For each μ_d, $O(N^2Q^2 + N^3)$ operations are required to obtain the reduced-basis solution and the bounds. Since $\dim(W_N) \ll \dim(Y)$, the cost to compute $s_N(\mu_d)$, $\overline{s}_N(\mu_d)$, and $\Delta_N(\mu_d)$ in the on-line stage will typically be much less than the cost to directly evaluate $u(\mu_d)$ and $s(\mu_d) = \ell^0(u(\mu_d))$ from (1).

Results

Instantiation: Fin Problem

To illustrate our method we consider the problem of designing the thermal fin of Figure 1 to cool (say) an electronic component at the fin base, Γ_1. The ith "radiator" of the fin has thermal conductivity k_i (normalized relative to the conductivity of the central post); and the fluid surrounding the fin is characterized by a heat convection coefficient expressed in nondimensional form by a Biot number, Bi. The fin geometry is described by the radiator length β and thickness α, both nondimensionalized with respect to the width of the fin base. We thus obtain $P = 7$, with a typical point in $\mathcal{D} \in \mathbb{R}^7$ given by $\mu = \{k_1, k_2, k_3, k_4, \text{Bi}, \alpha, \beta\}$. For the output of interest we choose the mean temperature of the base, $s(u) = \ell^0(u) = \int_{\Gamma_1} u$, which is directly related to the cooling efficiency of the fin.

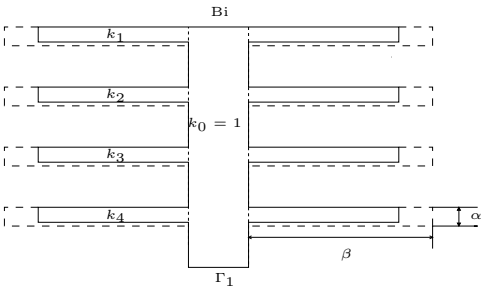

Figure 1

On the original domain the bilinear and linear forms are given by $\int_{\Omega_0} \nabla u \cdot \nabla v + \sum_{i=1}^{4} k_i \int_{\Omega_i} \nabla u \cdot \nabla v + Bi \int_{\partial\Omega \setminus \Gamma_1} uv$, and $\ell(v) = \int_{\Gamma_1} v$; here Ω_0 is the fin central post domain, and Ω_i is the ith radiator domain. (Note $\ell(v) = \ell^0(v)$, and thus the primal and dual problems coincide; this particular case is denoted compliance, and leads to considerable simplification of the numerical procedure.) We then map the domain Ω to a reference fin geometry $\hat{\Omega}$, shown by solid lines in Figure 1. The problem now takes the desired form (1) with $Y = H^1(\hat{\Omega})$ — more exactly, Y is a very fine (and hence very high-dimensional) finite element approximation of $H^1(\hat{\Omega})$ defined over a suitable triangulation of $\hat{\Omega}$. We can readily verify that the resulting form a is symmetric and positive-definite.

Taking advantage of the natural domain decomposition afforded by our mapping, it is then not difficult to cast the problem such that (3) is satisfied with $Q = 16$; the σ^q induced by the variable geometry appear as domain-dependent effective orthotropic conductivities and Bi numbers. Choosing $\hat{a}(u, v) = \sum_{q=1}^{Q} a^q(u, v) = \int_{\hat{\Omega}} \nabla u \cdot \nabla v + \int_{\partial\hat{\Omega} \setminus \Gamma_1} uv$, $g(\mu) = \min_{q \in \{1,\dots,Q\}} \sigma^q(\mu)$ (the σ^q are all bounded from below by a positive constant), we are able to verify (2). Thus all our requirements are honored, and the bound method can be applied.

Accuracy and Effectivity

We first investigate how the dimension of the reduced-basis space affects the accuracy of the bounds. We choose for the design space $\mathcal{D} = [0.1, 10]^4 \times [0.01, 1.] \times [0.1, 0.5] \times [2.0, 3.0]$, and for μ_d the value $\{0.4, 0.6, 0.8, 1.2, 0.1, 0.3, 2.8\}$. To form the reduced space we choose randomly $N/2$ points in $\mathcal{D}$. We plot in Table 1 the bound gap and effectivity as a function of N.

N	Δ_N	η_N
10	1.5987×10^{-1}	2.9947
20	1.5691×10^{-2}	2.8607
30	2.4267×10^{-3}	2.7557
40	7.2616×10^{-4}	2.6250
50	3.0620×10^{-4}	2.6085

Table 1

As we can see from Table 1, even for small N, the accuracy is very good; furthermore, convergence with N is quite rapid. This is particularly noteworthy given the high-dimensional parameter space; even with $N = 50$ points we have less than two points (effectively) in each parameter coordinate. We also note that the effectivity remains roughly constant with increasing N: the estimators are not only bounds, but relatively sharp bounds — good predictors for when N is "large enough." The behavior we observe at this particular value of μ_d is representative of most points in (a random sample over) $\mathcal{D}$, however there can certainly be points where the effectivity is larger: more systematic study is required.

Shape Optimization

Target Temperature

We suppose we wish to find the configuration which yields a base (e.g., chip) temperature of s_* (say 1.8) to within $\epsilon = .01$ by varying only the height (α) of the radiators. To start, we choose a relatively large number of basis functions in the design space $\mathcal{D}$ defined above, and perform the off-line stage of the blackbox method. For efficiency in the *on-line* stage, we then enlist only a subset of these basis functions [Kae00] — those which are closer in the design space to the desired evaluation point — and refine when higher accuracy is required. A binary chop algorithm, summarized below, is implemented to effect the coupled approximation–optimization; we assume monotonicity for simplicity of exposition.

```
for i = 1:maxiter
  Choose ᾱ := (α_l + α_r)/2
  blackbox for ᾱ ⇒ s_N^+, s_N^-
  d_1 := max(|s_* − s_N^+|, |s_* − s_N^-|)
  d_2 := min(|s_* − s_N^+|, |s_* − s_N^-|)
  if(d_2 > ε)
    if(s_N^+ > s_* and s_N^- > s_*) α_l := ᾱ
    if(s_N^+ < s_* and s_N^- < s_*) α_r := ᾱ
    else N := N + N^+
  if(d_1 < ε) stop
  else
    N := N + N^+
next
```

In the particular test case shown in Table 2, we begin with $N = 10$ points and set $N^+ = 10$ as well; we initialize $\alpha_l = 0.1$ and $\alpha_r = 0.5$. During the optimization process, refinement is effected twice, such that a total of $N = 30$ basis functions are invoked (considerably less than the 50 available). The savings are significant, yet we are still ensured, thanks to the bounds, that our design requirement is met to the desired tolerance of $\varepsilon = .01$. One can also apply a dynamic adaptation strategy in which only a minimal number of basis functions are generated (initially) in the off-line stage: if these prove inadequate, we return to the off-line stage for additional basis functions and also revision of the necessary matrices and inner products.

i	$\bar{\alpha}$	s_N^+	s_N^-	α_l	α_r
1	0.3	1.683	1.753	0.1	0.5
2	0.2	1.716	2.056	0.1	0.3
3	0.2	1.766	1.807	0.1	0.3
4	0.2	1.771	1.778	0.1	0.3
5	0.15	1.817	1.840	0.1	0.2
6	0.175	1.792	1.806	.15	0.2

Table 2

If we choose a tighter tolerance ε, or if we wish to investigate many different set points s_*, or if we perform the optimization permitting all 7 design parameters to vary, we would of course greatly increase the number of output predictions required — and hence greatly increase the efficiency of the reduced-basis blackbox technique relative to conventional approaches.

Achievable Set

In multicriterion optimization we consider various (competing) outputs of interest, say volume, $\mathcal{V}$, and root temperature, s. Changing the dimensions of the fin by selecting different α and β will (say) decrease the volume of the fin, and hence material requirements - but also (typically) increase the fin base temperature. It is thus of interest to determine all possible operating points, that is, to generate the map of the "achievable set." In general this will be prohibitively expensive unless one has recourse to a very low-dimensional representation such as the reduced-basis approximation.

We consider this problem for constant conductivities $k_i = 1.$, $i = 0, \ldots, 4$, and Biot number Bi$= 0.001$. We then select 100,000 points in the two dimensional design space $[\alpha, \beta] = [0.1, 0.5] \times [2.0, 3.0]$ and evaluate our bounds for s with an error tolerance of 0.1%. Since in this design we wish to be sure that the actual temperature will be less than our prediction, we choose to construct our map based on s_N^+. We are thus insured that at each design point the actual temperature will be lower than that on our curve.

Each evaluation produces a point on the s–$\mathcal{V}$ plane, thus generating the achievable set. Obvious optimality conditions require that we remain on the left or lower boundaries of the achievable set, known as the efficient frontier or trade-off curve in Pareto analysis. As we can see from Figure 2, we can decrease the volume with no real increase in temperature up to the point were the left and lower boundaries cross; after that, the small further possible volume reduction results in a steep rise in base temperature.

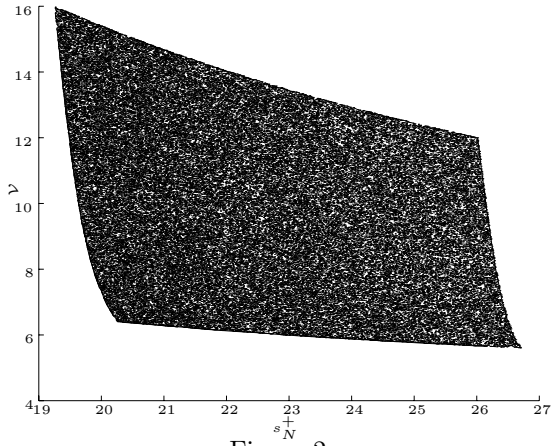

Figure 2

Acknowledgements This work was supported by the Singapore–MIT Alliance, by AFOSR Grant F49620-97-1-0052, and by NASA Grant NAG1-1978.

References

[ASB78]B. O. Almroth, P. Stern, and F. A. Brogan. Automatic choice of global shape functions in structural analysis. *AIAA Journal*, 16:525–528, May 1978.

[FR83]J. P. Fink and W. C. Rheinboldt. On the error behaviour of the reduced basis technique for nonlinear finite element approximations. *Z. Angew. Math. Mech.*, 63(1):21–28, 1983.

[Kae00]R. Kaenel. Reduced basis methods and output bounds for partial differential equations. Master's thesis, MIT, EPFL, 2000.

[MMO+00]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, Série I*, 2000.

[MPR00]Y. Maday, A. T. Patera, and D. V. Rovas. A blackbox reduced-basis output bound method for noncoercive linear problems. *College de France Series; also MIT-FML Report 00-2-1*, 2000.

[Nag79]D. A. Nagy. Modal representation of geometrically nonlinear behaviour by the finite element method. *Computers and Structures*, 10:683–688, 1979.

[NP80]A. K. Noor and J. M. Peters. Reduced basis technique for nonlinear analysis of structures. *AIAA Journal*, 18(4):455–462, April 1980.

[Por85]T. A. Porsching. Estimation of the error in the reduced basis method solution of nonlinear equations. *Mathematics of Computation*, 45(172):487–496, October 1985.