Adaptive reduced basis strategy based on goal oriented error assessment for stochastic problems

Eric Florentin a,*, Pedro Díez b

a LMTCachan (ENS Cachan) CNRS/UPMC PRES Université Paris XI, 61, avenue du Président Wilson, F-94235 Cachan, France
b Laboratori de Càlcul Numèric (LaCaN), Universitat Politècnica de Catalunya, C/O Campus Nord UPC, E-08034 Barcelona, Spain

Abstract

In the framework of stochastic non intrusive finite element modeling, a common practice is using Monte Carlo simulation. The main drawback of this approach is the computational cost, because it requires computing a large number of deterministic finite element solutions. The different Monte Carlo samplings correspond to realizations of the random variables characterizing the stochastic behavior of the model. Thus, this requires solving a set of deterministic problems with the same structure, that is with variations concerning the material parameters and the loading data. Consequently, the different problems to be solved are in practice similar to each other. The reduced basis strategy is therefore a sensible option to reduce computational cost, provided that the quality of the numerical solution is guaranteed. The paper introduces a goal oriented strategy allowing to successively enrich the reduced basis along the Monte Carlo process. The method is based on assessing the error of the reduced basis solution with a residual estimate for the prescribed quantity of interest. The efficiency of the proposed approach, which is particularly important if the number of independent random variables is large, is illustrated in 1D and 2D mechanical examples.

1. Introduction

Many authors have recently shown interest in developing numerical methods to analyze and design engineering systems which are modeled including randomness. In the mechanical engineering framework, the resulting governing equations are, in general, stochastic partial differential equations. Thus, stochastic finite element methods (FEM) are currently an essential tool for the quantitative prediction of the response of mechanical models. A key issue in this context, both for the research and industrial players, is reducing the computational cost in order to afford dealing with large scale applications. A state of the art for stochastic methods can be found in [33,27,34] and references therein.

The so called intrusive stochastic finite element methods are generalizations of the finite element method accounting for uncertainties associated with the parameters of the problem. Both the variation in the standard deterministic variables (space and time coordinates) and the random variables describing stochasticity are discretized using the standard approach in the FEM, that is the Galerkin formulation [11,8,23,2]. The key point of these methods is properly selecting the approximation space for the stochastic part. They are particularly well suited when the randomness is known and it is described by smooth functions. Despite of their high accuracy (the randomness is analytically accounted for), these methods require a huge computational effort because they lead to very large algebraic systems of equations. Moreover, the resulting structure of the codes is conditioned by the randomness and generally the standard code used for the deterministic simulations cannot be reused (this is the reason of naming these methods intrusive). In order to reduce the cost associated with these methodologies, different authors have proposed strategies based on designing iterative methods specially adapted to the structure of the resulting matrices [30,12,17] and using reduced basis methodologies for random fields [23,9,26].

The non intrusive stochastic techniques are more popular because they allow reusing standard deterministic codes. The random character of the model is accounted for using the Monte Carlo sampling approach, that is generating a large number of realizations of the random variables describing the stochasticity and computing the corresponding realization of the quantity of interest by solving a deterministic problem, see [28]. This approach is conceptually simple, robust and easy to implement but requires solving a large number of FE problems in order to generate an output sample allowing a proper statistical analysis.

Other associated methods, introducing further simplifications or increasing the computational efficiency are those using response surfaces [18], projections [31] or fitting by regression [35]. All

* Corresponding author.
E-mail address: eric.florentin@lmte.ens-cachan.fr (E. Florentin).
these methods are based on selecting the proper representative points in the random space in order to fit the function describing the stochastic variation.

Goal oriented error estimators address the issue of verification of the FE models, that is they focus on assessing the numerical quality or the accuracy of the FE solutions. These numerical tools have been developed and extensively used for standard deterministic FE models, see [20,32,29] for review. Generalizations of some of these error estimators to the stochastic framework have been published recently [19,17].

In [6], the authors introduced a reduced basis methodology to reduce the cost of Monte Carlo simulations, offering an attractive framework for solving stochastic problems with a large number of parameters. The idea is simple and effective because the different Monte Carlo shots lead to similar FE problems and therefore the reduced basis approach is highly performant. Other interesting recent references on this subject are [32,14,25,15,24,5].

The objective of the present work is offering an automatically adapted strategy to generate the reduced basis, which is enriched along the Monte Carlo process, including additional components to the basis only when they are needed to reach some prescribed accuracy. Thus, a criterion is needed to decide if the current sample is satisfactorily approximated with the current reduced basis or if it requires a complete resolution of the system (which will be added as a new component of the reduced basis). This criterion is based on goal oriented residual type error estimates, designed for some quantity of interest that has to be defined by the user.

The remainder of the paper is structured as follows: Section 2 presents the problem statement and discusses the non intrusive solving scheme. In Section 3, the standard reduced basis strategy is recalled and the algorithm generating the adapted reduced basis is presented. The performance of the proposed methodology is demonstrated in Section 4 with academic examples that can be compared with another strategy (a method using the concept of the response surface [18,4,22]). In Section 4, further numerical results are provided, including physical structures of engineering interest.

2. Stochastic problem

2.1. Problem statement equations

Let us consider a bounded domain Ω, representing a mechanical structure. The boundary of Ω, ∂Ω, is divided in two parts ∂hΩ and ∂dΩ such that ∂hΩ ∪ ∂dΩ = ∂Ω, ∂hΩ ≠ ∅ and ∂dΩ ∩ ∂hΩ = ∅. A prescribed displacement field uh is imposed on ∂hΩ and a traction gh is applied on ∂dΩ. The description of the loads is completed by a body force field fh in Ω.

The material model is assumed to be linear elastic, being K the Hooke tensor. The stochastic behavior of the model is introduced assuming that the Hooke tensor is a random field K(x, θ), where x ∈ Ω stands for the position vector and θ ∈ Θ is used to denote the randomness. The sample space Θ is the set of possible outcomes of θ. Without loss of generality, in the following random parameter is restricted to the material parameters introduced in K. In a general context, also f, g, and uh can be random fields.

Thus, the problem reads: find the unknown displacement field u(x, θ) (it is a random field, as indicated by its dependence on θ) such that

\[\begin{align*}
\text{div}(K(x, θ)u(x, θ)) + f(x, θ) &= 0 \quad \text{in } Ω, \\
K(x, θ)u(x, θ) \cdot n &= g(x, θ) \quad \text{on } ∂hΩ, \\
u(x, θ) &= u_0(x, θ) \quad \text{on } ∂dΩ
\end{align*}\]  

(1a) (1b) (1c)

The strong form of the problem, (1), is equivalent to the standard weak form reading as follows: find u(x, θ) such that

\[a(u(x, θ), w(x)) - \ell(w(x)) = 0 \quad \forall w(x) \in U\]  

(2)

where \(a(\cdot, \cdot)\) is the bilinear form expressing the energy product, \(\ell(\cdot)\) is the linear form including the loading and \(U\) the set of admissible displacements, satisfying (1c).

2.2. Non intrusive solving scheme

The non intrusive approach decouples the discretization of the physical space and the stochastic space, represented here by Ω and Θ. This can be described in two steps. In the first step, a standard finite element discretization is introduced to solve the problem (2) and obtain a numerical solution approximating \(u(x, θ)\) for a realization of θ (freezing the randomness). The second step consists in using a Monte Carlo simulation to obtain an approximation of the probability density function of \(u(x, θ)\). The two steps are described in order to introduce the necessary notations.

Step 1: Finite element discretization.

As mentioned above, in this phase, the problem is considered as if it was deterministic, that is for a given value of θ. The discrete finite element space is characterized by the standard finite element functions \(N_i(x)\), \(i = 1, 2, \ldots, N_{FE}\) generating the finite dimension sub space \(U_0 \subset U\)

\[U_0 = \text{span}\{N_1, N_2, \ldots, N_{N_{FE}}\}\]  

(3)

The numerical approximation in the space defined in (3) is \(u_0\), such that

\[u(x, θ) = u_0(x, θ) = \sum_{i=1}^{N_{FE}} u_i(θ) N_i(x), \quad N_i(x)U_0\]  

(4)

being \(u_0\) the finite element unknowns (nodal values), \(U_0 = \{u_0, u_1, \ldots, u_{N_{FE}}\}\) and \(N(x)\) \(\{N_1(x), N_2(x), \ldots, N_{N_{FE}}(x)\}\). The corresponding discretized form of (2) is the linear system of equations

\[K(θ)U(θ) = F(θ)\]  

(5)

where \(K(θ)\) is the finite element stiffness matrix (discretized form of \(K\)) and \(F(θ)\) is the load vector (discretized form of \(f\)).

Step 2: Monte Carlo simulation.

As previously mentioned, the stochastic character of the problem is described by the random variable θ. More precisely, the input data for the problem is the probability density function (PDF) of θ. Note that the solution \(U(θ)\) is also a random variable, resulting from a complex functional transformation of θ. The non intrusive approach seeks approximating the PDF of \(U(θ)\) (or some scalar quantity of interest (QoI) depending of it) without any analytical determination of this complex functional transformation. This is performed using the Monte Carlo technique.

The Monte Carlo technique consists in generating a number \(N_{MC}\) of realizations of θ. Note that these realizations are generated using the actual PDF \(f(θ)\). This is equivalent to determine \(N_{MC}\) realizations of \(K(θ)\) and therefore, solving \(N_{MC}\) linear systems of Eqs. (5), obtaining \(N_{MC}\) realizations of \(U(θ)\). Thus, the PDF of \(U(θ)\) is approximated from these realizations. In practice, the PDF is characterized by its moments (mean, variance) and the different percentiles.

In order to have an accurate approximation of the PDF of \(U(θ)\), the number of realizations \(N_{MC}\) must be very large. Note that the Monte Carlo simulation process converges very slowly to the exact answer, which is ideally obtained after an infinite number of realizations. The convergence rate is estimated to be of order \(O(1/\sqrt{N_{MC}})\), see for example [10].

Summarizing, the non intrusive strategy is extremely simple because it decouples the approximation of the stochastic behavior and the solution of the deterministic mechanical model. The deterministic numerical solver can be used as a black box, without any modification in the code associated with the stochasticity. The random character is accounted for by generating a large number of
realizations and solving the corresponding deterministic problems. The numerical cost is consequently very large.

2.3. Quantity of interest

The goal of the computation is often to obtain (or characterize) some specific Quantity of Interest. Thus, in this context, the QoI is also a random variable and attention is paid to the determination of its PDF. The QoI is denoted by \( Q(\theta) \) and it is given as a linear out put of \( u(x, \theta) \), namely

\[
Q(\theta) = \ell_Q(u(x, \theta)),
\]

being \( \ell_Q(\cdot) \) a deterministic linear form, in the same format as \( \ell(\cdot) \).

An auxiliary dual problem is introduced in order to obtain an error representation. This is standard in goal oriented error assessment techniques. The dual problem reads: find \( v(x, \theta) \) such that

\[
\alpha(v(x, \theta), w(x)) \mid \ell_Q(w(x)), \forall w(x) \in U
\]

The dual problem is discretized in the same finite element space as the direct problem. This results in a linear system of equation similar to (5)

\[
\mathbf{K}(\theta) v(x) = \mathbf{G}(\theta).
\]

where \( \mathbf{G} \) is the discrete counterpart of \( \ell_Q \) and \( V \) is the vector of finite element unknowns.

Then, for the space discretization corresponding to \( U_h \), the approximated value of the quantity of interest is directly computed as the following vector product

\[
Q_h = Q_h(\theta) = \ell_Q(u_h(x)) \mathbf{G^T}(\theta) U_h.
\]

Two alternative expressions for \( Q_h \) as a function of the solution of the primal (or direct) and the dual problems are recovered by recalling Eqs. (5) and (8),

\[
Q_h = \mathbf{V}^T \mathbf{U} \ell_Q(\mathbf{V}) \mathbf{U}^T \mathbf{F} \ell(\mathbf{U})
\]

where \( \ell_Q \) is the basis function associated with the solution of the dual problem. In practice, that means that, if the exact solution of the dual problem, \( v(x, \theta) \), was available, then the error in the QoI defined as

\[
e_\theta : Q \rightarrow Q_h
\]

is found to be an explicit function of the residual in the weak form of the problem, namely

\[
e_\theta = \ell(v(x, \theta)) - \alpha(u_h(x), v(x, \theta)) = R(u_h(x); v(x, \theta))
\]

where \( R(u_h(x); \cdot) \) is the residual associated with the numerical solution \( u_h(x) \).

2.4. Karhunen Loève decomposition

Recall that \( \mathbf{K}(x, \theta) \) is a random field, that is a function mapping each position vector \( x \) to a random variable, typically with all the same PDF and with cross correlation depending on the distance between the locations. Assuming that the spatial correlation is regular enough (essentially, it reduces to have square integrable probability distributions and symmetric correlation functions), the Karhunen Loève decomposition [21,16] allows representing a random field by a sum of independent scalar random variables multiplied by deterministic functions of \( x \), namely

\[
\mathbf{K}(x, \theta) = \mathbf{K}_0(x) + \sum_{i=1}^{N_N} \sqrt{\lambda_i} \mathbf{K}_i(x) \xi_i(\theta),
\]

where \( \mathbf{K}_0(x) \) and \( \lambda_i, i = 1, 2, \ldots, \infty \), are the eigenfunctions and eigenvalues of the covariance operator associated with the random field \( \mathbf{K}(x, \theta) \) and \( \mathbf{K}_0(x) \) is the mean value of \( \mathbf{K}(x, \theta) \). The infinite expansion is often fairly approximated by truncating the sum to a finite number of terms, \( N_N \), that is

\[
\mathbf{K}(x, \theta) \approx \mathbf{K}_0(x) + \sum_{i=1}^{N_N} \sqrt{\lambda_i} \mathbf{K}_i(x) \xi_i(\theta)
\]

The practical consequence of applying the (truncated) Karhunen Loève Decomposition is that the stochasticity of the system is simplified. In the original problem \( \mathbf{K}(x, \theta) \) is a random field, that is, in every point \( x \), \( \mathbf{K} \) is a random variable and all these random variables are correlated. Assuming that (12) holds, the stochasticity is reduced to \( N_N \) independent random variables \( \xi_i(\theta), i = 1, 2, \ldots, N_N \). Thus, generating a realization of \( \mathbf{K}(x, \theta) \) is reduced to generate the corresponding realizations of the \( N_N \) independent random variables \( \xi_i(\theta) \).

This also allows simplifying the methodology when it comes to generate the stiffness matrix corresponding to some realization. A direct consequence of (12) is that the stiffness matrix reads

\[
\mathbf{K}(\theta) \approx \mathbf{K}_0 + \sum_{i=1}^{N_N} \sqrt{\lambda_i} \mathbf{K}_i \xi_i(\theta)
\]

being \( \mathbf{K}_0 \) the (deterministic) stiffness matrix corresponding to the Hooke tensor distribution described by \( \mathbf{K}(x) \).

2.5. Monte Carlo simulation

As mentioned before, the Monte Carlo strategy is used in order to guarantee the non intrusive character of the stochastic simulation. A number of \( N_{MC} \) realizations of the independent random variables is generated. Each of these realizations is denoted as an occurrence \( \theta_k, k = 1, 2, \ldots, N_{MC} \), in the generic random space. Thus, the corresponding values of \( \xi_i(\theta_k) \) and, consequently, \( \mathbf{K}(\theta_k) \) (eventually also \( F(\theta_k) \) if the loads are assumed to be stochastic) are numerically generated.

The corresponding realization of system (5) corresponding to \( \theta_k \) is then to be solved in order to obtain \( U(\theta_k) \) and \( Q(\theta_k) \). Having assumed a large number of samples \( Q(\theta_k) \), \( k = 1, 2, \ldots, N_{MC} \), the mean value of \( Q(\theta) \), typically denoted by \( E(Q(\theta)) \) after the alternative naming of mathematical expectation, and the variance, \( \sigma^2(Q(\theta)) \) are readily approximated by

\[
E(Q(\theta)) \approx \frac{1}{N_{MC}} \sum_{k=1}^{N_{MC}} Q(\theta_k)
\]

and

\[
\sigma^2(Q(\theta)) \approx \frac{1}{N_{MC}} \sum_{k=1}^{N_{MC}} [Q(\theta_k) - E(Q(\theta))]^2
\]

Also the histogram characterizing the PDF of \( Q(\theta) \) and the corresponding percentiles are approximated after the Monte Carlo sampling. The percentile of \( p \) is defined as the value \( Q(\theta_k) \) such that the probability of the event \( Q(\theta) \leq Q(\theta_k) \) is equal to \( p \). Once the Monte Carlo sampling \( Q(\theta_k), k = 1, 2, \ldots, N_{MC} \) is available, each realization \( Q(\theta_k) \) is assumed to approximate the percentile \( Q(\theta_k) \) being \( p \) equal to the number of realizations \( Q(\theta_k) \) such that \( Q(\theta) \leq Q(\theta_k) \), divided by \( N_{MC} \).

3. Reduced basis strategy

The reduced basis method allows drastically reducing computational costs when solving a large number of problems with similar characteristics. Two contexts in which the reduced basis approach is particularly interesting are optimization (in order to reach the optimum any iterative method requires solving many problems with intermediate configurations) and stochastic modeling (as indicated above, Monte Carlo sampling requires solving many
instances of problem (5)). Note that the size of system (5) is \( N_{RF} \), which is assumed to be large. The idea of the reduced basis approach is to create a data base of solutions (corresponding to different realizations or instances of the problem to be solved). This collection of \( N_{RB} \) linearly independent solutions, \( \{ u_{h(1)}, u_{h(2)}, \ldots, u_{h(N_{RB})} \} \) described by the corresponding vectors of nodal values \( \{ U_1, U_2, \ldots, U_{N_{RB}} \} \), constitutes a basis of a subset of the discrete FE space,

\[
U_{RB} : \text{span}\{u_{h(1)}, u_{h(2)}, \ldots, u_{h(N_{RB})}\} \subset U,
\]

Then, the solution of a new instance of (5) is seek in \( U_{RB} \) instead of in \( U \). This requires solving a system of size \( N_{RB} \) instead of one of size \( N_{RF} \). It is assumed that an acceptable accuracy can be reached with a number of elements in the reduced basis much lower than the number of degrees of freedom in the FE mesh, that is \( N_{RB} \ll N_{RF} \). Here, the accuracy is associated with the truncation error introduced by the reduced basis approach with respect to the full FE solution.

The round off errors introduced by the numerical arithmetics are assumed to be negligible in front of truncation errors. Of course, this depends on the error prescribed by the user: in the limit case of an extremely demanding accuracy (small prescribed truncation error) \( N_{RB} \) would tend to \( N_{RB} \) because the full FE solution is required. Again, this does not hold if round off errors enter into the game: if the prescribed accuracy for the truncation error is of the same order as the round off error (or lower) none the FE or the RB solutions would provide a meaningful solution.

The reduced basis solution is therefore a linear combination of the elements of the reduced basis:

\[
U_{RB} = \sum_{i=1}^{N_{RB}} a_i U_i,
\]

where the matrix \( U_{RB} = \{ U_{1}, U_{2}, \ldots, U_{N_{RB}} \} \) (with \( N_{RB} \) rows and \( N_{RB} \) columns) describes the change of basis and \( a \) \( \{ a_1, a_2, \ldots, a_{N_{RB}} \} \) is the vector of unknowns.

As does not depend on the instance, it can be computed beforehand and be found as a contraction of (9),

\[
\left( \sum_{h=1}^{N_{RF}} K_\varepsilon(h) U_{RB} \right) a = \left( \sum_{h=1}^{N_{RF}} F(h) \right) \quad \text{readily rewritten as} \quad K_{RB}(\theta) a(\theta) = F_{RB}(\theta)
\]

by defining \( K_{RB}(\theta) := \sum_{h=1}^{N_{RF}} K_\varepsilon(h) U_{RB} \) and \( F_{RB}(\theta) := \sum_{h=1}^{N_{RF}} F(h) \). Then, once \( a \) is determined, the reduced basis solution is

\[
U_{RB}(\theta) = U_{RB} a(\theta).
\]

### 3.1. Goal oriented error assessment

The different simulations to be performed during the Monte Carlo process have the same structure. Typically, they correspond to the same problem statement with different materials or applied loads. The differences come from the variability associated with the randomness. In this context, the reduced basis methodology is expected to provide a good answer using (as reduced basis) a representative sample of snapshots, as it is shown in [6]. The novelty introduced in the current paper is to provide a criterion to select the basis, aiming at reducing the number of elements and improving their representativity. Thus, the reduced basis is not selected a priori using heuristic criteria but automatically generated along the process with an objective criterion, based on goal oriented error assessment. Hence, the resulting number of elements in the basis depends on the variability introduced by the random model: the larger is the dispersion, the larger is the number of required representative samples.

In other words, the reduced basis approach reduces drastically the computational cost because it requires solving system (18) (of size \( N_{RB} \)) instead of the complete FE system (5) (of size \( N_{RF} \gg N_{RB} \)). The question is whether the reduced basis solution \( U_{RB} \) is a sufficiently good approximation of the complete FE solution \( U \). This is equivalent to assess the error or to find out if the basis generating \( U_{RB} \) is representative enough. The remainder of this section is devoted to introduce a methodology allowing to answer this question. The basis is therefore enriched if the answer indicates that the error associated with the reduced basis approximation is too large.

At every instance \( \theta_k \) of the Monte Carlo process, the error introduced in the reduced basis phase is measured

\[
E_{RB} := U_{RB} - U,
\]

where, for the sake of simplifying the notation, the dependence on \( \theta_k \) is omitted in the expressions. Note that this error vector (expressed in the standard FE basis) corresponds to the approximation introduced by the reduced basis with respect to the complete FE solution. This is not assessing in any way the error introduced by the FE mesh (the difference between \( U \) and the exact solution \( u \)).

Multiplying Eq. (20) by matrix \( K \), an equation similar to (5) is recovered

\[
K E_{RB} = R_{RB} = K U_{RB} - F
\]

Note that the right hand side term is identified as the residual of (5) associated with the reduced basis solution. It follows from (21) that \( E_{RB} \) vanishes if \( R_{RB} \) is zero and, in general, the larger is the residual (which is computable), the larger is the error. The square norm of the residual, \( R_{RB}^T R_{RB} \), is in fact a good error indicator for the energy norm of the error.

The dual solution \( V \) introduced in (8) is used to assess the error in the evaluation of the QoI associated with the reduced basis. Note that the QoI associated with \( U_{RB} \) is readily computed as

\[
Q_{RB}(\theta) := Q_{RB}(\theta) = U_{RB}^T V_{RB} \quad \text{then, once} \quad V_{RB} \text{is determined, the QoI associated with the reduced basis is}
\]

\[
e_{RB}^2 := Q_{RB} - Q_{RB} G^T E_{RB} V_{RB}.
\]

Thus, the following expression holds, for the error in the QoI associated with the reduced basis approach:

\[
\epsilon_{RB}^2 := Q - Q_{RB} G^T E_{RB} V_{RB}.
\]

Note that if \( V \) is known, the error in the QoI associated with the reduced basis is computed explicitly using the right hand side term of (23), once the reduced basis solution is available. From the practical viewpoint, this has to be done for every instance of the Monte Carlo sampling. In the following it is assumed that the variation of \( V \) with the randomness (from one Monte Carlo shot to another) is small in such a way that \( V \) is kept constant in order to estimate the error in the QoI. Thus, the solution of the large dual system (8) is performed just once, for a value of \( \theta \) corresponding to the mean value of the random parameters. This allows assessing the error committed in every shot with a low computational cost and decide on the fly if the reduced basis is rich enough or if it has to be enriched further. The specific adaptive procedure using this idea is presented in detail in Section 3.2.

### 3.2. Adaptive algorithm for basis enrichment

As outlined above, the idea is to generate successively the reduced basis along the Monte Carlo simulation process, increasing the number of elements in the reduced basis when it is required. This is performed using a criterion based in the goal oriented error assessment given in the expression (23).

As a previous step, the dual problem is solved for the mean values of the random variables involved. Thus, the corresponding
solution of (8) yields \( V_A \), which is assumed to be the reference solution for the dual problem.

The Monte Carlo process starts as in the standard case, generating a realization of \( \theta \) denoted by \( \theta_1 \). This allows computing the realizations of the stiffness matrix a force vector, \( \mathbb{K}(\theta_1) \) and \( F(\theta_1) \). The corresponding system (5) is solved and the first solution \( U_1 \), \( U(\theta_1) \) is obtained. This solution is taken as the first reduced basis, at this phase \( U_{\text{ref}} = [U_1] \).

For the second Monte Carlo throw, \( \mathbb{K}(\theta_2) \) and \( F(\theta_2) \) are generated. Instead of considering the full system (5), the contracted system (18) is solved and \( U_{\text{ref}} \) is supposed to be an approximation to \( U(\theta_2) \). Note that size of system (18) is in this case, because the reduced basis has only one component. Thus, at this stage of the process, most probably, \( U_{\text{ref}} \) is not a good approximation of \( U(\theta_2) \).

The pertinence of the reduced basis approximation is checked with the proposed error assessment technique. First, the residual is computed as indicated in (21)

\[
R_{\text{ref}} = \mathbb{K}(\theta_2) U_{\text{ref}} - F(\theta_2)
\]

Then, the corresponding error in the QoI is estimated as

\[
\epsilon_{\text{ref}}^2 \approx V_1^T R_{\text{ref}}
\]

If the solution is admissible, i.e., the error in the QoI is small enough, then the Monte Carlo process is continued by generating the next realization. If it is not, then a new complete solution of (5), \( U_2 \), is computed and added to the reduced basis. Thus, if the admissibility criterion is not satisfied, then, at this stage \( U_{\text{ref}} = [U_1, U_2] \). This process is described compactly in Algorithm 1.

**Remark.** The reduced basis system of Eqs. (18) may suffer of potential ill conditioning if the elements in the basis are too similar. Notice that this occurs if one of the elements of the basis is nearly described as a linear combination of the others. The procedure introduced above is likely precluding this situation because a new element is introduced in the basis precisely only if it cannot be properly described as a linear combination of the existing ones. Nevertheless, if the prescribed tolerance, \( \epsilon_0 \) or \( \eta \), are too small (with respect to the roundoff error) the reduced basis matrix may, in practice, be ill conditioned. Moreover, the quality control is performed here only in terms of a specific QoI, and ill condition may appear because it results from having elements close to each other according to a different measure. Note that the standard approaches to prevent ill conditioning, alternatingly using Gram Schmidt or singular value decomposition, see [13], detect the relevant modes of the reduced basis and isolate the redundant information. Here, the strategy to incorporate elements in the reduced basis is likely providing only relevant elements introducing no redundant information in the basis.

### Algorithm 1: Adaptive algorithm for basis enrichment

**Data:** \( V^2 \), \( \epsilon_0 \) (or \( \eta \)), \( N_{\text{MC}} \) and the routines generating \( \mathbb{K}(\theta) \), \( F(\theta) \) and \( G(\theta) \).

**Result:** A number of \( N_{\text{MC}} \) realizations of the QoI, \( Q_i, i = 1, \ldots, N_{\text{MC}} \), with relative error lower than \( \eta \)

**First Monte Carlo throw**

1. generate \( \mathbb{K}(\theta_1), F(\theta_1), G(\theta_1) \);
2. compute \( U_1 \) solving \( \mathbb{K}(\theta_1) U_1 = F(\theta_1) \);
3. compute and store \( Q_1 = G(\theta_1)^T U_1 \);
4. initialize \( U_{\text{ref}} = [U_1] \) and \( N_{\text{ref}} = 1 \);

**Subsequent Monte Carlo throws**

for \( k = 2, \ldots, N_{\text{MC}} \) do

1. generate \( \mathbb{K}(\theta_k), F(\theta_k), G(\theta_k) \);
2. compute \( \mathbb{K}_{\text{ref}} := \int_\Omega^2 \mathbb{K}(\theta_k) U_{\text{ref}} \) and \( F_{\text{ref}} := \int_\Omega^2 F(\theta_k) U_{\text{ref}} \);
3. solving the reduced basis problem
4. compute a solution \( U_{\text{ref}} \)
5. compute \( U_k = U_{\text{ref}} \)
6. compute and store \( Q_k = G(\theta_k)^T U_k \);
7. assessing the error of the reduced basis solution
8. compute \( R_{\text{ref}} = \mathbb{K}(\theta_k) U_k - F(\theta_k) \);
9. compute the error estimate \( \epsilon_{\text{ref}}^2 \approx V_1^T R_{\text{ref}} \);
10. if \( \epsilon_{\text{ref}}^2 > \epsilon_0 \) or \( \epsilon_{\text{ref}}^2 > \eta \) then
11. solution not admissible: re-computing the full solution and enriching the basis
12. compute \( U_k \) solving \( \mathbb{K}(\theta_k) U_k = F(\theta_k) \);
13. compute and store \( Q_k = G(\theta_k)^T U_k \);
14. \( N_{\text{ref}} = N_{\text{ref}} + 1 \);
15. \( U_{\text{ref}} = [U_1, \ldots, U_{\text{ref}}] \)

end

### 4. Numerical examples

#### 4.1. Simple 1D example

A 1D test case is presented in this section. Both the geometry and the loads are simple, as well as the randomness behavior which is introduced by modeling the constant young modulus as a random variable. Despite of its simplicity, this example allows assessing the error introduced when the random behavior is approximated using a technique based on a response surface, see [19]. In this case, the reference case (obtained with a large number of Monte Carlo simulations) is assumed to be exact and, due to the simplicity of the problem, the Reduced Basis simulation with one degree of freedom is sufficient to capture the exact solution. The strategy to assess the error and the associated criterion to enrich the basis is, in this simple case, indicating that the one element basis is already furnishing an accurate enough solution. Thus, in this case, the Reduced Basis results (RB) are identical to the exact results which constitute the Reference and they are used to validate the outcome of the response surface approach.

This first example is also useful to introduce some of the notions and basic hypotheses used in the next section.

#### 4.1.1. Test description, analytical solution and RB solver

The unknown displacement \( u(x) \) describes the deformation of a bar of section \( S \) and length \( l \). (the domain \( \Omega = [0, l] \) is parametrized with the coordinate \( x \in \Omega \), as illustrated in Fig. 1. The bar is loaded with a deterministic traction \( F \) on \( x = l \) and clamped at \( x = 0 \).
material is assumed to be elastic with a non deterministic (but uniform) Young modulus \( E(\theta) \). The probability density function of \( E(\theta) \) is characterized as follows:

\[
E(\theta) = \mathbb{E}[\log(\xi(\theta))] \quad \text{with} \quad g(y) = \frac{2 \arcsin(\operatorname{Erf}(\frac{y}{\sqrt{2}}))}{\sqrt{\pi^2}} \frac{E_0}{8}
\]

where \( \xi(\theta) \) is a Gaussian centered random variable with standard deviation equal to one and \( \operatorname{Erf} \) stands for the error function. \( E_0 \) is the mean value of \( E(\theta) \) and \( \delta \) is the amplification coefficient for the random perturbation.

Note that the non standard nonlinear mapping \( g \) is chosen such that the probability density function of \( E(\theta) \), which is affine in \( g \), has a bounded support. Actually, the upper and lower limits for \( E \) are

\[
E_0 \left( 1 + \delta \sqrt{\frac{\pi^2}{\delta}} \right) \leq E(\theta) \leq E_0 \left( 1 + \delta \sqrt{\frac{\pi^2}{\delta}} \right).
\]

The definition allows excluding negative and physical values of the Young modulus. Positiveness is readily achieved taking \( E_0 > 1 \) and \( \delta > 0 \).

The equations to be solved read: find \( u(x, \theta) \) such that

\[
\frac{d}{dx} \left( u(x, \theta) E(\theta) \right) = 0 \quad \text{in} \quad x \in (0, L)
\]

\[
E(\theta) \frac{d}{dx} u(x, \theta) = F \quad \text{on} \quad x \in (0, L)
\]

The QoI is set to be the displacement at the right end, \( Q(\theta) = u(L, \theta) \).

An analytical expression for the exact solution of this very simple problem is available, namely

\[
Q(\theta) = \frac{F}{E(\theta) L}.
\]

Thus, a direct Monte Carlo simulation is computationally affordable because the QoI is obtained by simply replacing the value of the occurrence of \( E(\theta) \) in (28).

A full Monte Carlo simulation is performed using \( N_{\text{MC}} = 100,000 \) as a reference value. The corresponding occurrences for \( Q(\theta) \) are denoted \( Q_i(\theta) \) for \( i = 1, 2, \ldots, N_{\text{MC}} \). This allows computing a reference distribution for \( Q(\theta) \) and, in particular the mean, the standard deviation, the 25% percentile, the 50% percentile (median) and the 75% percentile, see Table 1.

4.2. Using a response surface based method

When applied to stochastic modeling, the response surface approach consists in interpolating the outcome of the model, here \( Q(\theta) \), from a number of different realizations of the random variable \( \theta \). Among the different possible interpolation techniques, here the simplest piecewise linear approximation is selected. Thus, the response surface approximation to \( Q(\theta) \), \( Q_{\text{RS}}(\theta) \), is computed by solving the full problem in a selected set of points \( \theta_j, j = 1, 2, \ldots, N_{\text{RS}} \) and computing the values of the occurrences \( Q(\theta_j) \).

The response surface solution a simple interpolation of the form

\[
Q_{\text{RS}}(\theta) = \sum_{j=1}^{N_{\text{RS}}} N_j(\theta) Q(\theta_j),
\]

being \( N_j(\theta) \) the classical linear finite element shape functions. This is the technique used in [19].

The parameter \( N_{\text{RS}} \) is tuning the cost of the method (number of complete solutions of the FE model) and how the method converges to the full Monte Carlo solution. Note that, once the response surface parametrization is set, each new occurrence requires just an explicit evaluation of (29). The results displayed in Table 1 show that the Response Surface Approach is converging to the exact value. Note that the results corresponding to \( N_{\text{RS}} = 27 \) are pretty accurate and required only 997 solutions of the full problem instead of 100,000.

Fig. 2 illustrates the convergence of the mean with \( N_{\text{MC}} \), different levels of refinement of the grid corresponding to \( N_{\text{RS}} \) are plotted.

In the next section, we will focus on the reduced basis method and the response surface method based on an exact solution, for 2D problems with Dirichlet boundary conditions, for a set of different structural data.

4.2. 2D elastic deformation, taking average vertical displacement as QoI

The deformation of the 2D elastic squared domain \( \Omega \) is described by the unknown displacement field \( u \). Displacements are set to be zero on the Dirichlet boundary, \( \partial D \Omega \), containing the lateral and the bottom sides. The deterministic distribution of tractions depicted in Fig. 3 is prescribed on the Neumann part of the boundary, \( \partial \Omega \), containing the top side. The traction density is non null on a centered strip \( \omega \) of length \( l_\omega \).

The elastic material is characterized by a stochastic Young modulus, described by the random field \( E(\theta) \). The mean value of \( E(x, \theta) \) is \( E_0 \), independent of \( x \), based on a arcsin description as in Eq. (25),

Table 1

<table>
<thead>
<tr>
<th>( N_{\text{RS}} )</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>25% percentile</th>
<th>Median</th>
<th>75% percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2.1791</td>
<td>1.7598</td>
<td>0.8778</td>
<td>0.9993</td>
<td>3.2767</td>
</tr>
<tr>
<td>7</td>
<td>1.149</td>
<td>0.6467</td>
<td>0.8214</td>
<td>0.9990</td>
<td>1.2522</td>
</tr>
<tr>
<td>13</td>
<td>1.1204</td>
<td>0.4277</td>
<td>0.8146</td>
<td>0.9988</td>
<td>1.3019</td>
</tr>
<tr>
<td>25</td>
<td>1.1145</td>
<td>0.4144</td>
<td>0.8118</td>
<td>0.9987</td>
<td>1.2953</td>
</tr>
<tr>
<td>49</td>
<td>1.1130</td>
<td>0.4125</td>
<td>0.8117</td>
<td>0.9987</td>
<td>1.2950</td>
</tr>
<tr>
<td>97</td>
<td>1.1127</td>
<td>0.4120</td>
<td>0.8116</td>
<td>0.9987</td>
<td>1.2949</td>
</tr>
<tr>
<td>Ref.</td>
<td>1.1125</td>
<td>0.4118</td>
<td>0.8116</td>
<td>0.9986</td>
<td>1.2948</td>
</tr>
</tbody>
</table>
The spatial correlation is characterized by an exponential behavior, such that

$$\text{cov}(E(x,\theta),E(x',\theta)) \rightarrow e^{-||x-x'||}$$

(31)

The quantity of interest considered here is the average vertical displacement on the strip $\omega$, that is

$$Q(\theta) = \frac{1}{E_g} \int_{\omega} u(x,\theta)dx$$

(32)

The discrete Karhunen–Loève expansion in (12) is used here to describe the spatial variability in terms of a finite number of $N_{KL}$ independent random modes. The random space is therefore $N_{KL}$ dimensional and a reduced order model is needed to preclude the so-called curse of dimensionality. In the current example, $N_{KL}$ is set to be equal to 20 after analyzing the evolution of the eigenvalues $\lambda_i$ in the expansion (12) corresponding to the spatial correlation introduced above. This evolution is graphically displayed in Fig. 5 where it can be observed that 20 modes are sufficient to capture the relevant features of the distribution. In a general case, a 20 dimensional space may seem not too large from the probabilistic viewpoint. However, it is sufficient for the current description of the spatial correlation and it is already interesting for reduced bases explorations. In the remainder, the error introduced by the Karhunen–Loève the truncation is considered to be negligible and the 20 dimensional space is taken as the reference problem.

Fig. 2. 1D example: Convergence of the mean with $N_{MC}$, for different values of $N_{MC}$.

Fig. 3. Illustration of the 2D problem statement.

Fig. 4. An occurrence of the random field $E(x,\theta)$.

Fig. 5. Values of the larger eigenvalues $\lambda_i$, $i = 1, \ldots, 200$, all the relevant amplitudes correspond to $i \leq 20$. 
Table 2
Parameters of the probability distribution of $Q$ (average vertical displacement) assessed using the reduced basis adaptive algorithm for different values of the prescribed error $e_0$ (upper rows) and reference value obtained solving $N_{MC}$ 100,000 different FE problems ($N_{MC}$ 944).

<table>
<thead>
<tr>
<th>$e_0$</th>
<th>$N_{RB}$</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>25 percentile</th>
<th>Median</th>
<th>75 percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1 \times 10^2$</td>
<td>1</td>
<td>3130.38563</td>
<td>130.18000</td>
<td>3030.48300</td>
<td>3124.85139</td>
<td>3225.34981</td>
</tr>
<tr>
<td>$1 \times 10^3$</td>
<td>9</td>
<td>3133.09896</td>
<td>127.46913</td>
<td>3035.00110</td>
<td>3127.75817</td>
<td>3226.99361</td>
</tr>
<tr>
<td>1</td>
<td>21</td>
<td>3133.72694</td>
<td>127.23239</td>
<td>3035.80352</td>
<td>3128.39911</td>
<td>3227.44973</td>
</tr>
<tr>
<td>$1 \times 10^4$</td>
<td>44</td>
<td>3133.72521</td>
<td>127.23357</td>
<td>3035.81726</td>
<td>3128.39412</td>
<td>3227.42919</td>
</tr>
<tr>
<td>$1 \times 10^2$</td>
<td>97</td>
<td>3133.72448</td>
<td>127.23713</td>
<td>3035.81932</td>
<td>3128.39521</td>
<td>3227.42800</td>
</tr>
<tr>
<td>$1 \times 10^3$</td>
<td>154</td>
<td>3133.72464</td>
<td>127.23715</td>
<td>3035.81903</td>
<td>3128.39334</td>
<td>3227.42781</td>
</tr>
<tr>
<td>$1 \times 10^4$</td>
<td>214</td>
<td>3133.72466</td>
<td>127.23718</td>
<td>3035.81884</td>
<td>3128.39341</td>
<td>3227.42794</td>
</tr>
</tbody>
</table>

$q_0$ is reduced (or as the prescribed accuracy is increased). Note that for all practical cases $N_{RB}$ is significantly lower than $N_{MC}$.

The error committed in the evaluation of the QoI for every occurrence is also analyzed. Having at hand the reference solution in each Monte Carlo simulation, that error $e_0$, associated with the RB computation of occurrence is controlled in the process. Actually, the correlation in Fig. 7 shows that the maximum error $e_{\text{max}}$ is, in practice and for a large number of Monte Carlo throws, equal to $e_0$.

Fig. 8 shows the dependence with the prescribed error $e_0$ of the different parameters of the distribution of the error $e_0$. It is clear that the choice of $e_0$ controls the error achieved in the different samplings, $e_k$.

Fig. 9 describes the evolution of the required number of elements in the reduced basis $N_{RB}$ as a function of the number of Monte Carlo throws, $N_{MC}$ for different values of $e_0$. As expected, $N_{RB}$ increases with $N_{MC}$ and gets stabilized after some threshold value, which depends on $e_0$.

4.3. 2D elastic deformation, taking average rotation as QoI

The same example as in Section 4.2 is now analyzed for a different quantity of interest. The new quantity of interest is the absolute value of the rotation of the strip $\omega$, defined by

$$Q(\omega) = \frac{1}{I_{G,\omega}} \int_{\omega} (\mathbf{z} \times \mathbf{x}_G) \cdot \mathbf{u}(\mathbf{x}, \theta) d\omega$$

where $G$ is the geometric center of section $\omega$, $I_{G,\omega}$ is the quadratic momentum along $x$, $G$ is the position of a point relative to $G$, and $\mathbf{V}$ is the vector normal to the plane.

Table 3 displays the parameters of the probability distribution of $e_0$ and maximum error $e_{\text{max}}$ (QoI average vertical displacement) for different values of the prescribed error $e_0$ and $N_{MC}$ 100,000.

<table>
<thead>
<tr>
<th>$e_0$</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>25 percentile</th>
<th>Median</th>
<th>75 percentile</th>
<th>$e_{\text{max}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1 \times 10^2$</td>
<td>3.3390361</td>
<td>2.928229</td>
<td>4.575123</td>
<td>9.619397</td>
<td>16.295682</td>
<td>57.4509362</td>
</tr>
<tr>
<td>$1 \times 10^3$</td>
<td>0.6248052</td>
<td>0.2319512</td>
<td>0.6907875</td>
<td>1.4663370</td>
<td>2.5131283</td>
<td>10.0730973</td>
</tr>
<tr>
<td>1</td>
<td>0.0000188</td>
<td>0.0000215</td>
<td>0.0000446</td>
<td>0.0000983</td>
<td>0.0001802</td>
<td>0.0011101</td>
</tr>
<tr>
<td>$1 \times 10^4$</td>
<td>0.0001806</td>
<td>0.0000427</td>
<td>0.0004581</td>
<td>0.0009901</td>
<td>0.0017716</td>
<td>0.0106919</td>
</tr>
</tbody>
</table>

The results of the modified algorithm are displayed in Tables 5 and 6 which compare to those of Tables 2 and 3. It is worth noting that for this test case, the results obtained using the approximate

4.4. Reference solution of the dual problem

The error assessment used above takes as the solution of the dual problem the occurrence $\theta_k$ the solution $V_0$ of the problem associated with the mean value of the random distribution, see Eq. (24).

Thus, it is assumed, for error assessment purposes, $V_0$ is a fair approximation of the actual solution of the dual problem. This approximation of the random variables corresponding to the current Monte Carlo throw, denoted as $V_k$ for $k = 1, 2, \ldots, N_{MC}$. The previous examples demonstrate that this working hypothesis is good enough, in the sense that the error in the quantity of interest, $e_0$, obtained with respect to the full FE solution, is effectively controlled in the process.

Here, an additional contrast of the hypothesis $V_0 \approx V_0^{\text{RB}}$, $V_0^{\text{RB}}$ is performed by solving the full dual problem at each Monte Carlo throw, that is, computing $V_k$ for each realization and therefore the actual error introduced by the hypothesis in the computation of the error in the QoI (associated with the RB strategy).
error assessment are very close to those obtained in this academic exercise consisting in computing the actual error at each Monte Carlo throw. There are not significant variations in the results replacing the approximated error assessment (proposed error estimate) by the actual error (comparing the RB solution with the full FE solution).

The possibility of solving the dual problem with an scheme that controls the quality of the dual solution is the object of current re
search. Note that, in order to increase the accuracy of the goal oriented error estimate, it suffices to increase the accuracy of the dual solution, that is finding a better approximation for $V_d$. A possible approach to address this question is using also a reduced basis strategy for the dual problem. It is not obvious whether using the same basis for both the primal and the dual problem or using different bases is a better option.

**Table 4**
Parameters of the probability distribution of $Q$ (average rotation) assessed using the reduced basis adaptive algorithm for different values of the prescribed error $\epsilon_g$ (upper rows) and reference value obtained solving $N_{ref}$ 100,000 different FE problems ($N_{ref}$ = 944).

<table>
<thead>
<tr>
<th>$\epsilon_g$</th>
<th>$N_{ref}$</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>25% percentile</th>
<th>Median</th>
<th>75% percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1 \times 10^{-3}$</td>
<td>9</td>
<td>22.828639</td>
<td>16.349056</td>
<td>9.484006</td>
<td>19.931768</td>
<td>33.288322</td>
</tr>
<tr>
<td>1</td>
<td>21</td>
<td>22.221976</td>
<td>15.970891</td>
<td>9.195373</td>
<td>19.376013</td>
<td>32.350854</td>
</tr>
<tr>
<td>$1 \times 10^{-1}$</td>
<td>45</td>
<td>22.214807</td>
<td>15.967786</td>
<td>9.214940</td>
<td>19.370062</td>
<td>32.336906</td>
</tr>
<tr>
<td>$1 \times 10^{-2}$</td>
<td>97</td>
<td>22.215679</td>
<td>15.970235</td>
<td>9.215144</td>
<td>19.368300</td>
<td>32.337244</td>
</tr>
<tr>
<td>$1 \times 10^{-3}$</td>
<td>147</td>
<td>22.215924</td>
<td>15.970409</td>
<td>9.21495</td>
<td>19.367796</td>
<td>32.337462</td>
</tr>
<tr>
<td>$1 \times 10^{-4}$</td>
<td>241</td>
<td>22.215934</td>
<td>15.970418</td>
<td>9.214382</td>
<td>19.367640</td>
<td>32.337302</td>
</tr>
<tr>
<td>Ref</td>
<td>-</td>
<td>22.215933</td>
<td>15.970418</td>
<td>9.214383</td>
<td>19.367634</td>
<td>32.337288</td>
</tr>
</tbody>
</table>
5. Conclusions

This paper describes a new approach to generate a reduced basis in the context of Monte Carlo strategies for stochastic modeling. The reduced basis is constructed automatically, ensuring a prescribed level of accuracy for the output of interest. The proposed methodology uses standard elements in goal oriented error assessment and adaptivity. Here, the error which is assessed is the approximation introduced by the reduced basis with respect to the complete FE solution, for a given mesh. That differs from the usual practice in the verification framework, in which the error introduced by the mesh is evaluated. Numerical 1D and 2D tests demonstrate the efficiency and robustness of the proposed strategy. The use of this methodfor 3D massive industrial examples, where the cost is of primary importance, is going to be the object of further research. This type of problems with large stochastic variations where the reduced basis can significantly increase is also an interesting point to be studied.

References


