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

Clémentine **PRIEUR**

MOMAS Multiphasiques 2015
Nice, October 5-7, 2015

INRIA RHÔNE-ALPES

Groupement de Recherche
MASCOT-NUM
Méthodes d'Analyse Stochastique pour les COdes et Traitements NUMériques

# My collaborators

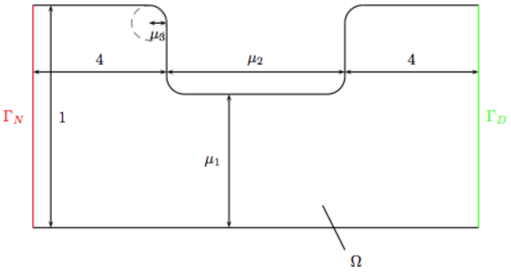
M. Nodet (UJF, Grenoble)

A. Janon (Orsay)

C. Prieur (CNRS, Grenoble)

Venturi's example [Rozza and A.T., 2008, Janon et al., 2012]

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



Let us define the continuous state variable $u_e = u_e(\mu) \in X_e$ as:

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

with

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

Define the output of interest as:

$$s(\mu) = s(u_e(\mu)) = \int_{\Gamma_N} u_e(\mu).$$

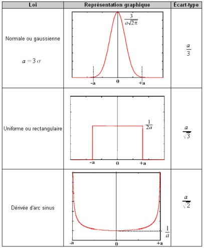
We are interested in determining the respective influence of parameters $\mu_1, \mu_2, \mu_3$ on $s(\mu)$.

Towards global sensitivity analysis:

Uncertainties on parameters $\mu_1, \mu_2, \mu_3$ are modeled by independent probability distributions.

independent random parameters

$\mu = (\mu_1, \mu_2, \mu_3)$



| Loi | Représentation graphique | Ecart-type |
| --- | --- | --- |
| Normale ou gaussienne $a = 3\sigma$ | | $\frac{a}{3}$ |
| Uniforme ou rectangulaire | | $\frac{a}{\sqrt{3}}$ |
| Dérivée d'arc sinus | | $\frac{a}{\sqrt{2}}$ |

$\longrightarrow$

random output

$s(\mu) \in \mathbb{R}$

In the previous example, let $\mu$ be modeled by a random vector distributed uniformly on $\mathcal{P} = [0.25, 0.5] \times [2, 4] \times [0.1, 0.2]$.

Is the output of interest $s(\mu)$ more or less variable when setting one of the parameters $\mu_i$, $i = 1, 2, 3$ to a nominal value?

$\mathrm{Var}\,(s(\mu)|\mu_i = \mu_{i,0})$, how to choose $\mu_{i,0}$?

## A toy example

Is the output of interest $s(\mu)$ more or less variable when setting one of the parameters $\mu_i$, $i = 1, 2, 3$ to a nominal value?

$\mathrm{Var}\left(s(\mu)|\mu_i = \mu_{i,0}\right)$, how to choose $\mu_{i,0}$? $\Rightarrow$ $\mathbb{E}\left[\mathrm{Var}\left(s(\mu)|\mu_i\right)\right]$

Is the output of interest $s(\mu)$ more or less variable when setting one of the parameters $\mu_i$, $i = 1, 2, 3$ to a nominal value?

$\mathrm{Var}\left(s(\mu)|\mu_i = \mu_{i,0}\right)$, how to choose $\mu_{i,0}$? $\Rightarrow \mathbb{E}\left[\mathrm{Var}\left(s(\mu)|\mu_i\right)\right]$

**Theorem (total variance)**

$\mathrm{Var}(s(\mu)) = \mathrm{Var}\left[\mathbb{E}\left(s(\mu)|\mu_i\right)\right] + \mathbb{E}\left[\mathrm{Var}\left(s(\mu)|\mu_i\right)\right].$

**Definition (First-order Sobol' indices)**

$i = 1, 2, 3$

$$0 \leq S_i = 1 - \frac{\mathbb{E}\left[\mathrm{Var}\left(s(\mu)|\mu_i\right)\right]}{\mathrm{Var}(s(\mu))} = \frac{\mathrm{Var}\left[\mathbb{E}\left(s(\mu)|\mu_i\right)\right]}{\mathrm{Var}(s(\mu))} \leq 1$$

A value close to 1 (*resp.* a small value) for the first-order Sobol' index $S_i$ means that $\mu_i$ has many (*resp.* little) influence on $s(\mu)$.

Estimation procedure:

Monte Carlo based estimation procedures for Sobol' indices require many evaluations of the output of interest $s(\mu)$. It is possible to obtain confidence intervals.

What happens if the output $s(\mu)$ is costly to evaluate?

It is then possible to approximate $s(\mu)$ by $\widetilde{s}(\mu) = s(\widetilde{u}(\mu))$ with $\widetilde{u}(\mu)$ an approximation of the state variable $u(\mu)$.

What about confidence intervals in that case?

The approximation error $|s(\mu) - \widetilde{s}(\mu)|$ has to be taken into account in their construction [Janon et al., 2014b, Janon et al., 2014a].

I- Problem statement

II- Goal oriented probabilistic error bound

III- Correction of the output

IV- Numerical example

Let $\mathcal{P}$ be the parameter space. For any $\mu \in \mathcal{P}$, let $u(\mu)$ be the solution of the linear system:

$$A(\mu)u(\mu) = f(\mu),$$

with $A(\mu)$ ($f(\mu)$) a known $\mathcal{N} \times \mathcal{N}$ matrix ($\mathcal{N} \times 1$ vector).

The linear output of interest is defined as: $s(\mu) = \ell^\top u(\mu)$.

Let $\mathcal{P}$ be the parameter space. For any $\mu \in \mathcal{P}$, let $u(\mu)$ be the solution of the linear system:

$$A(\mu)u(\mu) = f(\mu),$$

with $A(\mu)$ ($f(\mu)$) a known $\mathcal{N} \times \mathcal{N}$ matrix ($\mathcal{N} \times 1$ vector).

The linear output of interest is defined as: $s(\mu) = \ell^\top u(\mu)$.

Our toy example:

The variational formulation for the Venturi's example is:

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

Let $\mathcal{T}$ be a finite triangulation of $\Omega$ and $\mathbf{P}^1(\mathcal{T})$ the associated finite element subspace. The above problem is discretized as follows:

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

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

Reduced basis method:

Let $\widetilde{X}$ be a subspace of $X$ of dimension $n << \mathcal{N}$. Let $Z$ be the $\mathcal{N} \times n$ matrix whose columns are the components of a (reduced) basis of $\widetilde{X}$ in the canonical basis of $X$.

Let $\widetilde{u}(\mu)$ be the $n \times 1$ vector solution of:

$$(Z^\top A(\mu) Z)\widetilde{u}(\mu) = Z^\top f(\mu).$$

Define the approximated output as:

$$\widetilde{s}(\mu) = \ell^\top Z \widetilde{u}(\mu) \approx \ell^\top u(\mu) = s(\mu).$$

Under some (more or less technical) hypotheses on $A(\mu)$ and on the norm $\|\cdot\|$ (say, Euclidean norm), the reduced basis comes with an error bound $\epsilon^u(\mu)$:

$$\forall \mu \in \mathcal{P}, \|u(\mu) - Z\widetilde{u}(\mu)\| \leq \epsilon^u(\mu)$$

which can be computed *efficiently* (i.e., with the order of complexity of the computation of $\widetilde{u}(\mu)$).

**Question:** Can we deduce from it an error bound $\epsilon(\mu)$ on $s$:

$$\forall \mu \in \mathcal{P},\ |s(\mu) - \widetilde{s}(\mu)| \leq \epsilon(\mu)$$

which can be explicitly and efficiently computed ?

**Answer:** Yes, as the “Lipschitz bound” holds:

$$\forall \mu \in \mathcal{P}, |s(\mu) - \widetilde{s}(\mu)| \leq L\epsilon^u(\mu),\ \text{with } L = \sup_{\|v\|=1} \ell^\top v.$$

We aim at improving this "Lipschitz" bound.

By **efficiently computable**, one means that the computation may be divided in two phases:

an *offline* phase during which quantities not depending on $\mu$ are computed, this phase can be relatively costly;

an *online* phase during which quantities depending on $\mu$ are computed, this phase has to be efficient.

Indeed, let $c_{off}$ (*resp.* $c_{on}$) be the cost of the *offline* (*resp. online*) phase. If one wants to estimate all first-order Sobol' indices, one needs to evaluate $2N$ times the model, with $N$ the size of the Monte Carlo sample. Thus the cost is $c_{off} + N \times c_{on}$.

Starting point:

Usually, the bound $\epsilon^u(\mu)$ on $\|u(\mu) - \widetilde{u}(\mu)\|$ is based on the **residual:**

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

and its **norm** is efficiently computable.

We also want to exploit that the (say, Euclidean) scalar products of the residual, $\langle r(\mu), \phi \rangle$, are efficiently computable $\forall\, \phi \in X$, in the usual setting of reduced basis for affinely parametrized PDE.
Let $\{\phi_i\}_{i=1,\dots,\mathcal{N}}$ be an orthonormal basis of $X$ (to be chosen later).
We have:

$$\widetilde{s}(\mu) - s(\mu) = \sum_{i \geq 1} \langle w(\mu), \phi_i \rangle \langle r(\mu), \phi_i \rangle,$$

where $w(\mu)$ is the solution of the so-called **adjoint** (or **dual**) problem:

$$w(\mu) = A(\mu)^{-\top} \ell.$$

Let $K \in \mathbb{N}^*$. We have:

$$|\widetilde{s}(\mu) - s(\mu)| \leq \left|\sum_{i=1}^{K} \langle w(\mu), \phi_i\rangle \langle r(\mu), \phi_i\rangle\right| + \left|\sum_{i>K} \langle w(\mu), \phi_i\rangle \langle r(\mu), \phi_i\rangle\right|.$$

Consider first

$$\tau_1(\mu) := \left|\sum_{i=1}^{K} \underbrace{\langle w(\mu), \phi_i\rangle}_{\textbf{to bound}} \overbrace{\langle r(\mu), \phi_i\rangle}^{\text{computable}}\right|$$

We compute (once for all the values of $\mu$):

$$\beta_i^{min} = \min_{\mu \in \mathcal{P}} D_i(\mu), \quad \beta_i^{max} = \max_{\mu \in \mathcal{P}} D_i(\mu),$$

where: $D_i(\mu) = \langle w(\mu), \phi_i\rangle$, ($2K$ optimization problems on $\mathcal{P}$).

Let

$$\beta_i^{up}(\mu) = \begin{cases} \beta_i^{max} \text{ if } \langle r(\mu), \phi_i\rangle > 0 \\ \beta_i^{min} \text{ else,} \end{cases} \quad \beta_i^{low}(\mu) = \begin{cases} \beta_i^{min} \text{ if } \langle r(\mu), \phi_i\rangle > 0 \\ \beta_i^{max} \text{ else.} \end{cases}$$

We then have:

$$|\tau_1(\mu)| \leq \max\left( \left| \sum_{i=1}^{K} \langle r(\mu), \phi_i \rangle \beta_i^{low}(\mu) \right|, \ \left| \sum_{i=1}^{K} \langle r(\mu), \phi_i \rangle \beta_i^{up}(\mu) \right| \right) =: T_1(\mu).$$

We then have:

$$|\tau_1(\mu)| \leq \max\left( \left| \sum_{i=1}^{K} \langle r(\mu), \phi_i \rangle \beta_i^{low}(\mu) \right|, \ \left| \sum_{i=1}^{K} \langle r(\mu), \phi_i \rangle \beta_i^{up}(\mu) \right| \right) =: T_1(\mu).$$

Let now $\tau_2(\mu) = |\sum_{i>K} \langle w(\mu), \phi_i \rangle \langle r(\mu), \phi_i \rangle|$.

This term is not efficiently computable. We assume that $\mu$ is a random vector on $\mathcal{P}$, with known distribution. Let us control $\mathbb{E}_\mu [\tau_2(\mu)]$ by:

$$\frac{1}{2}\mathbb{E}_\mu \left( \sum_{i>K} \langle w(\mu), \phi_i \rangle^2 + \sum_{i>K} \langle r(\mu), \phi_i \rangle^2 \right) = \sum_{i>K} \langle G\phi_i, \phi_i \rangle$$

where $G$ is the positive, self-adjoint operator defined by:

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

Recall that:

$$\mathbb{E}_\mu\left[\tau_2(\mu)\right] \leq \sum_{i>K} \langle G\phi_i, \phi_i \rangle.$$

Let $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_{\mathcal{N}} \geq 0$ be the eigenvalues of $G$, and $\phi_i^G$ be a unitary eigenvector of $G$ with respect to $\lambda_i$. The term $\sum_{i>K} \langle G\phi_i, \phi_i \rangle$ is minimized for $\phi_i = \phi_i^G \; \forall i > K$.

With this choice for the orthonormal basis $\{\phi_1, \ldots, \phi_{\mathcal{N}}\}$, we get an *a priori* bound on $\tau_2$:

$$\mathbb{E}_\mu\left[\tau_2(\mu)\right] \leq \sum_{i>K} \lambda_i^2.$$

We now define $T_2 = \mathbb{E}_\mu\left[\tau_2(\mu)\right]$.

In [Janon et al., 2012] is described a way of estimating the $\phi_i^G$ with a cost independent of the dimension $\mathcal{N}$ of $X$.

The objective is some probabilistic error bound between $s(\mu)$ and $\widetilde{s}(\mu)$. In other words, one accepts the risk of this bound being violated for a set of parameters having "small" probability measure.

**Theorem ([Janon et al., 2012])**

*Let* $\alpha \in (0,1)$, $\mathbb{P}_\mu \left( |s(\mu) - \widetilde{s}(\mu)| > T_1(\mu) + \frac{T_2}{\alpha} \right) < \alpha$.

Proof:

$$\mathbb{P}\left( |\widetilde{s}(\mu) - s(\mu)| > T_1(\mu) + \frac{T_2}{\alpha} \right)$$

$$\leq \mathbb{P}\left( |\widetilde{s}(\mu) - s(\mu)| > \left| \sum_{i=1}^{K} \langle r(\mu), \phi_i \rangle \langle w(\mu), \phi_i \rangle \right| + \frac{T_2}{\alpha} \right)$$

$$\leq \mathbb{P}\left( \left| \sum_{i=K+1}^{\mathcal{N}} \langle r(\mu), \phi_i \rangle \langle w(\mu), \phi_i \rangle \right| > \frac{T_2}{\alpha} \right).$$

Then, by Markov Inequality and by definition of $T_2$ we get:

$$\mathbb{P}\left(\left|\sum_{i=K+1}^{\mathcal{N}} \langle r(\mu), \phi_i \rangle \langle w(\mu), \phi_i \rangle\right| > \frac{T_2}{\alpha}\right)$$

$$\leq \frac{\mathbb{E}_\mu\left(\left|\sum_{i=K+1}^{\mathcal{N}} \langle r(\mu), \phi_i \rangle \langle w(\mu), \phi_i \rangle\right|\right)}{\frac{T_2}{\alpha}} = \frac{\mathbb{E}_\mu\left[\tau_2(\mu)\right]}{\frac{T_2}{\alpha}} = \alpha. \quad \square$$

Monte Carlo estimation of $T_2$: in practice, we estimate $T_2$ by:

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

once for all the values of $\mu$.

The adjoint (dual) problem $A(\mu)^\top w(\mu) = \ell$, can also be projected by using a matrix $Z_d$:

$$[Z_d^\top A(\mu)^\top Z_d]\widetilde{w}(\mu) = Z_d^\top \ell,$$

providing an approximation $Z_d \widetilde{w}(\mu) \approx w(\mu)$.

Computation of $\widetilde{w}(\mu)$ generally doubles the computational time, but allows to compute a *corrected* output approximation for $s(\mu)$:

$$\widetilde{s}_c(\mu) = \widetilde{s}(\mu) - \langle Z_d \widetilde{w}(\mu), r(\mu)\rangle,$$

which is known to be "more precise" than $\widetilde{s}(\mu)$.

The adjoint (dual) problem $A(\mu)^\top w(\mu) = \ell$, can also be projected by using a matrix $Z_d$:

$$[Z_d^\top A(\mu)^\top Z_d]\widetilde{w}(\mu) = Z_d^\top \ell,$$

providing an approximation $Z_d \widetilde{w}(\mu) \approx w(\mu)$.

Computation of $\widetilde{w}(\mu)$ generally doubles the computational time, but allows to compute a *corrected* output approximation for $s(\mu)$:

$$\widetilde{s}_c(\mu) = \widetilde{s}(\mu) - \langle Z_d \widetilde{w}(\mu), r(\mu)\rangle,$$

which is known to be "more precise" than $\widetilde{s}(\mu)$.

More specifically, it has been proved in [Nguyen et al., 2005] that

$$|\widetilde{s}_c(\mu) - s(\mu)| \leq \epsilon_u(\mu)\epsilon_u^d(\mu),$$

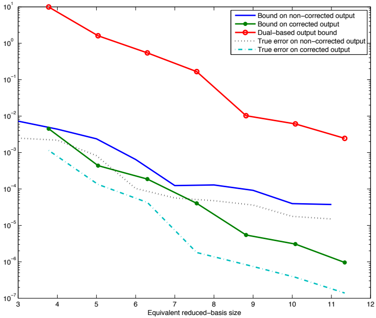
with $\|w(\mu) - Z_d \widetilde{w}(\mu)\| \leq \epsilon_u^d(\mu)$.

Our error bound can be readily extended so as to provide a bound $\epsilon_c(\mu)$ on the corrected output:

$$|\widetilde{s}_c(\mu) - s(\mu)| \leq \epsilon_c(\mu),$$

in probability (with respect to $\mu$), by changing every $w(\mu)$ by $w(\mu) - Z_d \widetilde{w}(\mu)$.

We come back to our toy model with $s(\mu) = \int_{\Gamma_N} u(\mu)$, $\mathcal{P} = [0.25, 0.5] \times [2, 4] \times [0.1, 0.2]$. We choose: $\mathcal{N} = 525$, the reduced basis is a POD computed from a snapshot of size 80, $K = 20$, $\#\Xi = 200$, the risk level $\alpha = 0.0001$



Results for first-order sensitivity analysis:

| input parameter | $\left[\widehat{S}_i^m; \widehat{S}_i^M\right]$ | $\left[\widehat{S}_{i,\alpha_{as}/2}^m; \widehat{S}_{i,1-\alpha_{as}/2}^M\right]$ |
|---|---|---|
| $\mu_1$ | [0.530352;0.530933] | [0.48132; 0.5791] |
| $\mu_2$ | [0.451537;0.452099] | [0.397962;0.51139] |
| $\mu_3$ | [0.00300247;0.0036825] | [-0.0575764;0.0729923] |

Table : Certification via bootstrap : size of the Monte Carlo sample 1000, bootstrap $B = 500$, reduced basis size $m = 10$, confidence level 0.95.

A second toy example: a non-homogeneous linear transport equation.

The continuous field $u_e = u_e(x,t)$ is the solution of:

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

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

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

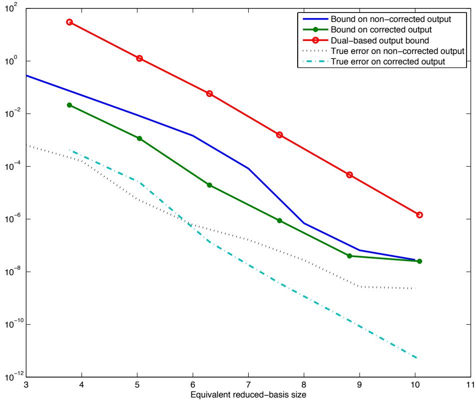
and boundary condition:

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

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

The discretization procedure is classical and described in [Janon et al., 2012].

"Tuning" parameters $K$, $\#\Xi$ and $\alpha$ are set as previously.




Comparisons are performed for the same online cost for the different bounds (factor of $\sqrt[3]{2}$).

## Conclusion

- we obtained a goal-oriented probabilistic error estimation for linear problems and linear outputs,
- this bound can be computed efficiently in an *offline*/*online* procedure,
- we applied such bounds to provide confidence intervals for sensitivity indices,
- during the talk I provided illustrations on toy examples.

**Generalizations, perspectives**

- application to more complex models, see *e.g.*, in [Janon et al., 2014c], an application to the sensitivity analysis for a flow control problem with linearized Shallow water equation,
- probabilistic goal-oriented error estimation for nonlinear models (ongoing work),
- probabilistic goal-oriented error estimation for nonlinear outputs (ongoing work),
- what happens if one consider sensitivity indices which are not based on variance?
- ...

# Some references I


Janon, A., Klein, T., Lagnoux, A., Nodet, M., and Prieur, C. (2014a).
Asymptotic normality and efficiency of two Sobol' index estimators.
*ESAIM P&S*, 18:342–364.

Janon, A., Nodet, M., and Prieur, C. (2012).
Goal-oriented error estimation for reduced basis method, with application to certified sensitivity analysis.
`http://hal.archives-ouvertes.fr/hal-00721616`.

Janon, A., Nodet, M., and Prieur, C. (2014b).
Uncertainties assessment in global sensitivity indices estimation from metamodels.
*International Journal for Uncertainty Quantification*, 4 (1):21–36.

Janon, A., Nodet, M., Prieur, C., and Prieur, C. (2014c).
Global sensitivity analysis for the boundary control of an open channel.
`http://hal.archives-ouvertes.fr/hal-01092590`.

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

Rozza, G. and A.T., P. (2008).
Venturi: Potential flow.
*http://augustine.mit.edu/workedproblems/rbMIT/venturi/*.