Goal-oriented error estimation for the reduced basis method, with application to sensitivity analysis

Alexandre Janon* Maëlle Nodet† Clémentine Prieur†

July 10, 2014

Abstract

The reduced basis method is a powerful model reduction technique designed to speed up the computation of multiple numerical solutions of parametrized partial differential equations. We consider a quantity of interest, which is a linear functional of the PDE solution. A new probabilistic error bound for the reduced model is proposed. It is efficiently and explicitly computable, and we show on different examples that this error bound is sharper than existing ones. We include application of our work to sensitivity analysis studies.

Keywords: reduced basis method, surrogate model, reduced order modelling, response surface method, scientific computation, sensitivity analysis, Sobol index computation, Monte-Carlo method
AMS MSC: 65M15

Introduction

A large number of mathematical models are based on partial differential equations (PDEs). These models require input data (e.g., the physical features of the considered system, the geometry of the domain, the external forces...) which enter in the PDE as *parameters*. In many applications (for instance, design optimization, data assimilation, or uncertainty quantification), one has to numerically compute the solution of a parametrized partial differential equation for a large number of values of the parameters. In such a case, it is generally interesting, in terms of computation time, to perform all possible parameter-independent computations in an *offline* phase, which is done only once, and to call an *online* phase for each required value of the parameter, during which the

*Laboratoire de Mathématiques d'Orsay, Université Paris-Sud (France)
†Laboratoire Jean Kuntzmann, Université Joseph Fourier, INRIA/MOISE

information gathered in the offline phase can be used to speed up the computation of an approximate solution of the PDE, and, hence, to reduce the marginal (ie., per parameter) computation cost.
The reduced basis method [12] is a way of specifying such offline and online phases, which has been successfully applied to various well-known PDEs [4, 7, 9, 21]. One should note that, in the reduced basis (RB) method, the online phase does not compute a solution which is strictly identical to the numerical PDE solution, but an approximation of it, obtained by projecting the original discretized equations onto a well-chosen basis. In the application cases given above, however, one is not interested in the solution by itself, but rather in a *quantity of interest*, or model *output*, which is a functional of this solution. Taking this functional into account when performing the model reduction leads to a so-called *goal-oriented* method. For instance, goal-oriented basis choice procedures have been tried with success in the context of dynamical systems in [6, 23], where the basis is chosen so as to contain the modes that are relevant to accurately represent the output of interest, and in a general context in [2], where the basis is chosen so as to minimize the overall output error. All those papers showed that using an adapted basis could lead to a great improvement of reduction error.
This paper is about goal-oriented error estimation, that is, the description of a rigorous and computable *error bound* between the model output and the reduced one. Two different reduced model outputs can be considered: the first (which we call *uncorrected* reduced output) is simply the output functional evaluated at the reduced output. The second (called *corrected* reduced output), described in [10–12], is the same, up to a correction term obtained from the solution of an auxiliary (dual) problem. The rate of convergence of the corrected output is better than the uncorrected one but the computation of the correction involves the application of the RB method to the dual problem, and this has generally the drawback of doubling offline *and online* computational times. Regarding output error estimation, an error bound for the difference between the corrected reduced output and the original output is provided in the papers cited above. In this paper, we propose two new goal-oriented error bounds: one for the uncorrected reduced output, and one for the corrected reduced output. We also show, in numerical examples, that our bound is more precise than the existing bound.
This paper is organized as follows: in the first part, we describe our output error bounds and explain how to compute them; in the second part, we see how to apply our error bound to sensitivity analysis studies; finally, the third and fourth parts present numerical applications.

1 Methodology

1.1 Preliminaries

Reference problem We begin by setting up the context of the reduced basis method for affine-parametrized linear partial differential equations presented in [12]. Our reference problem is the following: given a parameter tuple $\mu \in \mathcal{P} \subset \mathbb{R}^p$, and the vector space $X = \mathbb{R}^{\mathcal{N}}$ (for $\mathcal{N} \in \mathbb{N}$), find $u(\mu)$, the solution of:

$$A(\mu)u(\mu) = f(\mu), \tag{1}$$

where $A(\mu)$ is an invertible square matrix of dimension $\mathcal{N}$, and $f(\mu) \in X$, then compute the *output*:

$$s(\mu) = S(u(\mu)) \tag{2}$$

where $S : X \to \mathbb{R}$ is a linear form on X.

Choice of the inner product: We suppose that $X = \mathbb{R}^{\mathcal{N}}$ (with the standard basis), is endowed with the standard Euclidean inner product: $\langle u, v\rangle = u^t v$, with associated norm $\|u\| = \sqrt{\langle u, u\rangle}$

Affine decomposition hypothesis: We suppose that $A(\mu)$ and $f(\mu)$ admit the following so-called affine decomposition [12]:

$$\forall \mu \in \mathcal{P}, \ \ A(\mu) = \sum_{q=1}^{Q} \Theta_q(\mu) A_q, \ \ f(\mu) = \sum_{q'=1}^{Q'} \gamma_{q'}(\mu) f_{q'} \tag{3}$$

where $Q, Q' \in \mathbb{N}^*$, $\Theta_q : \mathcal{P} \to \mathbb{R}$ and $\gamma_{q'} : \mathcal{P} \to \mathbb{R}$ (for $q = 1, \ldots, Q$, $q' = 1, \ldots, Q'$) are given functions, A_q are square matrices of dimension $\dim X$ and $f_{q'} \in X$. This hypothesis is required by the reduced basis method.

Reduced basis method The dimension of the finite element subspace $\dim X$ is generally fairly large, so that the numerical computation of $u(\mu)$ from the inversion of $A(\mu)$ is expensive. The reduced basis aims at speeding up "many queries", that is, the computation of $u(\mu)$ for all parameters $\mu \in \mathcal{P}_0$ where $\mathcal{P}_0$ is a finite but "large" subset of the parameter set $\mathcal{P}$.

Reduced problem: We consider a subspace $\widetilde{X}$ of X, and a matrix Z whose columns are the components of a basis of $\widetilde{X}$ in a basis of X. This basis of $\widetilde{X}$ is called the *reduced basis* in the sequel. We denote by $\widetilde{u}(\mu)$ the components, in the reduced basis, of the solution of the projection of (1) onto $\widetilde{X}$, that is, the solution of:

$$Z^t A Z \widetilde{u}(\mu) = Z^t f(\mu) \tag{4}$$

(where, for any matrix M, M^t is the transpose of M).

Choice of the reduced subspace: There are different techniques for choosing the reduced basis (the Z matrix). This paper does not focus on this topic, but we cite the POD method (Proper orthogonal decomposition) [18], and the Greedy method [12].

Offline-online decomposition: The many-query computation can then be split into two parts: the first part (usually called the "offline phase"), which is done only once, begins by finding a reduced subspace, then the Q parameter-independent matrices:

$$\widetilde{A}_q = Z^t A_q Z, \ \ q = 1, \ldots, Q$$

and the Q' vectors:

$$\widetilde{f}_{q'} = Z^t f_{q'}, \quad q' = 1, \ldots, Q'$$

are computed and stored. In the second part (the "online phase"), we compute, for each value of the parameter μ:

$$\widetilde{A}(\mu) = \sum_{q=1}^{Q} \Theta_q(\mu)\widetilde{A}_q, \quad \widetilde{f}(\mu) = \sum_{q'=1}^{Q'} \gamma_q(\mu)\widetilde{f}_{q'} \tag{5}$$

and solve for $\widetilde{u}(\mu)$ satisfying:

$$\widetilde{A}(\mu)\widetilde{u}(\mu) = \widetilde{f}(\mu). \tag{6}$$

The key point is that the operations in (5) and (6) are performed on vectors and matrices of size $\dim \widetilde{X}$, and that the complexity of these operations is totally independent from the dimension of the underlying "truth" subspace X. In many cases, the smoothness of the map $\mu \mapsto u(\mu)$ allows to find (in a constructive way, ie., compute) $\widetilde{X}$ so that $\dim \widetilde{X} \ll \dim X$ while keeping $\|u(\mu) - Z\widetilde{u}(\mu)\|$ small, hence enabling significant computational savings.

Output approximation: The output $s(\mu)$ can also be approximated from $\widetilde{u}(\mu)$ using an efficient offline-online procedure: let $l \in X$ be so that:

$$s(u(\mu)) = \langle l, u \rangle \ \ \forall u \in X;$$

in the offline phase we compute and store:

$$\widetilde{l} = Z^t l$$

and in the online phase we take:

$$\widetilde{s}(\mu) = \langle \widetilde{l}, \widetilde{u}(\mu) \rangle$$

as an approximation for $s(\mu)$.

Reduced-basis error bounds

Bound on u: Under additional coercivity hypothesis on A, the reduced basis method [12] also provides an efficient offline-online procedure for computing $\epsilon^u(\mu)$ so that the approximation can be *certified*. This bound is based on the *dual norm of the residual*:

$$\rho(\mu) = \|r(\mu)\|_{RB} = \|A(\mu)Z\widetilde{u}(\mu) - f(\mu)\|_{RB}$$

where $\|\cdot\|_{RB}$ is a suitably chosen norm on X (not necessarily $\|\cdot\|$), and a *stability constant* bound, which can be written as:

$$0 < \alpha(\mu) \leq \inf_{v \in X, \|v\|_{RB}=1} \left| v^t A(\mu) v \right| \tag{7}$$

when A is symmetric. The inequality sign is due to the fact that the exact infimum can be costly to evaluate in the online stage; usually a procedure such as the successive constraints method [5] is used in order to quickly find a lower bound.
The bound $\epsilon^u(\mu)$ reads:

$$\forall \mu \in \mathcal{P} \quad \|u(\mu) - Z\widetilde{u}(\mu)\|_{RB} \leq \frac{\rho(\mu)}{\alpha(\mu)} := \epsilon^u(\mu),$$

The online procedure for the computation of $\epsilon(\mu)$ is also of complexity independent of $\dim X$.

Lipschitz bound on s: This online error bound can in turn be used to provide a certification on the output:

$$\forall \mu \in \mathcal{P} \quad |s(\mu) - \widetilde{s}(\mu)| \leq \underbrace{\|l\|_{RB}\, \epsilon^u(\mu)}_{=:\epsilon^L(\mu)} \tag{8}$$

We call this bound the "Lipschitz" bound, and denote it by $\epsilon^L(\mu)$. It is well-known that this bound is very pessimistic.

The aim of Section 1.2 is to bound $|s(\mu) - \widetilde{s}(\mu)|$ by a quantity which is smaller than $\epsilon^L(\mu)$ of (8) and can be computed using an efficient offline-online procedure which does not require computation of $\epsilon^u(\mu)$, described in Section 1.3. In Section 1.4, we consider a better approximation of $s(\mu)$ (denoted by $\widetilde{s}_c(\mu)$) which also depends on the solution of the adjoint equation of (1) projected on a suitably selected dual reduced basis, and we see how the proposed bound for $|s(\mu) - \widetilde{s}(\mu)|$ can be modified in order to bound $|s(\mu) - \widetilde{s}_c(\mu)|$.

1.2 Probabilistic error bound

In this section, we give the expression of our output error bound. We recall the notation for the residual $r(\mu)$:

$$r(\mu) = A(\mu)Z\widetilde{u}(\mu) - f(\mu) \in X,$$

and the adjoint problem solution (which will naturally appear in the proof of Theorem 1.1) by $w(\mu)$:

$$w(\mu) = A(\mu)^{-t} l.$$

Let, for any orthonormal basis $\Phi = \{\phi_1, \dots, \phi_{\mathcal{N}}\}$ of X, any $N \in \mathbb{N}^*$, and $i = 1, \dots, N$,

$$D_i(\mu, \Phi) = \langle w(\mu), \phi_i \rangle.$$

We take a partition $\{\mathcal{P}_1, \dots, \mathcal{P}_K\}$ of the parameter space $\mathcal{P}$, that is:

$$\mathcal{P} = \cup_{k=1}^K \mathcal{P}_k \quad \text{and} \quad k \neq k' \Rightarrow \mathcal{P}_k \cap \mathcal{P}_{k'} = \emptyset.$$

We set, for $i = 1, \dots, N$ and $k = 1, \dots, K$:

$$\beta_{i,k}^{min}(\Phi) = \min_{\mu \in \mathcal{P}_k} D_i(\mu, \Phi), \quad \beta_{i,k}^{max}(\Phi) = \max_{\mu \in \mathcal{P}_k} D_i(\mu, \Phi),$$

and:

$$\beta_i^{up}(\mu, \Phi) = \begin{cases} \beta_{i,k(\mu)}^{max}(\Phi) \text{ if } \langle r(\mu), \phi_i \rangle > 0 \\ \beta_{i,k(\mu)}^{min}(\Phi) \text{ else,} \end{cases}$$

$$\beta_i^{low}(\mu, \Phi) = \begin{cases} \beta_{i,k(\mu)}^{min}(\Phi) \text{ if } \langle r(\mu), \phi_i \rangle > 0 \\ \beta_{i,k(\mu)}^{max}(\Phi) \text{ else,} \end{cases}$$

where $k(\mu)$ is the only k in $\{1, \dots, K\}$ so that $\mu \in \mathcal{P}_k$. We also set:

$$T_1^{low}(\mu, N, \Phi) = \sum_{i=1}^{N} \langle r(\mu), \phi_i \rangle \beta_i^{low}(\mu, \Phi), \;\; T_1^{up}(\mu, N, \Phi) = \sum_{i=1}^{N} \langle r(\mu), \phi_i \rangle \beta_i^{up}(\mu, \Phi),$$

$$T_1(\mu, N, \Phi) = \max\left(\left| T_1^{low}(\mu, N, \Phi) \right|, \left| T_1^{up}(\mu, N, \Phi) \right| \right).$$

Finally, we suppose that μ is a random variable on $\mathcal{P}$ and set:

$$T_2(N, \Phi) = \mathbf{E}_\mu \left(\left| \sum_{i=N+1}^{\mathcal{N}} \langle w(\mu), \phi_i \rangle \langle r(\mu), \phi_i \rangle \right| \right).$$

We have the following theorem:

Theorem 1.1. *For any $\alpha \in]0;1[$ and for any $N \in \mathbb{N}^*$, we have:*

$$P\left(|s(\mu) - \widetilde{s}(\mu)| > T_1(\mu, N, \Phi) + \frac{T_2(N, \Phi)}{\alpha} \right) \leq \alpha.$$

Proof: We begin by noticing that:

$$A(\mu)^{-1} r(\mu) = Z\widetilde{u}(\mu) - u(\mu)$$

so that:

$$\widetilde{s}(\mu) - s(\mu) = \langle l, Z\widetilde{u}(\mu) - u(\mu) \rangle = \langle l, A(\mu)^{-1} r(\mu) \rangle = \langle w(\mu), r(\mu) \rangle.$$

We expand the residual in the Φ basis:

$$r(\mu) = \sum_{i \geq 1} \langle r(\mu), \phi_i \rangle \phi_i.$$

Hence:

$$\widetilde{s}(\mu) - s(\mu) = \sum_{i \geq 1} \langle l, A(\mu)^{-1} \phi_i \rangle \langle r(\mu), \phi_i \rangle = \sum_{i \geq 1} \langle w(\mu), \phi_i \rangle \langle r(\mu), \phi_i \rangle. \tag{9}$$

We clearly have that for any $N \in \mathbb{N}^*$:

$$\sum_{i=1}^{N}\langle r(\mu),\phi_i\rangle\beta_i^{low}(\mu,\Phi) \le \sum_{i=1}^{N}\langle r(\mu),\phi_i\rangle\langle w(\mu),\phi_i\rangle \le \sum_{i=1}^{N}\langle r(\mu),\phi_i\rangle\beta_i^{up}(\mu,\Phi)$$

and this implies:

$$\left|\sum_{i=1}^{N}\langle r(\mu),\phi_i\rangle\langle w(\mu),\phi_i\rangle\right| \le T_1(\mu,N,\Phi). \tag{10}$$

So we have:

$$\begin{aligned}
&P\left(|s(\mu)-\widetilde{s}(\mu)| > T_1(\mu,N,\Phi) + \frac{T_2(N,\Phi)}{\alpha}\right)\\
&\le P\left(|s(\mu)-\widetilde{s}(\mu)| > \left|\sum_{i=1}^{N}\langle r(\mu),\phi_i\rangle\langle w(\mu),\phi_i\rangle\right| + \frac{T_2(N,\Phi)}{\alpha}\right) \text{ by (10)}\\
&= P\left(|s(\mu)-\widetilde{s}(\mu)| - \left|\sum_{i=1}^{N}\langle r(\mu),\phi_i\rangle\langle w(\mu),\phi_i\rangle\right| > \frac{T_2(N,\Phi)}{\alpha}\right)\\
&\le P\left(\left|\sum_{i=N+1}^{\mathcal{N}}\langle r(\mu),\phi_i\rangle\langle w(\mu),\phi_i\rangle\right| > \frac{T_2(N,\Phi)}{\alpha}\right) \text{ by (9)}\\
&\le \alpha \text{ thanks to Markov's inequality.} \qquad \square
\end{aligned}$$

Choice of Φ The error bound given in Theorem 1.1 above is valid for any orthonormal basis Φ. For efficiency reasons, we would like to choose Φ so that the parameter-independent part $T_2(N,\Phi)$ is the smallest possible, for a fixed truncation index $N \in \mathbb{N}^*$.
To our knowledge, minimizing $T_2(N,\Phi)$ over orthonormal bases of X is an optimization problem for which no efficient algorithm exists. However, we can minimize an upper bound of $T_2(N,\Phi)$.
We define an auto-adjoint, positive operator $G : X \to X$ by:

$$\forall \phi \in X,\ \ G\phi = \frac{1}{2}\mathbf{E}_\mu\left(\langle r(\mu),\phi\rangle r(\mu) + \langle w(\mu),\phi\rangle w(\mu)\right). \tag{11}$$

Let $\lambda_1 \ge \lambda_2 \ge \dots \lambda_{\mathcal{N}} \ge 0$ be the eigenvalues of G. Let, for $i \in \{1,2\dots,\mathcal{N}\}$, ϕ_i^G be an unit eigenvector of G associated with the i^{th} eigenvalue, and $\Phi^G = \{\phi_1^G,\dots,\phi_{\mathcal{N}}^G\}$.
We can state that:

Theorem 1.2.

$$T_2(N,\Phi^G) \le \sum_{i=N+1}^{\mathcal{N}} \lambda_i^2.$$

Proof. We have:

$$T_2(N,\Phi) \le \frac{1}{2}\mathbf{E}_\mu \left(\sum_{i=N+1}^{\mathcal{N}} \langle w(\mu), \phi_i \rangle^2 + \sum_{i=N+1}^{\mathcal{N}} \langle r(\mu), \phi_i \rangle^2 \right) =: T_2^{sup}(N,\Phi) = \sum_{i=N+1}^{\mathcal{N}} \langle G\phi_i, \phi_i \rangle$$

Using Theorem 1.1 of [22], we get that the minimum of $T_2^{sup}(N,\Phi)$ is attained for $\Phi = \Phi^G$, and that minimum is $\sum_{i=N+1}^{\mathcal{N}} \lambda_i^2$. □

This theorem suggests to use $\Phi = \Phi^G$, so as to control $T_2(N,\Phi)$.

1.3 Monte-Carlo approximation of the error bound

In this Subsection, we present an implementable offline/online procedure for the estimation of the upper bound for $|\widetilde{s}(\mu) - s(\mu)|$ presented in Theorem 1.1.

Estimation of ϕ_i^G We fix a truncation index $N \in \mathbb{N}^*$, and we estimate $\{\phi_i^G\}_{i=1,\dots,N}$ by using a modification of the method of snapshots used in Proper Orthogonal Decomposition [18]. This estimation is performed during the offline phase. We begin by estimating the G operator by $\widehat{G}$, then we approximate ϕ_i^G by the appropriate eigenvectors of $\widehat{G}$.

Estimation of G: We take a finite (large), subset of parameters $\Xi \subset \mathcal{P}$, randomly sampled from the distribution of the parameter, and we approximate the G operator by:

$$\widehat{G}\phi = \frac{1}{2\#\Xi} \sum_{\mu \in \Xi} \left(\langle r(\mu), \phi \rangle r(\mu) + \langle w(\mu), \phi \rangle w(\mu) \right)$$

In other words, $\widehat{G}$ is a Monte-Carlo estimator of G. We take $\{\widehat{\phi}_i^G\}_{i=1,\dots,N}$ as the unit eigenvectors associated with the N largest eigenvalues of $\widehat{G}$.

Computation of the eigenvalues of $\widehat{G}$: The operator $\widehat{G}$ admits the following matrix representation:

$$\widehat{G} = \frac{1}{2\#\Xi} \left(WW^t + RR^t \right),$$

where W (resp. R) is the matrix whose columns are the components of $w(\mu)$ (resp. $r(\mu)$) in a basis of X, for $\mu \in \Xi$. These two matrices have $\#\Xi$ columns and $\dim X$ lines, which means that the matrix above is $\dim X \times \dim X$.
In general, we take $\#\Xi \ll \dim X$, and so it is computationally advantageous to notice that if ϕ is an eigenvector of $\widehat{G}$ associated with a nonzero eigenvalue λ, then:

$$\frac{1}{\lambda}\frac{1}{2\#\Xi} \left((WW^t\phi + RR^t\phi) \right) = \phi,$$

so that $\phi \in \mathrm{Im}\ W + \mathrm{Im}\ R =: \mathcal{V}$. Hence, if V is the matrix of an orthonormal basis of $\mathcal{V}$, then there exists ψ so that $\phi = V\psi$ and we have:

$$WW^t\phi + RR^t\phi = \lambda\phi \implies \left[V^t \frac{1}{2\#\Xi}\left((WW^t + RR^t)\right)V\right]\psi = \lambda\psi.$$

As a consequence, it is sufficient to find the dominant eigenvectors $\widehat{\psi}_1^G, \dots, \widehat{\psi}_N^G$ of the matrix $\Sigma = \frac{1}{2\#\Xi}V^t(WW^t + RR^t)V$ (of size $2\#\Xi$), and to deduce $\widehat{\phi}_i^G$ from $\widehat{\psi}_i^G$ by the relation $\widehat{\phi}_i^G = V\widehat{\psi}_i^G$. Besides, by writing Σ as:

$$\Sigma = \frac{1}{2\#\Xi}\left((V^tW)(W^tV) + (V^tR)(R^tV)\right),$$

it is possible to compute and store Σ without storing nor computing any dense $\dim X \times \dim X$ matrix.

Computation of $T_1(\mu, N, \Phi)$ All the quantities intervening in $T_1(\mu, N, \Phi)$ can be straightforwardly deduced from $\beta_{i,k}^{min,max}$ and $\langle r(\mu), \widehat{\phi}_i^G\rangle$.

Computation of $\beta_{i,k}^{min}(\Phi)$ *and* $\beta_{i,k}^{max}(\Phi)$*:* For $i = 1, \dots, N$ and $k = 1, \dots, K$, the reals $\beta_{i,k}^{min}(\Phi)$ and $\beta_{i,k}^{max}(\Phi)$ can be computed during the offline phase, as they are parameter-independent. If the Θ_qs functions are smooth enough, a quasi-Newton optimization such as L-BFGS [24] can be used so as to compute these reals; they can also be approximated by a simple discrete minimization:

$$\widetilde{\beta}_{i,k}^{min}(\Phi) = \min_{\mu \in \Xi \cap \mathcal{P}_k} D_i(\mu, \Phi), \quad \widetilde{\beta}_{i,k}^{max}(\Phi) = \max_{\mu \in \Xi \cap \mathcal{P}_k} D_i(\mu, \Phi). \tag{12}$$

One should note here that one has to be careful, as these numerical optimisation procedures are not exact. Indeed, the discrete optimisation is only an approximation, and the Quasi-Newton optimisation may be handled carefully, so as to avoid being trapped in local extrema.

Computation of $\langle r(\mu), \widehat{\phi}_i^G\rangle$*:* We denote by $\{\zeta_1, \dots, \zeta_n\}$ are the column vectors of Z, which form a basis of the reduced space $\widetilde{X}$. We can, during the offline phase, compute the following parameter-independent quantities:

$$\langle f_{q'}, \widehat{\phi}_i^G\rangle, \langle A_q\zeta_j, \widehat{\phi}_i^G\rangle \ (i = 1, \dots, N,\ j = 1, \dots, n,\ q = 1, \dots, Q,\ q' = 1, \dots, Q').$$

Let a parameter $\mu \in \mathcal{P}$ be given, and $\widetilde{u}_1(\mu), \dots, \widetilde{u}_n(\mu)$ be the components of the reduced solution $\widetilde{u}(\mu)$ in the reduced basis $\{\zeta_1, \dots, \zeta_n\}$.
By using the relation:

$$\langle r(\mu), \widehat{\phi}_i^G\rangle = \sum_{q=1}^{Q}\Theta_q(\mu)\sum_{j=1}^{n}\widetilde{u}_j(\mu)\langle A_q\zeta_j, \widehat{\phi}_i^G\rangle - \sum_{q'=1}^{Q'}\gamma_{q'}(\mu)\langle f_{q'}, \phi_i^G\rangle,$$

the dot products between the residual and $\widehat{\phi}_i^G$ can be computed in the online phase, with a complexity of $O(nQ+Q')$ arithmetic operations, $O(Q)$ evaluations of Θ functions and $O(Q')$ evaluations of γ functions, which is independent of $\dim X$.

Estimation of $T_2(N,\Phi)$ We approximate $T_2(N,\Phi)$ by computing the following Monte-Carlo estimator:

$$\widehat{T}_2(N,\Phi) = \frac{1}{2\#\Xi} \sum_{\mu\in\Xi} \left| \widetilde{s}(\mu) - s(\mu) - \sum_{i=1}^{N} \langle w(\mu), \phi_i \rangle \langle r(\mu), \phi_i \rangle \right|.$$

As this quantity is μ-independent, it can be computed once and for all during the offline phase.

Final error bound By using Theorem 1.1, we get that for $\epsilon(\mu,\alpha,N,\Phi) = T_1(\mu,N,\Phi) + T_2(N,\Phi)/\alpha$, we have:

$$P\left(|s(\mu) - \widetilde{s}(\mu)| \geq \epsilon(\mu,\alpha,N,\Phi)\right) \leq \alpha,$$

so we may take, as estimated (computable) error bound with risk α,

$$\widehat{\epsilon}(\mu,\alpha,N,\Phi) = T_1(\mu,N,\Phi) + \frac{\widehat{T}_2(N,\Phi)}{\alpha}. \tag{13}$$

In the rest of the text, this computable error bound is designated as the *error bound on the non-corrected output*, by contrast to the bound described in Section 1.2.

One may note that estimating $T_2(N,\Phi)$ by $\widehat{T}_2(N,\Phi)$ causes some error on the risk of the computable error bound $\widehat{\epsilon}(\mu,\alpha,N,\Phi)$. This error is analyzed in Appendix A.

1.4 Bound on the corrected output

The reduced output $\tilde{s}(\mu)$ is a natural reduced output that approximates $s(\mu)$. It is possible to solve an auxiliary problem in order to compute an error correction that improves the order of convergence of the reduced output. As we will see, the error bound presented above can be easily modified so as to certify the corrected output.

Output correction The idea of solving an adjoint problem in order to improve the order of convergence of a scalar output has originated in [10], with the first application in the reduced basis context in [11]. We introduced the so-called adjoint problem, whose solution $w(\mu)$ satisfies:

$$A(\mu)^t u_d w(\mu) = l,$$

and the solution $\tilde{u}_d\tilde{w}(\mu)$ of the reduced adjoint problem:

$$Z_d^t A(\mu)^t Z_d \tilde{u}_d \tilde{w}(\mu) = Z_d^t l,$$

where Z_d is the selected matrix of the reduced basis for the adjoint basis.

The corrected output is:

$$\tilde{s}_c(\mu) = \tilde{s}(\mu) - \langle Z\tilde{u}_d\tilde{w}(\mu), r(\mu)\rangle$$

One shoud note that the corrected output can be computed using an efficient offline-online procedure which requires *two* reduced basis solutions, hence roughly doubling (when Z_d has the same number of columns than Z) the offline and online computation times, except in the particular case where A is symmetric and l is proportional to f.

Existing error bound on the corrected output In [12], it is shown that:

$$|s(\mu) - \tilde{s}_c(\mu)| \leq \frac{\|r(\mu)\|' \, \|r_d(\mu)\|'}{\alpha(\mu)} =: \epsilon_{cc}(\mu) \tag{14}$$

where $\alpha(\mu)$ is the stability constant bound defined at (7) (in the symmetric case) and $r_d(\mu)$ is the dual residual:

$$r_d(\mu) = A^t(\mu) Z_d \tilde{u}_d \tilde{w}(\mu) - l(\mu).$$

Hereafter, the existing bound ϵ_{cc} is called *dual-based error bound*.

Proposed error bound It is clear that the work performed in the above sections can be reused so as to provide a probabilistic bound on $|s(\mu) - \tilde{s}_c(\mu)|$, by simply replacing $w(\mu)$ by:

$$w_c(\mu) = w(\mu) - Z\tilde{u}_d\tilde{w}(\mu), \tag{15}$$

and hence giving a competitor (called *error bound on the corrected output*) to the dual-based error bound.

1.5 Summary of the different bounds

To sum up, we have presented four output computable error bounds. Two of them are bounds for the uncorrected ouput error:

- the "Lipschitz" error bound ϵ^L (8);
- the estimated error bound on the uncorrected output $\widehat{\epsilon}$, that we propose in (13);

and two of them for the corrected output error:

- the existing dual-based error bound ϵ_{cc}, defined at (14);
- the estimated error bound on the corrected output $\widehat{\epsilon}_c$, that is (13) amended with (15).

Comparison of the various presented bounds. We can now discuss and compare the different considered error bounds (numerical comparisons are gathered in sections 3 and 4).
Online costs: using ϵ^{cc} or $\widehat{\epsilon}^c$ will require a double computation time (except if we are in a "compliant" case, i.e. $l = f$), when compared with using ϵ^L or $\widehat{\epsilon}$. However, this computational supplement enables error correction of the reduced output and an improved error estimation. Except for the possible supplemental cost for the adjoint problem resolution, the online cost of $\widehat{\epsilon}^c$ or $\widehat{\epsilon}$ should be comparable to the cost of ϵ^{cc}, if not slightly better for the former, as no optimization problem has to be solved in the online phase (while ϵ^{cc}, when using the Successive Constraints Method (SCM) [5], requires a linear programming in order to compute an online $\alpha(\mu)$).
Offline costs: the SCM procedure used for ϵ^{cc} and ϵ^L requires the offline resolution of a number of large eigenproblems on the "finite element space" X. Depending on the problem at hand, this may, or may not be more expensive than the optimization problems required to compute the $\widetilde{\beta}$ constants (see (12)) and the Monte-Carlo estimation of $\widehat{T}_2$. One of the methods can also be feasible and the other not.
Probabilistic vs. deterministic: one can also notice that the $\widehat{\epsilon}^c$ and $\widehat{\epsilon}$ bounds are probabilistic in nature; they increase when the risk level decreases. In practice, as shown in the numerical experiments of Section 4, our probabilistic bound is much shaper, even while choosing a very small risk level.
Accuracy: the classical bound introduces the dual norm of the residual, which causes a loss of accuracy. In this work, we avoid this step by using the majoration in (10) and a probabilistic argument.

To conclude this discussion, one can say that there may not be a definitive winner error bound, and that the best choice highly depends on the problem at hand (dimension of the X space, numbers of terms in the affine decompositions), the computational budget, the required precision, the number of online queries to perform, and the probability of failure of the error bound that one can afford.

2 Application to sensitivity analysis

Our error estimation method is applied in sensitivity analysis, so as to quantify the error caused by the replacement of the original model output by the reduced basis output during the Monte-Carlo estimation of the Sobol indices. For the sake of self-completeness, we briefly present the aim and the computation of these indices, and we refer to [17], [16] and [8] for details.

2.1 Definition of the Sobol indices

For $i = 1, \ldots, p$, the i^{th} Sobol index of a function of p variables $s(\mu_1, \ldots, \mu_p)$ is defined by:

$$S_i = \frac{\operatorname{Var}\left(\mathbb{E}(s(\mu_1, \ldots, \mu_p)|\mu_i)\right)}{\operatorname{Var}\left(s(\mu_1, \ldots, \mu_p)\right)}, \tag{16}$$

the variances and conditional expectation being taken with respect to a postulated distribution of the $(\mu_1, \dots, \mu_p)$ input vector accounting for the uncertainty on the inputs' value. These indices are well defined as soon as $s \in L^2(\mathcal{P})$ and $\mathrm{Var}\,(s(\mu_1, \dots, \mu_p)) \neq 0$. When $\mu_1, \dots, \mu_p$ are (stochastically) independent, the i^{th} Sobol index can be interpreted as the fraction of the variance of the output that is caused by the uncertainty on the i^{th} parameter μ_i. All the Sobol indices lie in $[0; 1]$; the closer to zero (resp., one) S_i is, the less (resp., the more) importance μ_i's uncertainty has on s's uncertainty.

2.2 Estimation of the Sobol indices

The conditional expectation and variances appearing in (16) are generally not amenable to analytic computations. In those cases, one can estimate S_i by using a Monte-Carlo estimate: from two random, independent samples of size M of the inputs' distribution, we compute $2M$ appropriate evaluations $\{s_j\}$ and $\{s'_j\}$ of s, and estimate S_i by:

$$\widehat{S}_i = \frac{\frac{1}{M}\sum_{j=1}^M s_j s'_j - \left(\frac{1}{M}\sum_{j=1}^M s_j\right)\left(\frac{1}{M}\sum_{j=1}^M s'_j\right)}{\frac{1}{M}\sum_{j=1}^M s_j^2 - \left(\frac{1}{M}\sum_{j=1}^M s_j\right)^2}. \tag{17}$$

When M and/or the required time for the evaluation of the model output are large, it is computationally advantageous to replace s by its surrogate model $\widetilde{s}$. By using (17) on $\widetilde{s}$ (hence with reduced model outputs $\{\widetilde{s}_j\}$ and $\{\widetilde{s}'_j\}$), one estimates the Sobol indices *of the surrogate model* rather than those of the true model. We presented in [8], Sections 3.1 and 3.2, a method to quantify the error made in the Sobol index estimation when replacing the original model by the surrogate one. We defined two estimators $\widehat{S}^m_{i,\alpha_{as}/2}$ and $\widehat{S}^M_{i,1-\alpha_{as}/2}$, relying on output error bound samples $\{\epsilon_j\}$ and $\{\epsilon'_j\}$, and proved that:

Theorem 2.1. *If:*

$$\forall j = 1, \dots, M, \quad |s_j - \widetilde{s}_j| \le \epsilon_j \;\; \textit{and} \;\; \left|s'_j - \widetilde{s}'_j\right| \le \epsilon'_j,$$

then we have:

$$P\left(S_i \in [\widehat{S}^m_{i,\alpha_{as}/2}; \widehat{S}^M_{i,1-\alpha_{as}/2}]\right) \ge 1 - \alpha_{as}.$$

In our case, the output error bound $\epsilon(\mu)$ of Theorem 1.1 does not satisfy the above hypothesis, but satisfies a weaker "probabilistic" statement. This is the object of the following Corollary:

Corollary 2.2. *If:*

$$\forall j = 1, \dots, M, \quad P\left(|s_j - \widetilde{s}_j| \ge \epsilon_j\right) \le \alpha \;\; \textit{and} \;\; \forall j = 1, \dots, M, \quad P\left(\left|s'_j - \widetilde{s}'_j\right| \ge \epsilon'_j\right) \le \alpha,$$

then we have:

$$P\left(S_i \in [\widehat{S}^m_{i,\alpha_{as}/2}; \widehat{S}^M_{i,1-\alpha_{as}/2}]\right) \ge (1 - \alpha_{as}) \times (1 - \alpha)^{2M}.$$

Proof. We easily have that:

$$\begin{aligned} P\left(S_i \in [\widehat{S}^m_{i,\alpha_{as}/2}; \widehat{S}^M_{i,1-\alpha_{as}/2}]\right) &\geq P\left(S_i \in [\widehat{S}^m_{i,\alpha_{as}/2}; \widehat{S}^M_{i,1-\alpha_{as}/2}] \,|\, \forall j, |s_j - \widetilde{s}_j| < \epsilon(\mu)\right) \\ &\quad \times P\left(\forall j, |s_j - \widetilde{s}_j| < \epsilon(\mu)\right) \\ &\geq (1-\alpha_{as}) \times (1-\alpha)^{2M}. \end{aligned}$$

□

3 Numerical results I: Diffusion equation

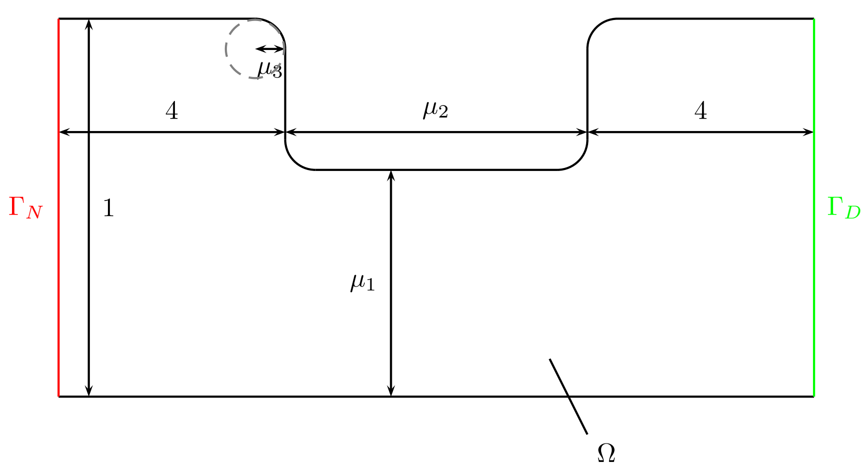
Important instances of problem (1) appear as discretizations of μ-parametrized linear partial differential equations (PDE); the X space is typically a finite element subspace (e.g., Lagrange P^1 finite elements), which we still see as identical to $\mathbb{R}^{\mathcal{N}}$. $A(\mu)$ and f are given by Galerkin projection of the weak form of the PDE onto a suitable basis of X. The boundary conditions of the PDE are usually either encoded in X or in $A(\mu)$, and the inner product used to perform the Galerkin projection is typically the L^2 or H^1 inner product. The use of the standard Euclidean product is justified by the fact that the relevant functional inner product has already been used to write $A(\mu)$ and $f(\mu)$, and the discrete matricial problem can be considered using the Euclidean inner product.

3.1 Benchmark problem

Our benchmark problem [15] is the following: given a parameter vector

$$\mu = (\mu_1, \mu_2, \mu_3) \in \mathcal{P} = [0.25, 0.5] \times [2, 4] \times [0.1, 0.2],$$

we consider the domain $\Omega = \Omega(\mu)$ below:



Our continuous field variable $u_e = u_e(\mu) \in X_e$ satisfies:

$$\begin{cases} \Delta u_e = 0 \text{ in } \Omega \\ u_e = 0 \text{ on } \Gamma_D \\ \frac{\partial u_e}{\partial n} = -1 \text{ on } \Gamma_N \\ \frac{\partial u_e}{\partial n} = 0 \text{ on } \partial\Omega \setminus (\Gamma_N \cup \Gamma_D) \end{cases} \tag{18}$$

where

$$X_e = \{v \in H^1(\Omega) \text{ s.t. } v|_{\Gamma_D} = 0\},$$

Δ denotes the Laplace operator, and $\frac{\partial}{\partial n}$ is the normal derivative with respect to $\partial\Omega$.

This continuous variable denotes the potential of a steady, incompressible flow moving in a tube whose profile is given by Ω, with open ends on Γ_N and Γ_D. The Neumann boundary condition on Γ_N states that the fluid enters by Γ_N with velocity equal to one, the condition on $\partial\Omega \setminus (\Gamma_N \cup \Gamma_D)$ states that the velocity field is tangential to the boundary of the tube; finally the Dirichlet condition on Γ_D guarantees well-posedness, as the potential field is determinated up to a constant.

The variational formulation of (18) states as follows: find $u_e = u_e(\mu) \in X_e$ so that:

$$\int_\Omega \nabla u_e \cdot \nabla v = -\int_{\Gamma_N} v, \quad \forall v \in X_e.$$

This variational problem is well-posed, as the bilinear form $(u, v) \mapsto \int_\Omega \nabla u \cdot \nabla v$ is coercive on X_e (see, for instance, [19], lemma A.14).

The above variational problem is discretized using a finite triangulation $\mathcal{T}$ of Ω and the associated $P^1(\mathcal{T})$ (see [3] or [14]) finite element subspace: find $u \in X$ so that

$$\int_\Omega \nabla u \cdot \nabla v = -\int_{\Gamma_N} v \quad \forall v \in X,$$

where $X = \{v \in \mathbf{P}^1(\mathcal{T}) \text{ s.t. } v|_{\Gamma_D} = 0\}$.

In our experiments, $\dim X = 525$.

The affine decomposition of the matrix of the bilinear form in the left-hand side of the above equation is obtained by using a piecewise affine mapping from $\Omega(\mu)$ to a reference domain $\bar{\Omega}$ as explained in [13], page 11.

Our scalar output of interest is taken to be:

$$s(\mu) = \int_{\Gamma_N} u(\mu),$$

and μ has uniform distribution on $\mathcal{P}$.

3.2 Results

We now present the numerical results obtained using the different error bounds on the output of the model described above. We report our bounds on the non-corrected and corrected outputs, as well as the dual-based output bound. Note that the stability constant $\alpha(\mu)$ is taken as the exact inf; this clearly advantages the dual-based output bound.

For the comparisons to be fair, one should compare the error bounds of same online cost. It is widely assumed that there exists a constant C so that this cost is $C \times 2(\dim \widetilde{X})^3$ for the dual-based method, and $C(\dim \widetilde{X})^3$ for our method, since dual-based method involves online inversion of two linear systems of size $\dim \widetilde{X}$, and one system of the same size for our method. Hence, the reported reduced basis sizes for the dual method are multiplied by a factor $\sqrt[3]{2}$.

In all cases, the reduced bases are computed using POD with snapshot size 80. To compute $\widehat{G}$, we use a snapshot of size 200. We also took $K = 1$ (ie., a trivial partition of $\mathcal{P}$). The truncation index N is taken equal to 20. We used a discrete minimization procedure to estimate the β_i constants.

In Figure 1, we compare the different error bounds on the non-corrected, and corrected output. For instance, for the error bound on the non-corrected output, we plot:

$$\bar{\epsilon} = \frac{1}{\#S} \sum_{\mu \in S} \widehat{\epsilon}(\mu, \alpha, N, \Phi)$$

where S is a random subset of $\mathcal{P}$ with size 200 and $\widehat{\epsilon}(\mu, \alpha, N, \Phi)$ is defined at (13). Other error bound means are computed accordingly.

We also computed the mean of the Lipschitz error bound ϵ^L. It is not reported here as it was way higher than dual-based output error bound. We observe that our new output error bound outperforms the dual-based error bound, at least for finite reduced bases sizes. Two reasons can be seen to this superiority: the difference in nature (probabilistic vs. sure) between the two bounds, and the fact that we make a crucial use of expansion (9) instead of using a Cauchy-Schwarz (or duality norm) argument. The rate of convergence (slope of the curve) of the corrected output is better than the non-corrected one, and this fact is reported by the two error bounds. Also, the expectation of $\widehat{T}_2$ was estimated at 10^{-12}, which allows to choose a low target risk and remain very competitive, as the α-dependency of the bound is in $\widehat{T}_2/\alpha$.

3.3 Application to sensitivity analysis

We estimate confidence intervals for the sensitivity indices of $s(\mu)$ by using the method described in [8], together with Corollary 2.2, and the non-corrected output.

We take $M = 1000$ as sample size, $B = 500$ as number of bootstrap replications (this parameter is used in the procedure which provides the confidence intervals from B replications of $\widehat{S}_i$, see [8]), $\dim \tilde{X} = 10$ as reduced basis size, $\alpha = 0.00001$ as output error bound risk, and $\alpha_{as} = 0.05$ as Monte-Carlo risk. The level of the combined confidence interval $\left[\widehat{S}^m_{i,\alpha_{as}/2}; \widehat{S}^M_{i,1-\alpha_{as}/2}\right]$ is then $(1-\alpha_{as})(1-\alpha)^M > 0.93$.

The results are gathered in Table 1. The spread between $\widehat{S}^m_i$ and $\widehat{S}^M_i$ accounts for the *metamodel-induced* error in the estimation of the Sobol indices. The remaining spread between $\widehat{S}^m_{i,\alpha_{as}/2}$ and $\widehat{S}^M_{i,1-\alpha_{as}/2}$ is the impact of the sampling

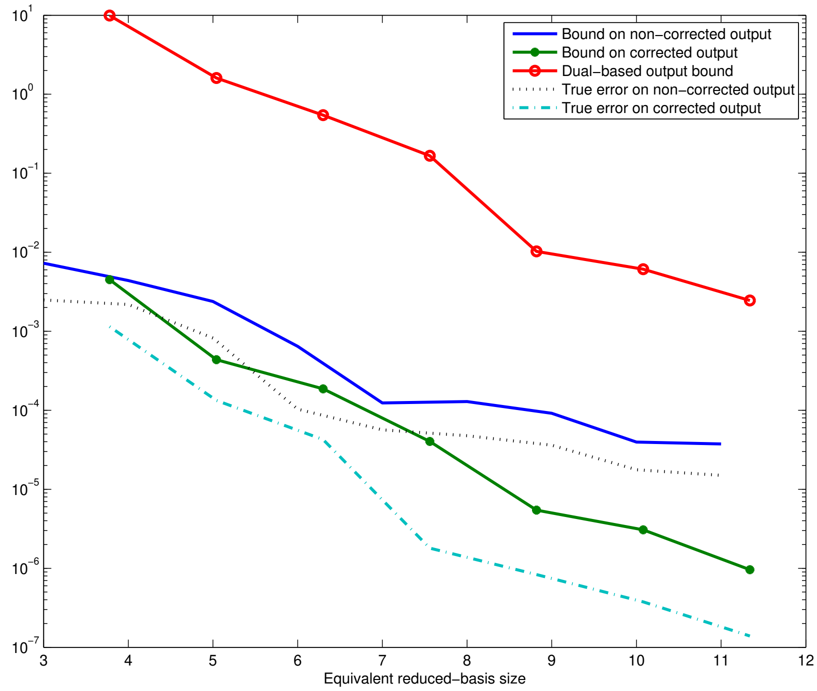


Figure 1: Comparison of the mean error bound on the non-corrected output, the mean dual-based error bound (ϵ_{cc}) and the mean error bound on the corrected output (for risk $\alpha = 0.0001$). The "equivalent" reduced basis sizes are in abscissae.

Input parameter	$\left[\widehat{S}_i^m; \widehat{S}_i^M\right]$	$\left[\widehat{S}_{i,\alpha_{as}/2}^m; \widehat{S}_{i,1-\alpha_{as}/2}^M\right]$
μ_1	[0.530352;0.530933]	[0.48132; 0.5791]
μ_2	[0.451537;0.452099]	[0.397962;0.51139]
μ_3	[0.00300247;0.0036825]	[-0.0575764;0.0729923]

Table 1: Results of the application of Section 2 to the estimation of the Sobol indices of the output of our benchmark model.

error (due to the replacement of the variances in the definition of the Sobol indices by their empirical estimators). We see that, in this case, the metamodel-induced error (certified by the use of our goal-oriented error bound) is very small with regard to the sampling error. We also notice that the estimate for the Sobol index for μ_3 is negative; this is not contradictory as it is the true value of the index that is in $[0, 1]$. For small indices, the estimate can be negative.

4 Numerical results II: transport equation

We now apply our error bound on a non-homogeneous linear transport equation. Compared to the previous example, the considered PDE is of a different kind (hyperbolic rather than elliptic).

4.1 Benchmark problem

In this problem, the continuous field $u_e = u_e(x,t)$ is the solution of the linear transport equation:

$$\frac{\partial u_e}{\partial t}(x,t) + \mu \frac{\partial u_e}{\partial x}(x,t) = \sin(x)\exp(-x)$$

for all $(x,t) \in]0,1[\times]0,1[$, satisfying the initial condition:

$$u_e(x, t=0) = x(1-x) \quad \forall x \in [0,1],$$

and boundary condition:

$$u_e(x=0,t) = 0 \quad \forall t \in [0,1].$$

The parameter μ is chosen in $\mathcal{P} = [0.5, 1]$ and $\mathcal{P}$ is endowed with the uniform measure.

We now choose a spatial discretization step $\Delta x > 0$ and a time discretization step $\Delta t > 0$, and we introduce our discrete unknown $u = (u_i^n)_{i=0,\ldots,N_x;n=0,\ldots,N_t}$ where

$$N_x = \frac{1}{\Delta x}, \quad \text{and} \quad N_t = \frac{1}{\Delta t}.$$

We note here that the considered PDE is hyperbolic and time-dependent, and that we perform the reduction on the space-time unknown u, of dimension

$(N_x+1)\cdot(N_t+1)$. This is different from reducing the space-discretized equation at each time step.
The u vector satisfies the discretized initial-boundary conditions:

$$\forall i, \;\; u_i^0 = (i\Delta x)(1 - i\Delta x) \tag{19}$$

$$\forall n, \;\; u_0^n = 0 \tag{20}$$

and the first-order upwind scheme implicit relation:

$$\forall i, n \;\; \frac{u_{i+1}^{n+1} - u_{i+1}^n}{\Delta t} + \mu \frac{u_{i+1}^{n+1} - u_i^{n+1}}{\Delta x} = \sin(i\Delta x)\exp(-i\Delta x). \tag{21}$$

Let us denote by $B = B(\mu)$ (resp. y) the matrix (resp. the vector) so that (19),(20) and (21) are equivalent to:

$$Bu = y \tag{22}$$

that is:

$$B^T B u = B^T y, \tag{23}$$

so that equation (23) is (1) with $A(\mu) = B^T B$ and $f = B^T y$.
The output of interest is: $s(\mu) = u_{N_x}^{N_t}$. In the following, we take $\Delta t = 0.02$ and $\Delta x = 0.05$. As in the previous example, the true stability constants are computed for the dual-based error bound.

4.2 Results

4.2.1 Comparison of the bounds

We took a very low risk level $\alpha = 0.0001$, a snapshot size of 200, $N = 20$ retained $\widehat{\phi}_i^G$ vectors and $K = 1$. The results (Figure 2) show that, once again, the error bounds we propose in this paper outperforms the dual-based error bound.

4.2.2 Choice of basis

For the comparison to be fair, we have also checked that using a POD basis does not penalize the dual-based error bound (ϵ^{cc}), by comparing the performance of this error bound when using a POD basis and a so-called "Greedy" [1] procedure. This procedure has a smaller offline cost than the POD, as it requires less resolutions of the reference problem (1). The results, shown in Figure 3, show that the Greedy procedure yields to inferior performance.

4.2.3 Coverage of the error bound

We have also estimated the actual error risk by doing computations of the error bound (on the corrected output), together with the actual error on a random sample of parameters of size 20000, using 10-sized bases for primal and dual problem. On this random sample, the estimated error bound is always greater than the true error; hence the error risk estimation appears conservative.

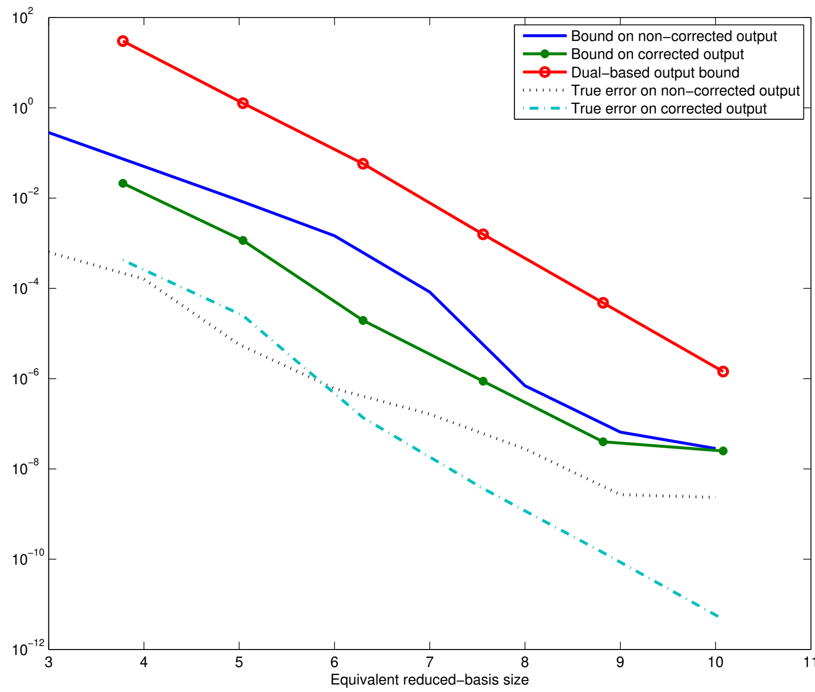


Figure 2: Comparison between the mean error bound (corr. and non-corr. outputs), the mean dual-based error bound, and the true (corr. and non-corr.) errors, for different reduced basis sizes.

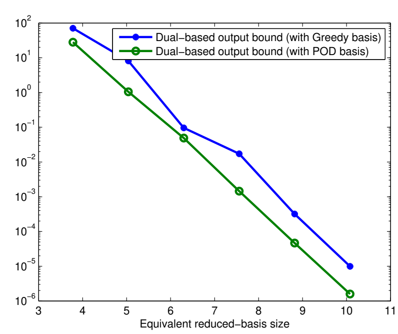


Figure 3: Mean of dual-based error bound, when using a POD or a Greedy basis selection procedure, for different reduced basis sizes.

Conclusion

We have presented a new explicitly computable output error bound for the reduced-basis method. We have shown, on two different practical examples, that this bound is clearly better than the naive Lipschitz bound and that, at the expense of a slight, controllable risk, the performances of this new bound are better than the ones of the existing dual-based output error bound.

Acknowledgements. This work has been partially supported by the French National Research Agency (ANR) through COSINUS program (project COSTA-BRAVA nr. ANR-09-COSI-015). We thank Anthony Nouy (École Centrale de Nantes) and Yvon Maday (Université Paris 6) for fruitful discussions, and the two anonymous referees for their pertinent remarks, which have greatly improved the quality of the paper.

A Estimation of $T_2(N, \Phi)$: error analysis

We now assess the error of estimation of $T_2(N, \Phi)$, and its consequence on the risk of the error bound $\widehat{\epsilon}(\mu, \alpha, N, \Phi)$.
First, by repeating the proof of Theorem 1.1, and by replacing (for the application of Markov inequality) $T_2(N, \Phi)/\alpha$ by $\widehat{T}_2(N, \Phi)/\alpha$, one gets

$$P\left(|s(\mu) - \widetilde{s}(\mu)| > T_1(\mu, N, \Phi) + \frac{\widehat{T}_2(N, \Phi)}{\alpha}\right) \leq \alpha \frac{T_2(N, \Phi)}{\widehat{T}_2(N, \Phi)}.$$

If $T_2(N, \Phi) = 0$, we have that

$$P\left(|s(\mu) - \widetilde{s}(\mu)| > T_1(\mu, N, \Phi) + \frac{\widehat{T}_2(N, \Phi)}{\alpha}\right) = 0,$$

hence the computable error bound has zero risk. So we can assume that $T_2(N, \Phi) \neq 0$.
The error on the risk of $\widehat{\epsilon}(\mu, \alpha, N, \Phi)$ is not easily attainable, because the same Ξ set of parameters is used to choose Φ and to compute $\widehat{T}_2(N, \Phi)$, leading to probabilistic dependence in the family:

$$\left\{ \left| \sum_{i=N+1}^{\mathcal{N}} \langle w(\mu), \Phi_i \rangle \langle r(\mu), \Phi_i \rangle \right| ; \mu \in \Xi \right\}.$$

However, one can take another random sample $\Xi' \subset \mathcal{P}$ of parameters, independent of Ξ, with size $M = \#\Xi'$, and define $\widehat{T}'_2(N, \Phi)$, the following estimator of $T_2(N, \Phi)$:

$$\widehat{T}'_2(N, \Phi) = \frac{1}{M} \sum_{\mu \in \Xi'} \left| \sum_{i=N+1}^{\mathcal{N}} \langle w(\mu), \Phi_i \rangle \langle r(\mu), \Phi_i \rangle \right|,$$

which gives in turn another error bound:

$$\widehat{\epsilon}'(\mu,\alpha,N,\Phi) = T_1(\mu,N,\Phi) + \frac{\widehat{T}_2'(N,\Phi)}{\alpha},$$

which is also computable in practice, but requires more solutions of the reference problem 1 during the offline phase.
Notice that $\widehat{T}_2'(N,\Phi)$ is a random variable with respect to the probability measure used to sample the Ξ' set. We denote by P' this probability measure. We denote by ρ the risk majorant of the computable error bound:

$$\rho = \alpha \frac{T_2(N,\Phi)}{\widehat{T}_2'(N,\Phi)},$$

which is also a random variable with respect to P'.
We have the following theorem:

Theorem A.1. *Let*

$$B(\mu) = \left| \sum_{i=N+1}^{\mathcal{N}} \langle w(\mu), \Phi_i\rangle \langle r(\mu), \Phi_i\rangle \right|, \quad \sigma^2 = \mathrm{Var}_{P'}(B(\mu)).$$

We have:

$$\sqrt{M}\,(\rho - \alpha) \underset{P'}{\longrightarrow} \mathcal{N}\left(0, \alpha^2 \frac{\sigma^2}{T_2^2}\right)$$

where $\underset{P'}{\longrightarrow}$ *denotes convergence in* P'*-distribution when* $M \to +\infty$ *and* $\mathcal{N}(0,\sigma)$ *is the centered gaussian distribution with variance* σ^2*.*

Hence ρ converges (in a probabilistic sense) to α with rate $1/\sqrt{M}$.

Proof. We have:

$$T_2(N,\Phi) = \mathbb{E}_{P'}(B(\mu)), \quad \widehat{T}_2' = \frac{1}{M} \sum_{\mu \in \Xi'} B(\mu).$$

Hence, by the central limit theorem [20],

$$\sqrt{M}(\widehat{T}_2 - T_2) \underset{P'}{\longrightarrow} \mathcal{N}(0,\sigma)$$

Now we define $g(x) = \alpha T_2(N,\Phi)/x$. As $T_2(N,\Phi) \neq 0$, g is differentiable in $T_2(N,\Phi)$ and one can use the Delta method [20] to write:

$$\sqrt{M}\left(g(\widehat{T}_2'(N,\Phi)) - g(T_2(N,\Phi))\right) \underset{P'}{\longrightarrow} \mathcal{N}\left(0, \alpha^2 \frac{\sigma^2}{T_2^2}\right),$$

which proves the theorem. $\square$

References

[1] 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. *Mathematical Modelling and Numerical Analysis*, 2009.

[2] T. Bui-Thanh, K. Willcox, O. Ghattas, and B. van Bloemen Waanders. Goal-oriented, model-constrained optimization for reduction of large-scale systems. *Journal of Computational Physics*, 224(2):880–896, 2007.

[3] P.G. Ciarlet. *The finite element method for elliptic problems*. Society for Industrial Mathematics, 2002.

[4] M.A. Grepl and A.T. Patera. A posteriori error bounds for reduced-basis approximations of parametrized parabolic partial differential equations. *Mathematical Modelling and Numerical Analysis*, 39(1):157–181, 2005.

[5] 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. *Comptes Rendus Mathematique*, 345(8):473–478, 2007.

[6] M. Ilak and C.W. Rowley. Modeling of transitional channel flow using balanced proper orthogonal decomposition. *Physics of Fluids*, 20:034103, 2008.

[7] A. Janon, M. Nodet, and C. Prieur. Certified reduced-basis solutions of viscous Burgers equations parametrized by initial and boundary values. Preprint available at `http://hal.inria.fr/inria-00524727/en`, 2010, *Accepted in Mathematical modelling and Numerical Analysis*.

[8] A. Janon, M. Nodet, and C. Prieur. Uncertainties assessment in global sensitivity indices estimation from metamodels. *International Journal for Uncertainty Quantification*, 2012.

[9] D.J. Knezevic, N.C. Nguyen, and A.T. Patera. Reduced basis approximation and a posteriori error estimation for the parametrized unsteady boussinesq equations. *Mathematical Models and Methods in Applied Sciences*, 2010.

[10] L. Machiels, A. T. Patera, J. Peraire, and Y. Maday. A general framework for finite element a posteriori error control: application to linear and nonlinear convection-dominated problems. In *Proceedings of ICFD Conference, Oxford*, 1998.

[11] Y. Maday, A.T. Patera, and D.V. Rovas. A blackbox reduced-basis output bound method for noncoercive linear problems. In Doina Cioranescu and Jacques-Louis Lions, editors, *Nonlinear Partial Differential Equations and their Applications Collège de France Seminar Volume XIV*, volume 31 of

Studies in Mathematics and Its Applications, pages 533 – 569. Elsevier, 2002.

[12] N.C. Nguyen, K. Veroy, and A.T. Patera. Certified real-time solution of parametrized partial differential equations. *Handbook of Materials Modeling*, pages 1523–1558, 2005.

[13] A. Quarteroni, G. Rozza, and A. Manzoni. Certified reduced basis approximation for parametrized partial differential equations and applications. *Journal of Mathematics in Industry*, 1(1):3, 2011.

[14] A.M. Quarteroni and A. Valli. *Numerical approximation of partial differential equations*. Springer, 2008.

[15] G. Rozza and Patera A.T. Venturi: Potential flow. *http://augustine.mit.edu/workedproblems/rbMIT/venturi/*, 2008.

[16] A. Saltelli. Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications*, 145(2):280–297, 2002.

[17] A. Saltelli, K. Chan, and E.M. Scott. *Sensitivity analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 2000.

[18] L. Sirovich. Turbulence and the dynamics of coherent structures. part i-ii. *Quarterly of applied mathematics*, 45(3):561–590, 1987.

[19] A. Toselli and O.B. Widlund. *Domain decomposition methods–algorithms and theory*. Springer Verlag, 2005.

[20] A. W. van der Vaart. *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998.

[21] K. Veroy and A.T. Patera. Certified real-time solution of the parametrized steady incompressible Navier-Stokes equations: Rigorous reduced-basis a posteriori error bounds. *International Journal for Numerical Methods in Fluids*, 47(8-9):773–788, 2005.

[22] S. Volkwein. Proper orthogonal decomposition and singular value decomposition. Spezialforschungsbereich F003 Optimierung und Kontrolle, Projektbereich Kontinuierliche Optimierung und Kontrolle, Bericht. Nr. 153, Graz, 1999.

[23] K. Willcox and J. Peraire. Balanced model reduction via the proper orthogonal decomposition. *AIAA journal*, 40(11):2323–2330, 2002.

[24] C. Zhu, R.H. Bryd, and J. Nocedal. *L-BFGS-B: Algorithm 778: L-BFGS-B, FORTRAN routines for large scale bound constrained optimization*, 1997.