Comput. Methods Appl. Mech. Engrg. 198 (2009) 3187–3206

Contents lists available at ScienceDirect

Comput. Methods Appl. Mech. Engrg.

journal homepage: www.elsevier.com/locate/cma

A reduced basis approach for variational problems with stochastic parameters: Application to heat conduction with variable Robin coefficient

Sébastien Boyaval [a,b,*], Claude Le Bris [a,b], Yvon Maday [c,d], Ngoc Cuong Nguyen [e], Anthony T. Patera [e]

[a] CERMICS, Ecole Nationale des Ponts et Chaussées, (ParisTech/Université Paris-Est), 6 and 8 Avenue Blaise Pascal, Cité Descartes, 77455 Marne-la-Vallée, Cedex 2, France
[b] MICMAC Team-Project, INRIA, Domaine de Voluceau, BP. 105-Rocquencourt, 78153 Le Chesnay, Cedex, France
[c] UPMC Univ Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France
[d] Division of Applied Mathematics, Brown University Providence, RI 02912, USA
[e] Massachusetts Institute of Technology, Department of Mechanical Engineering, Cambridge, MA 02139, USA

ARTICLE INFO

Article history:
Received 14 August 2008
Received in revised form 31 March 2009
Accepted 3 May 2009
Available online 14 June 2009

Keywords:
Stochastic parameterized partial differential equations
Karhunen–Loève
Monte Carlo
Reduced basis method
A posteriori error estimation

ABSTRACT

In this work, a Reduced Basis (RB) approach is used to solve a large number of boundary value problems parametrized by a stochastic input – expressed as a Karhunen–Loève expansion – in order to compute outputs that are smooth functionals of the random solution fields. The RB method proposed here for variational problems parametrized by stochastic coefficients bears many similarities to the RB approach developed previously for deterministic systems. However, the stochastic framework requires the development of new *a posteriori* estimates for "statistical" outputs – such as the first two moments of integrals of the random solution fields; these error bounds, in turn, permit efficient sampling of the input stochastic parameters and fast reliable computation of the outputs in particular in the many-query context.

1. Introduction

1.1. Overview

Let $U(\cdot,\omega)$ be a scalar random field solution to a (presumed well-posed) Boundary Value Problem (BVP) involving a Stochastic Partial Differential Equation (SPDE). For instance, if variations in the probability space $(\Omega,\mathscr{F},\mathbf{P})$ are denoted by the variable ω, we take $U(\cdot,\omega)$ as the **P**-almost sure (a.s.) solution to the Partial Differential Equation (PDE) in a (smooth) physical domain $\mathscr{D}$

$$-\mathrm{div}(\mathbf{a}(x)\nabla U(x,\omega)) = 0, \quad \forall x \in \mathscr{D} \tag{1}$$

supplied with a stochastic Robin Boundary Condition (BC) on the boundary $\partial\mathscr{D}$ parametrized by a random input field $\mathrm{Bi}(\cdot,\omega)$

$$\mathbf{n}(x)^{\mathrm{T}}\mathbf{a}(x)\nabla U(x,\omega) + \mathrm{Bi}(x,\omega)\ U(x,\omega) = g(x), \quad \forall x \in \partial\mathscr{D}. \tag{2}$$

Here, **a** takes symmetric positive definite matrix values, the random field $\mathrm{Bi}(\cdot,\omega)$ (Biot number [22]) is non-zero (non-degenerate positive) on some subset $\Gamma_{\mathrm{B}} \subset \partial\mathscr{D}$ (with non-zero measure), $\mathbf{n}(x)$ is the outward unit normal at $x \in \partial\mathscr{D}$ and T denotes the transpose.

We consider the rapid and reliable computation of statistical outputs associated with $U(\cdot,\omega)$ such as the expected value $\mathbf{E}_{\mathbf{P}}(S)$ and the variance $\mathbf{Var}_{\mathbf{P}}(S)$ of a random variable $S(\omega) = \mathscr{E}(U(\cdot,\omega))$ given by a linear (scalar) functional $\mathscr{E}$ of the trace of $U(\cdot,\omega)$ on $\Gamma_{\mathrm{R}} \subset \partial\mathscr{D}$ (where $\Gamma_{\mathrm{R}} \cap \Gamma_{\mathrm{B}} = \emptyset$)

$$\mathscr{E}(U(\cdot,\omega)) = \int_{\Gamma_{\mathrm{R}}} U(\cdot,\omega). \tag{3}$$

One possible strategy is to evaluate the statistical outputs as Monte-Carlo (MC) sums of the random variable S

$$E_M[S] = \frac{1}{M}\sum_{m=1}^{M} S^m, \quad V_M[S] = \frac{1}{M-1}\sum_{m=1}^{M}\left(E_M[S] - S^m\right)^2, \tag{4}$$

using M independent random variables $(S^m)_{1\leqslant m\leqslant M}$ with the same distribution law as S. But M can be very large, and hence these MC evaluations can be very demanding (for each m, one must solve a BVP PDE in $\mathscr{D}$). Furthermore, in actual practice, and as developed subsequently in this paper, we are often interested in evaluating our statistical outputs for different values of deterministic parameters, say ϱ – which even further increases the computational challenge. For this reason we develop a Reduced Basis (RB) approach: to decrease the computational cost of the M realizations of the Finite

* Corresponding author. Address: CERMICS, Ecole Nationale des Ponts et Chaussées, (ParisTech/Université Paris-Est), 6 and 8 Avenue Blaise Pascal, Cité Descartes, 77455 Marne-la-Vallée, Cedex 2, France. Tel.: +33 1 64 15 35 79; fax: +33 1 64 15 35 86.
E-mail addresses: boyaval@cermics.enpc.fr, sboyaval@mit.edu (S. Boyaval).

0045-7825/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cma.2009.05.019

Element (FE) approximations $U_{\mathcal{N}}(\cdot,\omega) \approx U(\cdot,\omega)$ required in the Monte-Carlo sums.

Toward this goal, we first rewrite the parametrization of the BVP using a Karhunen–Loève (KL) expansion of the random input field (see Section 2 for details)

$$\mathrm{Bi}(x,\omega) = \overline{\mathrm{Bi}}\left(G(x) + \sum_{k=1}^{\mathscr{K}} \Phi_k(x)\, Y_k(\omega)\right), \quad \forall x \in \partial\mathscr{D}, \tag{5}$$

where $\mathscr{K}$ is the rank (possibly infinite) of the covariance operator for $\mathrm{Bi}(\cdot,\omega)$ with eigenvectors $(\Phi_k)_{1\leqslant k\leqslant\mathscr{K}}$, the positive number $\overline{\mathrm{Bi}} = \int_\Omega d\mathbf{P}(\omega) \int_{\partial\mathscr{D}} \mathrm{Bi}(\cdot,\omega)$ is an intensity factor and the random variables $(Y_k)_{1\leqslant k\leqslant\mathscr{K}}$ are mutually uncorrelated in $L^2_{\mathbf{P}}(\Omega)$ with zero mean. Next, we define a function $\mathrm{bi}(\cdot;\overline{\mathrm{Bi}},y)$ parametrized by $\overline{\mathrm{Bi}} \in \mathbb{R}_{>0}$ and the (possibly infinite) sequence $y = (y_1, y_2, \ldots) \in \Lambda^y \subset \mathbb{R}^{\mathscr{K}}$

$$\mathrm{bi}(x;\overline{\mathrm{Bi}},y) = \overline{\mathrm{Bi}}\left(G(x) + \sum_{k=1}^{\mathscr{K}} \Phi_k(x) y_k\right), \ \forall x \in \partial\mathscr{D}, \tag{6}$$

such that for all $\overline{\mathrm{Bi}} \in \mathbb{R}_{>0}$ and $y \in \Lambda^y$ the parametrized function $\mathrm{bi}(\cdot;\overline{\mathrm{Bi}},y)$ is well defined; we also define truncated versions $y^K = (y_1, y_2, \ldots, y_K, 0, 0, \ldots) \in \Lambda^y$ up to order $K \leqslant \mathscr{K}$ of the deterministic parameter sequence y.

For any positive integer $K \leqslant \mathscr{K}$, we then define a solution $U_K(\cdot,\omega)$ to the BVP in which the KL expansion of $\mathrm{Bi}(\cdot,\omega)$ in the Robin BCs is replaced by a truncated version at order K

$$\mathrm{Bi}_K(\cdot,\omega) = \mathrm{bi}(\cdot;\overline{\mathrm{Bi}}, Y^K(\omega)),$$

using truncated versions Y^K (with $K \leqslant \mathscr{K}$) of the (possibly infinite) sequence $Y = (Y_k)_{1\leqslant k\leqslant\mathscr{K}}$ of random variables. For almost all fixed $x \in \mathscr{D}$, the random variable $U_K(x,\cdot)$ is clearly $\sigma(Y^K)$-measurable and, by the Doob–Dynkin lemma [37], we have $U_K(x,\omega) = u_K(x; Y^K(\omega))$ for almost all $(x,\omega) \in \mathscr{D}\times\Omega$, where $u_K(\cdot;y^K)$ solves a y^K-parametrized BVP PDE problem ($y^K \in \Lambda^y$):

$$\begin{cases} -\mathrm{div}(\mathbf{a}(x)\nabla u_K(x;y^K)) = 0, & \forall x \in \mathscr{D}, \\ \mathbf{n}(x)^{\mathrm{T}}\mathbf{a}(x)\nabla u_K(x;y^K) + \mathrm{bi}(x;\overline{\mathrm{Bi}},y^K)u_K(x;y^K) = g(x), & \forall x \in \partial\mathscr{D}. \end{cases} \tag{7}$$

The problem (7) is well-posed under standard hypotheses for all $y^K \in \Lambda^y$ in the range of Y.

The statistical outputs for $S_K(\omega) = \mathscr{E}(U_K(\cdot,\omega))$ obtained after truncation of the KL expansion

$$E_M[S_K] = \frac{1}{M}\sum_{m=1}^{M} S_K^m, \quad V_M[S_K] = \frac{1}{M-1}\sum_{m=1}^{M}\left(E_M[S_K] - S_K^m\right)^2, \tag{8}$$

can then be obtained as, respectively, $E_M[s_K(Y^K)]$ and $V_M[s_K(Y^K)]$, using $s_K(y^K) = \mathscr{E}(u_K(\cdot;y^K))$ and M independent random vectors $(Y_m^K)_{1\leqslant m\leqslant M}$ with the same distribution law as Y^K. Clearly, the error in these outputs due to truncation of the KL expansion must be assessed; we discuss this issue further below. (We must also ensure that M is large enough; we address this question in the context of our numerical results.)

In Section 3, we develop a Reduced Basis (RB) approach [1,12,34,41] for the parametrized (deterministic) BVP (7) and outputs (8) for the case in which the random variables $Y_k, 1 \leqslant k \leqslant K(\leqslant \mathscr{K})$, are bounded (uniformly if $\mathscr{K} = +\infty$) such that the KL expansion is positive for any truncation order K (and converges absolutely a.e. in $\partial\mathscr{D}$ when $\mathscr{K} = +\infty$); the latter ensures well-posedness of the BVPs obtained after truncation at any order $1 \leqslant K \leqslant \mathscr{K}$. We shall present numerical results for a random input field $\mathrm{Bi}(\cdot,\omega)$ whose spatial autocovariance function is a Gaussian kernel such that the KL spectrum decays rapidly.

In particular, we shall show that our RB approach significantly reduces the computational cost of the MC evaluations with no sensible loss of accuracy compared to a direct Finite Element (FE) approach. For instance, with truncated KL expansions of order $K \leqslant 20$, the RB computational time for solutions to (7) is reduced by a factor of $\frac{1}{45}$ relative to direct FE, and the (relative) approximation error in the expectation due to both RB and KL truncation is controlled *and* certified to 0.1% (for $K = 20$). Our RB approach thus also straightforwardly permits rapid exploration of the dependence of the outputs $E_M[s_K(Y^K)]$ and $V_M[s_K(Y^K)]$ on variations in additional *deterministic* parameters ϱ entering the problem. (In the limit of many evaluations at different ϱ, computational savings relative to FE can be as much as $O(200)$).

1.2. Relation to prior work

The computation of BVPs involving SPDEs has been identified as a demanding task [3,9,10,29] for several years, whatever the numerical approach used to discretize the SPDE. For instance, among those numerous numerical approaches, the popular *spectral (stochastic) Galerkin* discretizations [16], based on a (generalized) Polynomial Chaos (PC) expansion of the solution [55,56], consists in solving a variational problem in a *high-dimensional* tensor–product functional space on $\mathscr{D}\times\Lambda^y$, which is computationally (very) expensive. Hence several reduction techniques have been proposed recently for the spectral Galerkin approach, in particular:

- sparse/adaptive methods [13,50],
- efficient iterative algorithms for the fully discretized problems, using parallel solvers, preconditioners and/or Krylov projections [40,25], sometimes termed "stochastic RB" Krylov methods [30,44,49],
- POD approaches for PC discretizations of the functions in the stochastic variable (combined with a two-scale approach in the physical space) [11,15],
- POD approaches for PC-FE discretizations of the functions defined on the whole tensor–product space, termed "generalized spectral decomposition" [35,36],
- and stochastic collocation approaches [57,2,33].

These reduction techniques have shown good performance on test cases. However, the sparse/adaptive methods require substantial implementation efforts, the Krylov methods and the POD approaches do not yet provide rigorous *a posteriori* analysis to control the output approximation error, and the stochastic collocation method still invokes numerous (expensive) FE solutions – at each collocation point. The RB method described here – albeit for a limited class of problems – focuses on simple implementation, rigorous *a posteriori* error bounds, and parsimonious appeal to the FE "truth".

The formulation of the RB method presented herein can be straightforwardly applied to discretizations of the SPDE that lead to the solution of many *decoupled* variational formulations of the BVP on $\mathscr{D}$ for many fixed given values of the random input in Λ^y (like (7)). In the present work, the RB method is only applied to Monte-Carlo/Galerkin (in fact, Finite-Element) discretizations of the SPDE, as described earlier in this introduction. That is, the statistical outputs like mean and variance of some functional of the random variable solution to the SPDE are computed through Monte-Carlo (MC) evaluations of the random variable $S_K = s_K(Y^K)$, and not through quadrature or collocation formulæ for the (weighted) integration of the function $y^K \to s_K(y^K)$ over $y^K \in \Lambda^y$.

However, the RB method could be applied as well to many numerical approaches where integration in the stochastic space is discretized by collocation at many points in the range of the random input, where at each of these points one has to solve a PDE parametrized *only* by the value of the random input at the same point. In particular, the RB method proposed in this paper can be viewed as an *accelerator* of the stochastic collocation approach de-

scribed in [2], where a basis of orthogonal polynomials in the stochastic variables is substituted for the standard PC basis. As a matter of fact, the stochastic collocation approach is just a *pseudo-spectral* Galerkin discretization: it applies quadrature formulæ for the computation of the outputs $\mathbf{E_P}\left(s_K(Y^K)\right)$ and $\mathbf{Var_P}\left(s_K(Y^K)\right)$ so as to split the variational formulation for $u_K(\cdot;\cdot)$ on the high-dimensional tensor–product space $(x,y^K)\in\mathscr{D}\times\Lambda^y$ into many variational formulations on the lower-dimensional space $\mathscr{D}$ parametrized by $y^K\in\Lambda^y$. Clearly, we may replace s_K by a (certified) RB approximation to further reduce the computational effort[1]; equivalently, we may replace the MC sums of our current approach with the quadrature rules developed in [2,33]. Future work will investigate this promising opportunity.

Compared with numerical approaches developed previously for SPDEs, the main features of our RB approach are the following:

(a) the solution $U_K(\cdot,\omega)$ to the original *stochastic* BVP is mapped to the distribution of Y^K,

$$U_K(x,\omega)=u_K(x;Y^K(\omega)) \quad \text{for almost every (a.e.)} \quad x\in\mathscr{D} \quad \text{and} \quad \mathbf{P}-\text{a.e. outcome} \quad \omega\in\Omega,$$

through the solution $u_K(\cdot;y^K)$ to a *deterministic* BVP, the variational formulation of which must have an *affine* parametrization[2] (*affine* in the sense that the weak form can be expressed as a sum of products of parameter-dependent functions and parameter-independent forms) – as typically provided by a KL expansion of the random input field which decouples the dependencies on the probability and physical spaces;

(b) a large number of variational approximations for the solutions $u_K(\cdot;y^K)$ to the *deterministic* BVP, defined over the (relatively) low-dimensional physical space $\mathscr{D}$ and parametrized by y^K, must be computed for each MC evaluation of the statistical outputs (and for each value of the additional parameter ϱ) – as opposed to spectral Galerkin variational methods in which $u_K(\cdot;\cdot)$ is discretized on the high-dimensional tensor–product space $(x,y^K)\in\mathscr{D}\times\Lambda^y$ such that only one, very expensive, solution is required (for each value of the additional parameter ϱ);

(c) the "deterministic" RB approach [27,42,43] is then applied to the deterministic BVP to yield – based on a many-query Offline–Online computational strategy – greatly reduced computational cost at little loss in accuracy or, thanks to rigorous *a posteriori* bounds, certainty.

Of course our approach also bears many similarities to earlier proposals, most notably reliance on the Kolmogorov strong law of large numbers (for the MC evaluations to converge), on the KL expansion of the random input field, and on smoothness with respect to the parameter y^K.

Note that the usual RB method can be extended to the SPDE framework thanks to new error bounds (to take into account the effect of the truncation of the KL expansion, and to assess the efficiency of the reduction, that is to control the RB error in outputs that are sums over many parameter realizations). But the idea behind the RB method remains the same as in the usual case of parametrized (deterministic) PDEs, even though SPDEs typically result in *many* ($>K$) deterministic parameters (y^K,ϱ). The rapid convergence of the RB method we observe here – that does not break but at least moderates the curse of dimensionality – relies heavily not only on the smoothness of $u_K(\cdot;y^K)$ with respect to y^K, but also on the limited range of the y_k component of y^K when $k\gg 1$; the latter, in turn, derives from the (assumed) smoothness of the autocovariance function (rapid decay of the eigenvalues of the Hilbert–Schmidt integral operator with the autocovariance function as kernel). It is imperative to choose K as small as possible.

2. Variational formulation of a boundary value problem with stochastic parameters

2.1. Stochastic partial differential equations

The modeling of multiscale problems in science and engineering is often cast into the following framework. At the macroscopic scale at which important quantities must be computed, a (possibly multi-dimensional) field variable $U(\cdot,\omega)$ is assumed to satisfy a PDE on a physical domain $\mathscr{D}\subset\mathbb{R}^d$ ($d=2,3$, or 4 for common applications)

$$A(\cdot,\omega)\,U(\cdot,\omega)=f(\cdot,\omega) \quad \text{in } \mathscr{D}, \tag{9}$$

supplied with Boundary Conditions (BC) on the (sufficiently smooth) boundary $\partial\mathscr{D}$,

$$B(\cdot,\omega)\,U(\cdot,\omega)=g(\cdot,\omega) \quad \text{in } \partial\mathscr{D}; \tag{10}$$

here the differential operators $A(\cdot,\omega),B(\cdot,\omega)$ and the source terms $f(\cdot,\omega),g(\cdot,\omega)$ are parametrized at each point of the physical domain by a variable ω describing the state of some generalized local microstructure. We shall not discuss other possible formulations for multiscale problems, such as integral equations; furthermore, the formulation above will be assumed well-posed in the sense of Hadamard for the case in which A,B,f and g vary with the microstructure ω (extensions of this work to distributions, that is, generalized functions of ω, are not straightforward).

To model the "fluctuations" of the underlying microstructure, whose impact on the macroscopic quantities of interest is to be evaluated, we can assume – without invoking detailed information about the microstructure – that the input is random. To this aim, one can introduce an abstract probability space to model the fluctuations, the latter being then described through variations within the set of elementary events $\omega\in\Omega$ (similar arguments are often developed to model material properties[3], see *e.g.* [38,58]). The outputs of such models are then also random by nature. The Eqs. (9) and (10) are then generally called Stochastic PDEs (SPDEs). SPDEs are useful when one cannot, or does not want to, describe precisely the microstructure. Examples include uncertainty quantification for structures in civil engineering [8,48], for complex flows in fluid dynamics [28], or for multiphase flows in porous media [14].

2.2. Problem statement: stochastic robin boundary condition

The RB method has been introduced earlier for the many-query evaluation of outputs for various parametrized variational problems [27,42,43] in a deterministic framework (deterministic PDE and BC). In this work, we shall choose only one (simple) example to illustrate the stochastic case; however, it should be clear that

[1] In [33], it is even shown that one can minimize the number of collocation points, which correspond to zeros of the family of orthogonal polynomials substituted for the PC basis, with a view to "optimally" describing the range of the random input.

[2] Non-affine (but piecewise smooth) parametrizations can also be treated by the so-called *magic points* to "empirically" interpolate the coefficients entering the variational formulation [4,17].

[3] We note that by choosing the microscopic fluctuations as stationary ergodic random fields, the numerical treatment of averaged outputs for SPDEs also applies to many situations considered in stochastic homogenization theory [5,23], in which a powerful and elegant analysis of (weak) convergence allows one to reduce the modeling of complex multiscale problems to a more tractable set of sub-problems. Note that the RB approach has been applied to efficient numerical treatment of multiscale problems with locally periodic fluctuations within the context of deterministic homogenization theory [6].

the approach admits a general abstraction applicable to a wide class of problems.[4] We now pose our particular problem.

We shall let $(\Omega, \mathscr{F}, \mathbf{P})$ be a complete probability space where Ω is the set of outcomes ω, $\mathscr{F}$ is the σ-algebra of events among all subsets of Ω, and $\mathbf{P}$ is a probability measure (notice that this definition itself is often a practical issue for the modeller). And we shall let the physical domain $\mathscr{D}$ be an open, bounded, connected subset of $\mathbb{R}^2$ ($d = 2$) with Lipschitz polyhedral boundary, which we classically equip with the usual Borel σ-algebra and the Lebesgue measure. We recall that random fields are collections of scalar random variables that can be mapped to a physical domain; for instance, functions are defined on $\partial\mathscr{D}$ and take values in $L^2_{\mathbf{P}}(\Omega)$ – the space of square-integrable functions on the probability space $(\Omega, \mathscr{F}, \mathbf{P})$.

Let us introduce some further notations:

- $L^2(\mathscr{D})$ the Hilbert space of Lebesgue square integrable functions in $\mathscr{D}$;
- $H^1(\mathscr{D})$ the usual Sobolev space (with Hilbert structure) of functions in $L^2(\mathscr{D})$ that have gradient in $[L^2(\mathscr{D})]^2$, imbued with the usual Hilbert norm $\|\cdot\|_{1,\mathscr{D}}$;
- $L^2(\partial\mathscr{D})$ the Hilbert space of the Lebesgue square integrable functions in the manifold $\partial\mathscr{D}$ equipped with its Borel σ-algebra, imbued with the Hilbert norm $\|\cdot\|_{0,\partial\mathscr{D}}$;
- $L^\infty(\partial\mathscr{D})$ the Banach space of essentially bounded functions on the manifold $\partial\mathscr{D}$, imbued with its usual norm $\|\cdot\|_{\infty,\partial\mathscr{D}}$.

We also recall that functions $v \in H^1(\mathscr{D})$ have a trace $v \in L^2(\partial\mathscr{D})$ on $\partial\mathscr{D}$ that satisfies

$$\|v\|_{0,\partial\mathscr{D}} \leqslant \tilde{\gamma}\|v\|_{1,\mathscr{D}}, \tag{11}$$

where $\tilde{\gamma} \equiv \tilde{\gamma}(\mathscr{D})$ is a constant positive real number that depends only on $\mathscr{D}$.

In the following, we shall deal with SPDEs in which only the boundary differential operator $B(\omega)$ is parametrized by a random scalar input field, in particular $\mathrm{Bi}(\cdot,\cdot)$: $\partial\mathscr{D} \times \Omega \to \mathbb{R}$. We identify in (9) and (10)

$$A(x,\omega) = -\mathrm{div}(\mathbf{a}(x)\nabla\cdot), \quad f(x,\omega) = 0, \quad \forall x \in \mathscr{D},$$
$$B(x,\omega) = \mathbf{n}^{\mathrm{T}}(x)\ \mathbf{a}(x)\ \nabla\cdot + \mathrm{Bi}(x,\omega)\cdot, \quad g(x,\omega) = g(x), \quad \forall x \in \partial\mathscr{D}.$$

The case in which the other terms also depend on a single scalar random field $\mathrm{Bi}(\cdot,\omega)$ is a straightforward extension, provided the problem (9) and (10) remains well-posed in the sense of Hadamard with respect to the variations $\omega \in \Omega$. Note that the divergence div and gradient ∇ operators imply differentiations with respect to the physical variable x only, and not with respect to the probability variable ω. The scalar random field $U(\cdot,\omega)$ with $x \in \mathscr{D}$ is defined as the **P**-a.s. solution to the Robin BVP (1) and (2).

The deterministic (strictly positive) diffusion matrix $\mathbf{a}$ is assumed isotropic though non-constant for all $x \in \mathscr{D}$ (the function κ is specified below to get a simple "additional" deterministic parameter ϱ),

$$\mathbf{a}(x) = \begin{bmatrix} \kappa(x) & 0 \\ 0 & \kappa(x) \end{bmatrix}, \quad \forall x \in \mathscr{D}.$$

We shall assume $0 < \kappa_{\min} \leqslant \kappa(x) \leqslant \kappa_{\max} < +\infty$ for well-posedness. The boundary $\partial\mathscr{D}$ is divided into three non-overlapping open subsets

$$\partial\mathscr{D} \subset \left(\overline{\Gamma_{\mathrm{N}}} \cup \overline{\Gamma_{\mathrm{R}}} \cup \overline{\Gamma_{\mathrm{B}}}\right).$$

[4] We shall limit attention to those simple SPDEs which are not generalizations of Stochastic Differential Equations (SDEs) to multi-dimensional derivatives – where outcomes of the random input are distributions (generalized functions). Such interesting cases will be the subject of future work.

The boundary (Root) source term g is taken as deterministic (constant), non-zero on Γ_{R} only,

$$g(x) = 1_{\Gamma_{\mathrm{R}}}, \quad \forall x \in \partial\mathscr{D},$$

while the *Biot number* Bi is taken as a positive random field, non-degenerate on Γ_{B} only,

$$\mathrm{Bi}(x,\omega) = \mathrm{Bi}(x,\omega)\ 1_{\Gamma_{\mathrm{B}}}, \quad \forall x \in \partial\mathscr{D}.$$

Note that on Γ_{N}, Eq. (2) thus reduces to homogeneous Neumann conditions.

The physical interpretation is simple: if T_0 is the constant temperature of the ambient medium, $T_0 + U$ is the steady-state temperature field in a domain $\mathscr{D}$ (comprised of an isotropic material of thermal conductivity κ) subject to zero heat flux on boundary Γ_{N} (either by contact with a thermal insulator or for reasons of symmetry), constant flux at boundary Γ_{R} (contact with a heat source), and a random heat transfer coefficient Bi at boundary Γ_{B} (contact with a convective fluid medium). Note that the Biot number Bi is a fashion for decoupling the solid conduction problem from the exterior fluid convection problem: it is at best an engineering approximation, and at worst a rough average – often not reflecting the environmental details; it thus makes sense to model the unknown Bi variations as a random (but typically rather smooth) field $\mathrm{Bi}(\cdot,\omega)$ in order to understand the sensitivity of output quantities to heat transfer coefficient uncertainties.

With a view to specify parameters which will then be used in the numerical application of Section 3, we shall more precisely consider the steady heat conduction problem (1) and (2) inside the T-shaped heat sink $\mathscr{D}$ as shown in Fig. 1. The heat sink comprises a 2×1 rectangular substrate (spreader) $\mathscr{D}_2 \equiv (-1,1) \times (0,1)$ on top of which is situated a 0.25×4 thermal fin $\mathscr{D}_1 \equiv (-0.25, 0.25) \times (1,5)$. (In effect, all lengths will be nondimensionalized relative to the side-length of the substrate.) We also specify the diffusion coefficient, which we shall take as a (normalized) piecewise constant

$$\kappa(x) = 1_{\mathscr{D}_1} + \kappa 1_{\mathscr{D}_2}, \quad \forall x \in \mathscr{D},$$

where $1_{\mathscr{D}_i}$ is the characteristic function of domain $\mathscr{D}_i$ ($i = 1, 2$). On Γ_{B}, the two sides of the fin, we shall impose a *stochastic* convection/Robin BC with a non-zero random Biot number Bi (built as a random field $\mathrm{Bi}(\cdot,\omega)$ with *a priori* known mean and autocovariance function, see Section 2.4.1); on Γ_{R}, the root, we impose unit flux $g(x) = 1$; and on Γ_N, we impose zero flux.

We recall that the outputs of interest will be the first two moments of a (scalar) linear functional $\mathscr{E}$ of the random solution field $U(\cdot,\omega)$ defined in (3) as the (random) integrated trace $S(\omega) = \mathscr{E}(U(\cdot,\omega))$ on the edge Γ_{R} of the domain $\mathscr{D}$ (corresponding

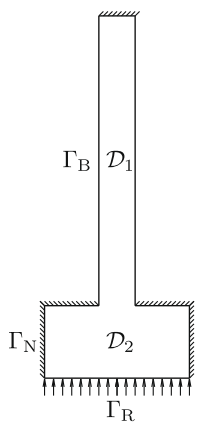


Fig. 1. Geometry of the heat sink: a spreader $\mathscr{D}_2$ and a fin $\mathscr{D}_1$.

to the location of the heat source – the point at which we wish to control the temperature):

$$\mathbf{E}_{\mathbf{P}}(S) := \int_{\Omega} S(\omega)\ d\mathbf{P}(\omega), \quad \mathbf{Var}_{\mathbf{P}}(S) : = \int_{\Omega} S(\omega)^2\ d\mathbf{P}(\omega) - \mathbf{E}_{\mathbf{P}}(S)^2, \tag{12}$$

provided the random variable S is sufficiently regular (for instance in $L^2_{\mathbf{P}}(\Omega)$).

Remark 2.1 (*Outputs*). It is possible to consider other (and multiple) outputs within the RB approach. Essentially these outputs should be empirical estimations for functionals of $U(\cdot,\omega)$ that are continuous with respect to some $L^p_{\mathbf{P}}(\Omega, H^1(\mathscr{D}))$ topology $(1 \leqslant p \leqslant +\infty)$. Note that interesting cases such as $p = +\infty$ above, and pointwise values of a cumulative distribution function

$$\mathbf{P}\{\omega \in \Omega | \mathscr{E}(U(\cdot,\omega)) \leqslant \mathscr{E}_0\}$$

for some finite numbers $\mathscr{E}_0 \in \mathbb{R}$, are covered by this first RB approach. Indeed, assuming smoothness in ω, one can bin the range of the random variable $\mathscr{E}(U(\cdot,\omega))$, and use a tree algorithm to account for the variations inside the confidence interval obtained for each realization S^m $(1 \leqslant m \leqslant M)$ of $S(\omega) = \mathscr{E}(U(\cdot,\omega))$ using the RB approach. If a confidence interval Δ^0_m is associated to each realization S^m and overlaps $n_m \in \mathbb{N}$ bins, then computing the confidence interval for the output cumulative distribution function amounts to a search for the extreme variations in the output among the $(\Pi_{m=1}^M n_m)$ leaves of the tree.

In the numerical application of Section 3, the statistical outputs (12) (expected value and variance of the integrated temperature at the bottom surface Γ_{R} of the heat sink) will be explored in a many-query context (of design optimization for instance) as functions of the "additional" deterministic parameter $\varrho = (\kappa, \overline{\mathrm{Bi}})$ in the range $\mathcal{A}^{\varrho}$, where

$$\overline{\mathrm{Bi}} := \frac{1}{|\Gamma_{\mathrm{B}}|}\int_{\Gamma_{\mathrm{B}}} \mathbf{E}_{\mathbf{P}}(\mathrm{Bi}).$$

A detailed stochastic description of the random field $\mathrm{Bi}(\cdot,\omega)$ used for numerics is given in Section 2.4.

2.3. Different discretization formulations

Much attention has been devoted recently to the development and the numerical analysis of various numerical approaches for BVPs involving SPDEs *e.g.* [2,3,9,13,16,25,29,33,44,51,53,54,56]. Our RB approach specifically aims at reducing the number of computations in many of the previously developed frameworks without any loss in precision by (i) splitting the computations into *Offline* and *Online* steps, and (ii) maintaining accuracy control through *a posteriori* error estimation of the outputs. The RB approach applies to those formulations that are variational with respect to variables in the physical space $\mathscr{D}$, which we denote $\mathscr{D}$-weak formulations, and can be combined with different treatments of the probabilistic dependence. The latter fall into two main categories: the Ω-strong/$\mathscr{D}$-weak formulations; and the Ω-weak/$\mathscr{D}$-weak formulations. Although we shall only deal with Ω-strong/$\mathscr{D}$-weak formulations in the rest of this paper, our RB approach applies equally well to many Ω-weak/$\mathscr{D}$-weak formulations, as already discussed in the introduction. It is for this reason that we briefly summarize the principles of each of the different formulations so as to make it clear how our RB approach would adapt to Ω-weak/$\mathscr{D}$-weak formulations. (Both formulations have been studied extensively before, though typically by different authors; a few studies already compare both formulations [29,3], but it may be interesting to reevaluate such comparisons between formulations from the viewpoint of our RB approach.)

2.3.1. Strong–weak formulations

If the Biot number $\mathrm{Bi}(\cdot,\omega)$ is a non-degenerate positive random field on the (non-negligible) subset Γ_{B} of $\partial\mathscr{D}$, that is if there exist two constants $0 < \bar{b}_{\min} < \bar{b}_{\max} < +\infty$ such that $\mathbf{P}$-a.s.

$$\mathrm{Bi}(\cdot,\omega) \in (\bar{b}_{\min}, \bar{b}_{\max}) \quad \text{a.e. in } \Gamma_{\mathrm{B}}, \tag{13}$$

or equivalently $\mathrm{Bi}(\cdot,\omega), \mathrm{Bi}^{-1}(\cdot,\omega) \in L^{\infty}_{\mathbf{P}}(\Omega, L^{\infty}_{\mathbf{P}}(\Gamma_{\mathrm{B}}))$, then, by virtue of the Lax–Milgram theorem, there exists a unique (weak) solution $U(\cdot,\omega) \in H^1(\mathscr{D})$ to (1) and (2), satisfying (14) $\mathbf{P}$-a.s.:

$$\int_{\mathscr{D}_1} \nabla U(\cdot,\omega)\cdot\nabla v + \kappa\int_{\mathscr{D}_2} \nabla U(\cdot,\omega)\cdot\nabla v + \int_{\Gamma_{\mathrm{B}}} \mathrm{Bi}(\cdot,\omega)\ U(\cdot,\omega)v = \int_{\Gamma_{\mathrm{R}}} v, \quad \forall v \in H^1(\mathscr{D}). \tag{14}$$

Furthermore, from (13), we have the stability result:

$$\|U(\cdot,\omega)\|_{1,\mathscr{D}} \leqslant \frac{C_1(\mathscr{D})}{\min\{1, \kappa_{\min}, \bar{b}_{\min}\}}, \tag{15}$$

and $\|U(\cdot,\omega)\|_{1,\mathscr{D}} \in L^{\infty}_{\mathbf{P}}(\Omega)$ (with $C_1(\mathscr{D})$ a constant positive real number that depends only on $\mathscr{D}$).

Strong–weak formulations then use the fact that we also have $S \in L^{\infty}_{\mathbf{P}}(\Omega) \subset L^2_{\mathbf{P}}(\Omega)$, where the functional $S(\omega) = \mathscr{E}(U(\cdot,\omega))$ makes sense since, using (11) and (15), the trace of $U(\cdot,\omega)$ on the boundary segment Γ_{R} is well-defined. The outputs $\mathbf{E}_{\mathbf{P}}(S), \mathbf{Var}_{\mathbf{P}}(S)$ are thus approximated as the empirical Monte-Carlo estimations (4) where $\{S^m, m = 1,\ldots,M\}$ are M independent copies (with same law) of the random variable S, and with the following convergence properties (by virtue of the Strong Law of Large Numbers)

$$E_M[S] \underset{M\to+\infty}{\overset{\mathbf{P}-\text{a.s.}}{\longrightarrow}} \mathbf{E}_{\mathbf{P}}(S), \quad V_M[S] \underset{M\to+\infty}{\overset{\mathbf{P}-\text{a.s.}}{\longrightarrow}} \mathbf{Var}_{\mathbf{P}}(S). \tag{16}$$

Hence a major advantage of the Ω-strong/$\mathscr{D}$-weak formulations is to permit the direct application of classical computational procedures (in particular, FE) for the numerical approximation of deterministic BVPs such as (14) in their usual form, without any modification. Many (many...) computations of such parametrized approximate solutions can then be combined – according to (the numerical simulation of) the law of the random field parameter $\mathrm{Bi}(\cdot,\omega)$ – to form the MC evaluations. Such formulations are thus very simple from the implementation viewpoint, presuming that we can readily simulate the law of $\mathrm{Bi}(x_k,\omega)$ at those discrete (e.g., quadrature or nodal) points x_k in the physical domain $\mathscr{D}$ required by the numerical approximation of (14). Note that the latter point is of course true for all formulations, but seems less stringent for the Ω-strong/$\mathscr{D}$-weak formulation (see Section 2.4.1).

However, the convergence (in probability) of SLLN will be slow – the rate of convergence for $E_M[S]$ is governed by the ratio of the variance of S (or its MC counterpart $V_M[S]$) to $\sqrt{M}$ by virtue of the Central Limit Theorem (CLT). This slow convergence is a strong limitation in the application of Ω-strong/ $\mathscr{D}$-weak formulations. Variance reduction techniques, such as Quasi-Monte-Carlo (QMC) methods based on low-discrepancy sequences of random numbers [51], have been developed to reduce the statistical error of the empirical estimations (4). And the RB approach itself brings new possibilities to addressing this slow convergence problem, not by directly reducing the number of necessary outcomes in the MC sums, but rather by improving the numerical treatment of many slow-varying outcomes.

In Section 3, we shall show how to apply our RB approach to the numerical approximation of Ω-strong/ $\mathscr{D}$-weak formulations by taking advantage of the parametrized character of the BVP. We first map outcomes of stochastic coefficients to deterministic values of

the parameters; we then reduce the computational cost of numerical approximations of the BVP for many values of the parameter by splitting the computations into Offline–Online steps; finally, we introduce *a posteriori* error control on the accuracy of the RB-KL approximations (relative to very accurate approximations in high-dimensional discretization-probability space). (We do not consider here variance reduction strategies.)

2.3.2. Weak–weak formulations

Assuming (13) again for well-posedness, the Ω-weak/ $\mathscr{D}$-weak formulations discretize a variational formulation of the original BVP on the full tensor–product space $\Omega\times\mathscr{D}$

$$\begin{aligned}&\int_\Omega d\mathbf{P}(\omega)\int_{\mathscr{D}_1}\nabla U(\cdot,\omega)\cdot\nabla v(\cdot,\omega)+\kappa\int_\Omega d\mathbf{P}(\omega)\int_{\mathscr{D}_2}\nabla U(\cdot,\omega)\cdot\nabla v(\cdot,\omega)\\&+\int_\Omega d\mathbf{P}(\omega)\int_{\Gamma_B}\mathrm{Bi}(\cdot,\omega)U(\cdot,\omega)v(\cdot,\omega)\\&=\int_\Omega d\mathbf{P}(\omega)\int_{\Gamma_R}v(\cdot,\omega),\quad\forall v(\cdot,\omega)\in L^2_{\mathbf{P}}\left(\Omega,H^1(\mathscr{D})\right)\end{aligned}\tag{17}$$

to compute an approximation of a weak solution $U(\cdot,\omega)\in L^2_{\mathbf{P}}(\Omega,H^1(\mathscr{D}))$ satisfying (17), typically through Galerkin projections over tensor–product approximation subspaces of the Hilbert space $L^2_{\mathbf{P}}(\Omega,H^1(\mathscr{D}))$ defined over the (high-dimensional) domain $\Omega\times\mathscr{D}$. The computations of $\mathbf{E}_{\mathbf{P}}(S)$ and $\mathbf{Var}_{\mathbf{P}}(S)$ are then effected by quadrature (or collocation) formulæ in $\Omega\times\mathscr{D}$ once discrete approximations for $U(\cdot,\omega)$ have been computed.

The weak–weak formulations may thus require less regularity (in fact, this seems very useful for input random fields that do not fulfill (13) but only a weaker assumption for well-posedness), although it also seems essential to the Ω-weak/ $\mathscr{D}$-weak formulations that $\mathrm{Bi}(\cdot,\omega)$ be compatible with tensor–product approximations (see Section 2.4.1: this adds condition on $\mathrm{Bi}(\cdot,\omega)$ in comparison with the Ω-strong/ $\mathscr{D}$-weak formulations). The weak–weak formulations essentially provide greatly improved convergence relative to SLLN (in fact, convergence is often improved only for small dimensions, where numerical approaches for this formulation are sufficiently simple).

For instance, after substituting in (17) a truncated version (6) of the KL expansion (5) of $\mathrm{Bi}(\cdot,\omega)$ using K (with $1\leqslant K\leqslant\mathscr{K}$) independent identically distributed (i.i.d.) random variables in a complete set $\{Z_k,\ k\in\mathbb{N}\}$ of $L^2_{\mathbf{P}}(\Omega)$, the seminal work [16] used so-called spectral (stochastic) Galerkin methods, in which $L^2_{\mathbf{P}}(\Omega,H^1(\mathscr{D}))$ is discretized by tensor products of classical discrete approximations for the variational formulation of a BVP in $H^1(\mathscr{D})$ (such as FE) multiplied by orthogonal polynomials $\{H_n,n\in\mathbb{N}\}$ of the random variables $\{Z_k,k\in\mathbb{N}\}$

$$H_0,\quad H_1(Z_k(\omega)),\quad H_2(Z_{k_1}(\omega),Z_{k_2}(\omega)),\ldots,\quad k,k_1,k_2\in\mathbb{N},\\ k_1\geqslant k_2\geqslant 0,\ldots.$$

(In the original Polynomial Chaos (PC) expansion of Wiener [55] for $L^2_{\mathbf{P}}(\Omega)$, the H_n are Hermite polynomials and the variates Z_k are Gaussian; this expansion has then been generalized to other couples of polynomials and probability distributions [56,47].) The Galerkin projections in the stochastic variable that truncate the PC expansions at polynomial order $L\in\mathbb{N}_{>0}$ ($L\geqslant K$), hence using $D=K+L-1$ i.i.d variates $Z_k(\omega)$, then result in a p-dimensional vector space

$$\mathbf{Span}\big(H_l(Z_{k_1},\ldots,Z_{k_l})\mid 0\leqslant l\leqslant L,K+L-1\geqslant k_l>\cdots>k_1\geqslant 1,\\ \{k_1,\ldots,k_l\}\cap\{1,\ldots,K\}\neq\emptyset\big)$$

with $p=1+\sum_{l=1}^{L}\sum_{k=1}^{l}\binom{K}{k}\binom{L-1}{l-k}$. Equivalently, the variational formulation (17) is projected onto the (very high) $(d+D)$-dimensional domain in which $(x,Z_1,\ldots,Z_D)$ take its values. (Alternatively, the discretization level in each direction of the tensor–product Galerkin approximations can be tailored to achieve rapid convergence with respect to the number of degrees of freedom (d.o.f.). In fact, *a posteriori* error indicators and reduced spaces – though quite different from the error bounds and reduced basis spaces presented in the present paper – can serve to identify efficient truncations [54].)

A major limitation of such spectral Galerkin methods is the high-dimensionality of the approximation spaces for (truncated) PC expansions (p increases rapidly with K and L), which necessitates complicated (though certainly often efficient) numerical strategies in order to maintain sparsity on the discretization grid [3,13,29,50,54]. There are many approaches to this *curse of dimensionality*, most of which have already been mentioned in the introduction. The essential features of our RB approach compared to the other reduction techniques previously applied to SPDEs have also been discussed in the introduction. Clearly, the efficiency of the reduction methods – which are not necessarily incompatible between one another and may thus be combined in future studies – only makes sense in a precise context, where it is clear what has to be reduced, why, and for what purpose.

2.4. Random input field

2.4.1. Karhunen–Loève expansions of random fields

To develop efficient numerical procedures for SPDEs, it has been noted in the above Section 2.3 that it was essential to discretize the (scalar) random input field $\mathrm{Bi}(\cdot,\omega)$ consistently with the discretization of the BVP problem (whatever the formulation). Besides, the (de)coupling of variations of $\mathrm{Bi}(x,\omega)$ on the space variable $x\in\mathscr{D}$ and on the probability variable $\omega\in\Omega$ is also an important feature of the variational problems resulting from our numerical approach. It indeed leads to a parametrized weak form where the parametrization is *affine* (see Section 1.2 for a definition). We thus do not only need to assume the non-degeneracy of the random field $\mathrm{Bi}(\cdot,\omega)$ on Γ_B for well-posedness of the BVP, but also the possibility to rewrite it in a decoupled manner like in the KL expansion (5).

In the present work, we introduce general random input fields $\mathrm{Bi}(x,\omega)$ at a continuous level, defined by an *infinite* collection of correlated random numbers mapped to an *infinite* number of points in the physical domain $\mathscr{D}$. This is typically a situation where the fluctuations are modeled following physical assumptions (statistical mechanics for instance). More precisely, we deal with a random process $(\mathrm{Bi}(x,\cdot))_{x\in\partial\mathscr{D}}$ where $\mathrm{Bi}(x_1,\cdot)$ and $\mathrm{Bi}(x_2,\cdot)$ are not necessarily decorrelated when $x_1\neq x_2$.

For well-posedness of the BVP, we only consider random input fields that satisfy (13). Now, such random fields are in $L^2_{\mathbf{P}}(\Omega,L^2(\partial\mathscr{D}))$. Thus, assuming (13), the random input fields $\mathrm{Bi}(\cdot,\omega)$ in this work always have a KL expansion and can always be generated by decoupled variations in x and ω (possibly asymptotically if $\mathscr{K}$ is infinite) after the well-known Proposition 1 (recalled below). Yet, in cases where there is no other motivation like well-posedness for assuming (13), one should still keep in mind that specific assumptions may be necessary to fulfill the requirement of decoupling – by the way, other expansions than KL might also fulfill that requirement.

Note that in practical engineering situations, $\mathrm{Bi}(\cdot,\omega)$ is often not given but rather constructed from a few measurements only, after solving an inverse problem to assimilate (or calibrate) statistical data (see *e.g.* [24]). Since the inverse problem is solved at the discrete level[5], this yields a *finite* collection of random numbers mapped to a *finite* number of points in the physical domain $\mathscr{D}$, and the assumptions may be simplified.

[5] It is interesting to note that inverse problems are usually solved through optimization algorithms that define a typical many-query context where a RB approach for parametrized PDEs is well motivated.

Proposition 1. *Random fields* $\mathrm{Bi}(\cdot,\omega) \in L^2_{\mathbf{P}}(\Omega, L^2(\partial\mathscr{D}))$ *are in one-to-one correspondence with couples* $(\mathbf{E_P}(\mathrm{Bi}), \mathbf{Cov_P}(\mathrm{Bi})) \in L^2(\partial\mathscr{D}) \times L^2(\partial\mathscr{D}\times\partial\mathscr{D})$ *supplied with a collection of mutually uncorrelated random variables* $\{Z_k(\omega); 1 \leqslant k \leqslant \mathscr{K}\}$ *in* $L^2_{\mathbf{P}}(\Omega)$ *with zero mean and unit variance*

$$\mathbf{E_P}(Z_k) = 0 \quad \mathbf{E_P}(Z_k Z_{k'}) = \delta_{k,k'} \quad \forall 1 \leqslant k,k' \leqslant \mathscr{K}$$
$$\text{(with Kronecker notations, hence } \mathbf{Var_P}(Z_k) = 1),$$

when the kernel $\mathbf{Cov_P}(\mathrm{Bi})$ *defines a positive, self-adjoint, trace class linear operator*

$$\widetilde{T} \in \mathscr{L}\Big(L^2(\partial\mathscr{D}), L^2(\partial\mathscr{D})\Big), \quad (\widetilde{T}f)(x) = \int_{\partial\mathscr{D}} \mathbf{Cov_P}(\mathrm{Bi})(x,y)\, f(y)\, dy, \quad \forall f \in L^2(\partial\mathscr{D}) \tag{18}$$

of (possibly infinite) rank $\mathscr{K}$. *Furthermore, random fields* $\mathrm{Bi}(\cdot,\omega) \in L^2_{\mathbf{P}}(\Omega, L^2(\partial\mathscr{D}))$ *have the following Karhunen–Loève expansion* [26]

$$\mathrm{Bi}(x,\omega) = \mathbf{E_P}(\mathrm{Bi})(x) + \sum_{k=1}^{\mathscr{K}} \sqrt{\tilde{\lambda}_k}\Phi_k(x)\, Z_k(\omega), \quad x \in \partial\mathscr{D}, \tag{19}$$

where $\{\tilde{\lambda}_k; 1 \leqslant k \leqslant \mathscr{K}\}$ *are the positive eigenvalues (in descending order) of the positive, self-adjoint, trace class operator* $\widetilde{T}$ *associated with eigenvectors* $\{\Phi_k(x) \in L^2(\partial\mathscr{D}); 1 \leqslant k \leqslant \mathscr{K}\}$ *(orthonormal in the* $L^2(\partial\mathscr{D})$*-inner-product),*

$$(\widetilde{T}f)(x) = \sum_{1\leqslant k\leqslant\mathscr{K}} \tilde{\lambda}_k \left(\int_{\partial\mathscr{D}} \Phi_k(y) f(y)\, dy\right)\Phi_k(x), \quad \forall f \in L^2(\partial\mathscr{D}),$$

and the random variables $\{Z_k\}$ *are defined by*

$$Z_k(\omega) = \frac{1}{\sqrt{\tilde{\lambda}_k}} \int_{\partial\mathscr{D}} (\mathrm{Bi}(\cdot,\omega) - \mathbf{E_P}(\mathrm{Bi}))\Phi_k, \quad \forall 1 \leqslant k \leqslant \mathscr{K}.$$

Since $L^2(\partial\mathscr{D})$ and $L^2_{\mathbf{P}}(\Omega)$ are Hilbert spaces, the Proposition 1 can be easily proved using Riesz representation theorem, and the Hilbert–Schmidt theorem for bounded (linear) operators of the trace class (then compact) like $\widetilde{T}$ (see *e.g.* [45]).

In the following, we rewrite the usual representation (19) with a scaling parameter $\widetilde{\Upsilon} > 0$,

$$\widetilde{\Upsilon}^2 := \int_{\Gamma_\mathrm{B}}\int_{\Gamma_\mathrm{B}} \mathbf{Cov_P}(\mathrm{Bi})(x,y)\, dx\, dy = \int_{\Gamma_\mathrm{B}} \mathbf{Var_P}(\mathrm{Bi}) = \mathrm{tr}(\widetilde{T}) = \sum_{1\leqslant k\leqslant\mathscr{K}} \tilde{\lambda}_k,$$

and then re-scale the collection of positive eigenvalues as

$$\lambda_k := \frac{\tilde{\lambda}_k}{\widetilde{\Upsilon}^2}, \quad \forall 1 \leqslant k \leqslant \mathscr{K}$$

to obtain the following KL expansion from Proposition 1

$$\mathrm{Bi}(x,\omega) = \mathbf{E_P}(\mathrm{Bi})(x) + \widetilde{\Upsilon}\sum_{k=1}^{\mathscr{K}} \sqrt{\lambda_k}\Phi_k(x) Z_k(\omega), \quad x \in \partial\mathscr{D}.$$

Lastly, when $\mathscr{K}$ is infinite or too large, numerical approaches exploit, instead of the full KL expansion, KL truncations of order K ($K \in \mathbb{N}, 0 < K < \mathscr{K}$) which we write as

$$\mathrm{Bi}_K(x,\omega) = \mathbf{E_P}(\mathrm{Bi})(x) + \widetilde{\Upsilon}\sum_{k=1}^{K} \sqrt{\lambda_k}\Phi_k(x) Z_k(\omega), \quad x \in \partial\mathscr{D}.$$

The truncation error satisfies

$$\mathbf{E_P}\Big((\mathrm{Bi} - \mathrm{Bi}_K)^2\Big) = \widetilde{\Upsilon}^2 \sum_{k=K+1}^{\mathscr{K}} \lambda_k \Phi_k^2(x) \underset{K\to\mathscr{K}}{\to} 0 \quad \text{in } L^1(\partial\mathscr{D}). \tag{20}$$

2.4.2. Additional assumptions on the random input field

In the numerical applications of the next section, we shall require (13) for well-posedness of the BVP. This implies $\mathrm{Bi}(\cdot,\omega) \in L^\infty_{\mathbf{P}}(\Omega, L^\infty(\Gamma_\mathrm{B}))$, thus $\mathrm{Bi}(\cdot,\omega)$ is fully determined by (Proposition 1)

(i) an expected value function $\mathbf{E_P}(\mathrm{Bi})$: $x \in \Gamma_\mathrm{B} \to \mathbf{E_P}(\mathrm{Bi})(x) \in \mathbb{R}$ in $L^\infty(\Gamma_\mathrm{B}) \subset L^2(\Gamma_\mathrm{B})$,

(ii) a covariance function $\mathbf{Cov_P}(\mathrm{Bi})$: $(x,y) \in \Gamma_\mathrm{B}\times\Gamma_\mathrm{B} \to \mathbf{Cov_P}(\mathrm{Bi})(x,y) \in \mathbb{R}$ in $L^2(\Gamma_\mathrm{B}\times\Gamma_\mathrm{B})$, thus the kernel of a positive self-adjoint trace class operator of rank $\mathscr{K}$ with eigenpairs $(\widetilde{\Upsilon}^2\lambda_k, \Phi_k)$ $(\lambda_k \geqslant \lambda_{k+1} > 0, 1 \leqslant k \leqslant \mathscr{K})$ satisfying $\sum_{k=1}^{\mathscr{K}}\lambda_k = 1$ and

$$\int_{\Gamma_\mathrm{B}} \mathbf{Cov_P}(\mathrm{Bi})(x,y)\ \Phi_k(y)\ dy = \widetilde{\Upsilon}^2\lambda_k\Phi_k(x), \qquad \forall x \in \Gamma_\mathrm{B}, \tag{21}$$

(iii) and mutually uncorrelated random variables $\{Z_k \in L^\infty_{\mathbf{P}}(\Omega) \subset L^2_{\mathbf{P}}(\Omega); 1 \leqslant k \leqslant \mathscr{K}\}$ with zero mean and unit variance, through the Karhunen–Loève (KL) expansion

$$\mathrm{Bi}(x,\omega) = \overline{\mathrm{Bi}}\left(G(x) + \Upsilon\sum_{k=1}^{\mathscr{K}} \sqrt{\lambda_k}\Phi_k(x)Z_k(\omega)\right), \tag{22}$$

where $G \in L^\infty(\Gamma_B)$ is a prescribed (deterministic) positive function such that $\mathbf{E_P}(\mathrm{Bi})(\cdot) = \overline{\mathrm{Bi}}\, G(\cdot)$, and $\frac{1}{|\Gamma_B|}\int_{\Gamma_B} G(x)dx = 1$, using the scaling parameters $\overline{\mathrm{Bi}} = \frac{1}{|\Gamma_B|}\int_{\Gamma_B}\mathbf{E_P}(\mathrm{Bi})(x)\,dx$ and $\Upsilon = \tilde{\Upsilon}/\overline{\mathrm{Bi}}$.

For all nonnegative integer $1 \leqslant K \leqslant \mathscr{K}$, we introduce the truncation of KL expansion (22)

$$\mathrm{Bi}_K(x,\omega) = \overline{\mathrm{Bi}}\left(G(x) + \Upsilon\sum_{k=1}^{K} \sqrt{\lambda_k}\ \Phi_k(x)Z_k(\omega)\right). \tag{23}$$

For the sake of consistency of the numerical discretization, we require

$$\|\mathrm{Bi}(\cdot,\omega) - \mathrm{Bi}_K(\cdot,\omega)\|_{L^\infty_{\mathbf{P}}(\Omega,L^\infty(\Gamma_\mathrm{B}))} \overset{K\to\mathscr{K}}{\to} 0, \tag{24}$$

which is stronger than (20) and can be achieved for instance by choosing

(H1) a smooth covariance function $\mathbf{Cov_P}(\mathrm{Bi})$ such that

(H1a)
the eigenvectors are uniformly bounded by some positive real number $\phi > 0$

$$\|\Phi_k\|_{L^\infty(\Gamma_\mathrm{B})} \leqslant \phi, \quad 1 \leqslant k \leqslant \mathscr{K}, \tag{25}$$

(H1b)
the eigenvalues decay sufficiently rapidly,

$$\sum_{k=1}^{\mathscr{K}} \sqrt{\lambda_k} < \infty, \tag{26}$$

(H2) uniformly bounded random variables (say) $\{Z_k; |Z_k(\omega)| < \sqrt{3}, \mathbf{P}-\text{a.s.}\}$.In the numerical results we shall consider Gaussian covariances $\mathbf{Cov_P}(\mathrm{Bi})(x,y) = (\overline{\mathrm{Bi}}\Upsilon)^2 e^{-\frac{(x-y)^2}{\delta^2}}$, with δ a positive real constant, which complies with the requirements above [13]. The fast decay of the eigenvalues in the Gaussian case play an important role in the fast convergence of any numerical discretization based on KL expansions of the input random field; as we shall see, this is true also for our RB approach – the eigenvalues determine the ranges of the parameters, which in turn affect the dimension of the RB space. Next, we shall also insist upon

(H3) *independent* (thus mutually uncorrelated) random variables $\{Z_k; 1 \leqslant k \leqslant \mathscr{K}\}$,
(H4) $Z_k, 1 \leqslant k \leqslant K$, i.i.d. according to the uniform density with respect to the Lebesgue measure on $\mathbb{R}$ in the range $(-\sqrt{3}, \sqrt{3})$,
(H5) Υ chosen such that

$$\tau_0 := \sqrt{3}\Upsilon \sum_{k=1}^{\mathscr{K}} \sqrt{\lambda_k}\|\Phi_k\|_{L^\infty(\Gamma_B)} \leqslant \frac{\min_{x\in\Gamma_B} G(x)}{2}. \tag{27}$$

Then, under our assumptions, the truncation error is bounded above $\forall 1 \leqslant K \leqslant \mathscr{K}$:

$$\begin{aligned} &\|\mathrm{Bi}(\cdot,\omega) - \mathrm{Bi}_K(\cdot,\omega)\|_{L^\infty_{\mathbf{P}}(\Omega,L^\infty(\Gamma_B))} \leqslant \overline{\mathrm{Bi}}\ \tau_K, \\ &\tau_K := \sqrt{3}\Upsilon \sum_{k=K+1}^{\mathscr{K}} \sqrt{\lambda_k}\|\Phi_k\|_{L^\infty(\Gamma_B)}, \end{aligned} \tag{28}$$

and furthermore for $0 < \bar{b}_{\min} \leqslant \frac{\overline{\mathrm{Bi}}}{2}\left(\min_{x\in\Gamma_B} G(x)\right)$ we have **P**-a.s.

$$\mathrm{Bi}_K(\cdot,\omega) \geqslant \bar{b}_{\min} > 0 \quad \text{a.e. in } \mathscr{D}, \quad 1 \leqslant K \leqslant \mathscr{K}. \tag{29}$$

Remark 2.2 (*Choice of the random variables* $\{Z_k\}$). Note that there are many other interesting cases where, for a given smooth covariance function, the random variables $\{Z_k\}$ are not uniformly distributed. These cases will be considered in future studies as they necessitate refinements that would complicate this first exposition of our viewpoint.

3. Reduced basis approach for Monte-Carlo evaluations

3.1. Discretization of a test problem in strong–weak formulation

We now equip the Sobolev space $X := H^1(\mathscr{D})$ with the following inner product for all $w, v \in X$

$$(w, v)_X = \int_{\mathscr{D}_1} \nabla w \cdot \nabla v + \int_{\mathscr{D}_2} \nabla w \cdot \nabla v + \int_{\Gamma_B} wv, \tag{30}$$

and induced norm $\|v\|_X = \sqrt{(v,v)_X}$. It is a standard result that the norm $\|\cdot\|_X$ is equivalent to the usual norm $\|\cdot\|_{1,\mathscr{D}}$ defined previously. We also introduce a finite element (FE) subspace $X_{\mathscr{N}} \subset X$ of dimension $\mathscr{N}$ which inherits the inner product and norm of X. For functions $v \in X_{\mathscr{N}}$, it is possible to define a trace $v \in L^2(\Gamma_B)$ which satisfies

$$\|v\|_{0,\Gamma_B} \leqslant \gamma_{\mathscr{N}}\|v\|_X, \tag{31}$$

where the constant $\gamma_{\mathscr{N}}$ depends only on $\mathscr{D}$ and is bounded above for all $\mathscr{N}$ since

$$\gamma_{\mathscr{N}} \equiv \gamma_{\mathscr{N}}(\mathscr{D}) = \sup_{v\in X_{\mathscr{N}}} \frac{\int_{\Gamma_B} v^2}{\|v\|_X^2} \leqslant \gamma \equiv \sup_{v\in X} \frac{\int_{\Gamma_B} v^2}{\|v\|_X^2}. \tag{32}$$

(Note $\tilde{\gamma}$ of (11) differs from γ of (32) only because of the choice of norm.)

For a given positive scalar κ and a given random input field $\mathrm{Bi}(\cdot,\omega)$, we define

(a) the temperature distribution $U(\cdot,\omega) \in X$ in $\mathscr{D}$,
(b) a FE approximation $U_{\mathscr{N}}(\cdot,\omega) \in X_{\mathscr{N}}$ to the temperature distribution in $\mathscr{D}$,
as the respective solutions to the following variational formulations (33),

$$\begin{aligned} &\int_{\mathscr{D}_1} \nabla U_{(\mathscr{N})}(\cdot,\omega)\cdot\nabla v + \kappa\int_{\mathscr{D}_2} \nabla U_{(\mathscr{N})}(\cdot,\omega)\cdot\nabla v \\ &\quad + \int_{\Gamma_B} \mathrm{Bi}(\cdot,\omega)U_{(\mathscr{N})}(\cdot,\omega)v = \int_{\Gamma_R} v, \quad \forall v \in X_{(\mathscr{N})}, \end{aligned} \tag{33}$$

and, when $\mathrm{Bi}(\cdot,\omega)$ is approximated by $\mathrm{Bi}_K(\cdot,\omega)$,

(c) an approximation $U_K(\cdot,\omega) \in X$ to $U(\cdot,\omega)$,
(d) and a FE approximation $U_{\mathscr{N},K}(\cdot,\omega) \in X_{\mathscr{N}}$ to $U_{\mathscr{N}}(\cdot,\omega)$
as the respective solutions to the following variational formulations (34)

$$\begin{aligned} &\int_{\mathscr{D}_1} \nabla U_{(\mathscr{N}),K}(\cdot,\omega)\cdot\nabla v + \kappa\int_{\mathscr{D}_2} \nabla U_{(\mathscr{N}),K}(\cdot,\omega)\cdot\nabla v \\ &\quad + \int_{\Gamma_B} \mathrm{Bi}_K(\cdot,\omega)U_{(\mathscr{N}),K}(\cdot,\omega)v = \int_{\Gamma_R} v, \quad \forall v \in X_{(\mathscr{N})}, \end{aligned} \tag{34}$$

where the same subscripts into brackets $(\cdot)$ are simultaneously active or not, which means in (a), (b) that (33) holds for $U(\cdot,\omega)$ and X, or $U_{\mathscr{N}}(\cdot,\omega)$ and $X_{\mathscr{N}}$ respectively, in (c), (d) that (34) holds for $U_K(\cdot,\omega)$ and X, or $U_{\mathscr{N},K}(\cdot,\omega)$ and $X_{\mathscr{N}}$, respectively.

With a similar use of the subscripts in $(\cdot)$, we also define (intermediate) outputs as

$$S_{(\mathscr{N})(,K)}(\omega) := \mathscr{E}\big(U_{(\mathscr{N})(,K)}(\cdot,\omega)\big) = \int_{\Gamma_R} U_{(\mathscr{N})(,K)}(\cdot,\omega). \tag{35}$$

We are interested in evaluating the expected value and variance of the integrated temperature $S_{(\mathscr{N})(,K)}(\cdot,\omega)$, which are our (final) statistical outputs:

$$\mathbf{E}_{\mathbf{P}}\big(S_{(\mathscr{N})(,K)}\big) = \int_\Omega S_{(\mathscr{N})(,K)}(\omega)d\mathbf{P}(\omega), \tag{36}$$

$$\mathbf{Var}_{\mathbf{P}}\big(S_{(\mathscr{N})(,K)}\big) = \int_\Omega \big(\mathbf{E}_{\mathbf{P}}\big(S_{(\mathscr{N})(,K)}\big) - S_{(\mathscr{N})(,K)}(\cdot,\omega)\big)^2 d\mathbf{P}(\omega). \tag{37}$$

Since $\mathrm{Bi}_K(\cdot,\omega)$ is **P**-a.s. strictly positive on Γ_B and every $1 \leqslant K \leqslant \mathscr{K}$ (by assumption), the variational problems (33) and (34) are well-posed, and the solutions satisfy the following bound **P**-a.s.

$$\|U_{(\mathscr{N})(,K)}(\cdot,\omega)\|_X \leqslant \frac{C_1'(\mathscr{D})}{\min\{1,\kappa,\bar{b}_{\min}\}} \tag{38}$$

for some positive constant $C_1'(\mathscr{D})$. In addition, we have

Proposition 2. *Under standard regularity hypotheses (as* $\mathscr{N}\to\infty$*) on the family of FE spaces* $X_{\mathscr{N}}$*, the FE approximation converges as* $\mathscr{N}\to\infty$*. Furthermore, under the hypotheses of* Section 2.4.2*, the KL approximation converges as* $K\to\mathscr{K}$*. Finally, the following convergence holds* **P***-a.s.*

$$\begin{array}{ccccc} & S_{\mathscr{N},K}(\omega) & \xrightarrow{\mathscr{N}\to\infty} & S_{,K}(\omega) & \\ K\to\mathscr{K} & \downarrow & & \downarrow & K\to\mathscr{K} \\ & S_{\mathscr{N}}(\omega) & \xrightarrow{\mathscr{N}\to\infty} & S(\omega) & \end{array} \tag{39}$$

Proof. First, for any fixed $1 \leqslant K \leqslant \mathscr{K}$, the **P**-a.s. convergence, as $\mathscr{N}\to 0$, of $U_{\mathscr{N}(,K)}(\cdot,\omega) \to U_{(,K)}(\cdot,\omega)$ in X as $\mathscr{N}\to\infty$ follows under standard hypotheses on the FE spaces $X_{\mathscr{N}}$. Then, by subtracting the variational formulation (34) for $U_{(\mathscr{N}),K}(\cdot,\omega)$ from (33) for $U_{(\mathscr{N})}(\cdot,\omega)$ (in $X_{(\mathscr{N})}$) with $v = U_{(\mathscr{N})}(\cdot,\omega) - U_{(\mathscr{N}),K}(\cdot,\omega)$, we get **P**-a.s.

$$\begin{aligned} &\|U_{(\mathscr{N})}(\cdot,\omega) - U_{(\mathscr{N}),K}(\cdot,\omega)\|_{H^1(\mathscr{D})} \\ &\quad \leqslant C_2(\mathscr{D},\bar{b}_{\min})\|\mathrm{Bi}(\cdot,\omega) - \mathrm{Bi}_K(\cdot,\omega)\|_{L^\infty(\Gamma_B)}\|U_{(\mathscr{N}),K}(\cdot,\omega)\|_{L^2(\Gamma_B)} \end{aligned} \tag{40}$$

for some positive real number $C_2(\mathscr{D},\bar{b}_{\min})$ depending only on $\mathscr{D}$ and $\bar{b}_{\min}$. By compactness of the trace mapping from $H^1(\mathscr{D})$ into $L^2(\partial\mathscr{D})$, the uniform bound (38) for all K and the continuity (24) of $\mathrm{Bi}(\cdot,\omega)$ with respect to the $L^\infty(\Gamma_B)$ norm, we get the **P**-a.s. convergence of $U_{(\mathscr{N})}(\cdot,\omega) \to U_{(\mathscr{N}),K}(\cdot,\omega)$ in X as $K\to\mathscr{K}$. So the following diagram of convergence holds:

$$K \to \mathcal{K} \quad \begin{array}{ccc} U_{\mathcal{N},K}(\cdot,\omega) & \xrightarrow{\mathcal{N}\to\infty} & U_{,K}(\cdot,\omega) \\ \downarrow & & \downarrow \\ U_{\mathcal{N}}(\cdot,\omega) & \xrightarrow{\mathcal{N}\to\infty} & U(\cdot,\omega) \end{array} \quad K \to \mathcal{K} \text{ in } L^\infty_{\mathbf{P}}(\Omega, X) \tag{41}$$

Finally, because $S_{(\mathcal{N})(,K)}(\omega)$ are linear functionals of $U_{(\mathcal{N})(,K)}(\cdot,\omega)$ and by continuity of the trace of $U_{(\mathcal{N})(,K)}(\cdot,\omega) \in H^1(\mathscr{D})$ on Γ_B, the diagram of convergences (39) holds. □

Proposition 3. *Under the same standard regularity hypotheses (as $\mathcal{N} \to \infty$) on the family of FE spaces $X_{\mathcal{N}}$ as in* Proposition 3, *the following convergence holds*

$$K \to \mathcal{K} \quad \begin{array}{ccc} (\mathbf{E}_{\mathbf{P}}(S_{\mathcal{N},K}), \mathbf{Var}_{\mathbf{P}}(S_{\mathcal{N},K})) & \xrightarrow{\mathcal{N}\to\infty} & (\mathbf{E}_{\mathbf{P}}(S_{,K}), \mathbf{Var}_{\mathbf{P}}(S_{,K})) \\ \downarrow & & \downarrow \\ (\mathbf{E}_{\mathbf{P}}(S_{\mathcal{N}}), \mathbf{Var}_{\mathbf{P}}(S_{\mathcal{N}})) & \xrightarrow{\mathcal{N}\to\infty} & (\mathbf{E}_{\mathbf{P}}(S), \mathbf{Var}_{\mathbf{P}}(S)) \end{array} \quad K \to \mathcal{K}\,; \tag{42}$$

Proof. Because $S_{(\mathcal{N})(,K)}(\omega) \in L^\infty_{\mathbf{P}}(\Omega) \subset L^2_{\mathbf{P}}(\Omega)$, we simply use the following estimates which hold for any two linear functionals S_1, S_2 of random fields $U_1(\cdot,\omega), U_2(\cdot,\omega)$ in $L^\infty_{\mathbf{P}}(\Omega, X)$ and some positive constant C_0,

$$\begin{aligned} |\mathbf{E}_{\mathbf{P}}(S_1) - \mathbf{E}_{\mathbf{P}}(S_2)| &\leqslant \int_\Omega d\mathbf{P}(\omega) \int_{\Gamma_R} |U_1(\cdot,\omega) - U_2(\cdot,\omega)| \\ &\leqslant |\Gamma_R| \|U_1(\cdot,\omega) - U_2(\cdot,\omega)\|_{0,\partial\mathscr{D}\times\Omega}, \end{aligned} \tag{43}$$

$$\begin{aligned} &|\mathbf{Var}_{\mathbf{P}}(S_1) - \mathbf{Var}_{\mathbf{P}}(S_2)| \\ &\quad\leqslant C_0 \max_{i=1,2} \|U_i(\cdot,\omega)\|_{0,\partial\mathscr{D}\times\Omega} \|U_1(\cdot,\omega) - U_2(\cdot,\omega)\|_{0,\partial\mathscr{D}\times\Omega} \end{aligned} \tag{44}$$

as well as the uniform bound (38) for all $U_{(\mathcal{N})(,K)}(\cdot,\omega), 1 \leqslant K \leqslant \mathcal{K}$, and the compactness of the trace mapping from $H^1(\mathscr{D})$ into $L^2(\partial\mathscr{D})$. □

Lastly, for all positive integer M, we define, akin to (4), M i.i.d. copies $(S^m_{(\mathcal{N})(,K)})_{1\leqslant m\leqslant M}$ of $S_{(\mathcal{N})(,K)}$ and empirical estimators for the expected values $(\mathbf{E}_{\mathbf{P}}(S_{(\mathcal{N})(,K)}), \mathbf{Var}_{\mathbf{P}}(S_{(\mathcal{N})(,K)}))$ as

$$E_M[S_{(\mathcal{N})(,K)}] = \frac{1}{M}\sum_{m=1}^M S^m_{(\mathcal{N})(,K)}, \tag{45}$$

$$V_M[S_{(\mathcal{N})(,K)}] = \frac{1}{M-1}\sum_{m=1}^M \left(S^m_{(\mathcal{N})(,K)} - E_M[S_{(\mathcal{N})(,K)}]\right)^2. \tag{46}$$

The results in (42) for real numbers $(\mathbf{E}_{\mathbf{P}}(S_{(\mathcal{N})(,K)}), \mathbf{Var}_{\mathbf{P}}(S_{(\mathcal{N})(,K)}))$ also clearly hold **P**-a.s. for the discrete sums $(E_M[S_{(\mathcal{N})(,K)}], V_M[S_{(\mathcal{N})(,K)}])$ for any $M > 0$; and by SLLN, it also **P**-a.s. holds:

$$(E_M[S_{(\mathcal{N})(,K)}], V_M[S_{(\mathcal{N})(,K)}]) \overset{\mathbf{P}-a.s.}{\underset{M\to\infty}{\to}} (\mathbf{E}_{\mathbf{P}}(S_{(\mathcal{N})(,K)}), \mathbf{Var}_{\mathbf{P}}(S_{(\mathcal{N})(,K)})).$$

Now, assume sufficient regularity on the PDE data such that the FE approximations $U_{\mathcal{N}}(\cdot,\omega)$ are **P**-a.s. sufficiently close to $U(\cdot,\omega)$ (for some large $\mathcal{N}$), and that furthermore the accuracy required in the evaluation of the outputs $\mathbf{E}_{\mathbf{P}}(S_{(,K)}), \mathbf{Var}_{\mathbf{P}}(S_{(,K)})$ (respectively $E_M[S_{(,K)}], V_M[S_{(,K)}]$) is provided by $\mathbf{E}_{\mathbf{P}}(S_{\mathcal{N}(,K)}), \mathbf{Var}_{\mathbf{P}}(S_{\mathcal{N}(,K)})$ (respectively $E_M[S_{\mathcal{N}(,K)}], V_M[S_{\mathcal{N}(,K)}]$). Even then, the empirical estimations (45) and (46) will still typically converge slowly: many evaluations of the FE approximation are required (M should be large) for the empirical estimations to be good approximations of the required statistical outputs.

In addition, even if, for a given (supposedly large) M, empirical estimations (45) and (46) are assumed both sufficiently close to the required outputs and accessible to numerical computation for given parameters κ and $\mathrm{Bi}(\cdot,\omega)$, the evaluation of $E_M[S_{\mathcal{N}}]$ and $V_M[S_{\mathcal{N}}]$ for many values of these parameters in a many-query context is arguably prohibitive for a direct FE method.

In summary, the FE method with large $\mathcal{N}$ is too expensive to permit the rapid evaluation of empirical estimations (45) and (46), first for a given large M, and second for many values of the (deterministic and stochastic) parameters κ and $\mathrm{Bi}(\cdot,\omega)$ in a many-query context in which M is fixed (presumably large).

Our reduced basis approach aims at reducing the computational cost of multiple (many) FE computations – without sacrificing certified accuracy – by exploiting the parametric structure of the problem through Offline–Online decompositions.

3.2. Reduced-basis approximation

3.2.1. A deterministic parametrized problem

As mentioned in the introduction, we would like to map the sequence of random variables $(Z_k)_{1\leqslant k\leqslant\mathcal{K}}$ in (22) to random solution fields $U_{(\mathcal{N})(,K)}(\cdot,\omega)$, through the solutions $u_{(\mathcal{N})(,K)}(\cdot;y^{(K)})$ of deterministic BVP PDE problems parametrized by deterministic coefficients $y^{(K)}$, invoking the Doob–Dynkin lemma [37].

Moreover, we would like to study variations of the statistical outputs on an "additional" deterministic parameter ϱ, corresponding to many given values of the (deterministic and stochastic) parameters κ and $\mathrm{Bi}(\cdot,\omega)$; this has also been mentioned previously. In the following, we take as "additional" deterministic parameter

$$\varrho = (\kappa, \overline{\mathrm{Bi}}) \in \Lambda^\varrho.$$

We recall that truncations at order K of $Y = (Y_k)_{1\leqslant k\leqslant\mathcal{K}}$ $(1 \leqslant K \leqslant \mathcal{K})$ have been defined in the introduction as

$$Y^K(\omega) := (Y_1(\omega), \ldots, Y_K(\omega), 0, 0, \ldots), \quad \text{where}$$
$$Y_k(\omega) = \Upsilon\sqrt{\lambda_k}\, Z_k(\omega), \quad 1 \leqslant k \leqslant K.$$

We also recall that has been set

$$y := (y_1, y_2 \ldots) \in \Lambda^y \subset \mathbb{R}^{\mathbb{N}}$$
such that for all finite positive integer $1 \leqslant K \leqslant \mathcal{K}$,

$$y^K := (y_1, \ldots, y_K, 0, 0, \ldots) \in \Lambda^y \quad \text{and the range } \Lambda^y \text{ is the cylinder}$$

$$\Lambda^y := \left[-\sqrt{3}\Upsilon\sqrt{\lambda_1}, +\sqrt{3}\Upsilon\sqrt{\lambda_1}\right] \times \left[-\sqrt{3}\Upsilon\sqrt{\lambda_2}, +\sqrt{3}\Upsilon\sqrt{\lambda_2}\right] \times \cdots \subset \mathbb{R}^{\mathcal{K}}.$$

It is important to note that when the eigenvalues λ_k decay rapidly with k, the extent $2\sqrt{3}\Upsilon\sqrt{\lambda_k}$ of the intervals $[-\sqrt{3}\Upsilon\sqrt{\lambda_k}, +\sqrt{3}\Upsilon\sqrt{\lambda_k}]$ will also shrink rapidly. This small range in the y_k for larger k is one of the reasons the RB approximation developed subsequently will converge quickly.[6] A function $\mathrm{bi}(\cdot;\overline{\mathrm{Bi}}, y)$ has been defined on the boundary, parametrized by $\overline{\mathrm{Bi}}$ and by the deterministic parameters $y_k \in [-\sqrt{3}\Upsilon\sqrt{\lambda_k}, +\sqrt{3}\Upsilon\sqrt{\lambda_k}]$ $(1 \leqslant k \leqslant \mathcal{K})$

$$\mathrm{bi}(x;\overline{\mathrm{Bi}}, y) := \overline{\mathrm{Bi}}\left(G(x) + \sum_{k=1}^{\mathcal{K}} y_k \Phi_k(x)\right), \quad \forall x \in \partial\mathscr{D}; \tag{47}$$

note that the function $\mathrm{bi}(\cdot;\overline{\mathrm{Bi}}, y)$ is well defined since, by assumption, the series (47) absolutely converges in $L^\infty(\Gamma_B)$ for a.e. $y \in \Lambda^y$ (see Section 2.4.2). Lastly, we denote the full parameter as $\mu := (\kappa, \overline{\mathrm{Bi}}, y) \in \Lambda^\mu$ with countably (possibly infinite) entries, and truncated versions with $K+2$ entries (for any finite integer $1 \leqslant K \leqslant \mathcal{K}$)

$$\mu_K := (\kappa, \overline{\mathrm{Bi}}, y^K) \in \Lambda^\mu \equiv \Lambda^\varrho \times \Lambda^y$$

[6] Note we can treat with a single RB many different covariance functions of varying smoothness if we introduce the parameters y_k in the interval (say) $[-\sqrt{3}\Upsilon, \sqrt{3}\Upsilon]$ independent of k such that $y \equiv (y_1, \ldots, y_K) \in \mathscr{L}^y_K \equiv [-\sqrt{3}\Upsilon, \sqrt{3}\Upsilon]^K \subset \mathbb{R}^K$. However, in this case the reduced basis approximation will converge much more slowly since the parameter space $\mathscr{L}^y_K$ is much larger.

where $\Lambda^{\varrho} \subset \mathbb{R}^2_{>0}$ denotes the range of $\varrho = (\kappa, \overline{\mathrm{Bi}})$ (at this point, there is no *a priori* assumption on Λ^{ϱ}: it is some subset of $\mathbb{R}^2_{>0}$ that will be made precise later in the numerical part).

Let us now introduce a deterministic BVP PDE problem parametrized by the deterministic (full) parameter $\mu \in \Lambda^{\mu}$. For every $\mu \in \Lambda^{\mu}$, with notations obviously in accordance with the previous Section 3.1, we define $u(\mu), u_K(\mu_K) \in X$ and $u_{\mathcal{N}}(\mu), u_{\mathcal{N},K}(\mu_K) \in X_{\mathcal{N}}$ as solutions to the respective variational formulations

$$a(u_{(\mathcal{N})(,K)}(\mu_{(K)}), v; \mu_{(K)}) = f(v), \quad \forall v \in X_{(\mathcal{N})}, \tag{48}$$

where the subscripts $(\mathcal{N})$ and (K) are simultaneoulsy active everywhere or not, and where the functional $f(\cdot)$ and the parametrized bilinear form $a(\cdot,\cdot;\mu)$ are given by:

$$f(v) = \int_{\Gamma_{\mathrm{R}}} v, \quad \forall v \in X, \tag{49}$$

$$a(w, v; \mu) = \int_{\mathscr{D}_1} \nabla w \cdot \nabla v + \kappa \int_{\mathscr{D}_2} \nabla w \cdot \nabla v + \int_{\Gamma_{\mathrm{B}}} \mathrm{bi}(\cdot; \overline{\mathrm{Bi}}, y) w v, \quad \forall w, v \in X. \tag{50}$$

We may then define our realization output as

$$s_{(\mathcal{N})(,K)}(\mu_{(K)}) = f\left(u_{(\mathcal{N})(,K)}(\mu_{(K)})\right). \tag{51}$$

Clearly, there exists a sequence $\mathscr{M}$ of random variables in $L^{\infty}_{\mathbf{P}}(\Omega)$, with range Λ^{μ} such that for a.e. ω in Ω it holds $\mathscr{M}(\omega) = (\kappa, \overline{\mathrm{Bi}}, Y(\omega))$. We then define truncations such that **P**-a.s., $\forall 1 \leqslant K \leqslant \mathscr{K}$

$$\mathscr{M}_K(\omega) = (\kappa, \overline{\mathrm{Bi}}, Y^K(\omega)),$$

which implies in return, provided $U_{(\mathcal{N})(,K)}(\cdot, \omega)$ is well defined, that **P**-a.s. holds

$$u_{(\mathcal{N})(,K)}(\mathscr{M}_{(K)}(\omega)) = U_{(\mathcal{N})(,K)}(\cdot, \omega), \quad s_{(\mathcal{N})(,K)}(\mathscr{M}_{(K)}(\omega)) = S_{(\mathcal{N})(,K)}(\omega).$$

Moreover, for each $M > 0$, we define M i.i.d. copies $(\mathscr{M}^m)_{1 \leqslant m \leqslant M}$ of the random variable $\mathscr{M}$ such that the empirical estimations

$$E_M[s_{(\mathcal{N})(,K)}(\mathscr{M}_{(K)})] = \frac{1}{M} \sum_{m=1}^{M} s_{(\mathcal{N})(,K)}(\mathscr{M}^m_{(K)}), \tag{52}$$

$$V_M[s_{(\mathcal{N})(,K)}(\mathscr{M}_{(K)})] = \frac{1}{M-1} \sum_{m=1}^{M} \left(E_M[s_{(\mathcal{N})(,K)}] - s_{(\mathcal{N})(,K)}(\mathscr{M}^m_{(K)})\right)^2, \tag{53}$$

coincide **P**-a.s. with $E_M[S_{(\mathcal{N})(,K)}]$ and $V_M[S_{(\mathcal{N})(,K)}]$ as statistical approximations of the expected value and variance $\mathbf{E_P}(S_{(\mathcal{N})(,K)})$ and $\mathbf{Var_P}(S_{(\mathcal{N})(,K)})$, respectively. Note that all the convergence results established in the previous Section 3.1 for $\mathcal{N}, K \to \infty$ still hold for $s_{(\mathcal{N})(,K)}(\mu_{(K)})$ and a fixed parameter value μ.

In the following, we shall develop a Reduced Basis (RB) approximation and associated *a posteriori* error estimator which will permit rapid and reliable evaluation of the empirical approximations (52) and (53) for the outputs of interest (the expected value and variance $(\mathbf{E_P}(S), \mathbf{Var_P}(S))$). Our RB approximation will be based upon, and the RB error will be measured relative to, the FE approximation $u_{\mathcal{N},K}(\mu_K)$ of (48), for a fixed parameter value $\mu \in \Lambda^{\mu}$. Note we assume that $\mathcal{N}$ is chosen sufficiently large *a priori* to provide the desired accuracy relative to the exact solution; we shall thus concentrate our *a posteriori* estimation and control on the RB approximation and on the KL truncation (note it is very simple to change the order of KL truncation in a strong–weak formulation). As we shall see, the total RB cost (Offline and Online, see Section 3.4) will actually depend rather weakly on $\mathcal{N}$, and hence $\mathcal{N}$ may be chosen conservatively.

3.2.2. RB approximation

Let $N_{\max} \leqslant X$-orthonormalized basis functions $\zeta_n \in X_{\mathcal{N}}, 1 \leqslant n \leqslant N_{\max} (N_{\max} \leqslant \mathcal{N})$ be given, and define the associated hierarchical Lagrange [41] RB spaces $X_N \subset X_{\mathcal{N}}, 1 \leqslant N \leqslant N_{\max}$, as

$$X_N = \mathrm{span}\ \{\zeta_n, 1 \leqslant n \leqslant N\},\ N = 1, \ldots, N_{\max}. \tag{54}$$

In practice (see Section 3.4), the spaces X_N will be generated by a Greedy sampling procedure [32,43]; for our present purpose, however, X_N can in fact represent any sequence of (low-dimensional) hierarchical approximation spaces. Let the KL expansion of the random input field be truncated at some finite order K, the (N,K)-RB approximation of the problem (48) then reads: Given $\mu \in \Lambda^{\mu}$, we look for an RB approximation $u_{N,K}(\mu_K) \in X_N$ such that

$$a_K(u_{N,K}(\mu_K), v; \mu_K) = f(v), \quad \forall v \in X_N. \tag{55}$$

We then calculate the RB realization output as

$$s_{N,K}(\mu_K) = \int_{\Gamma_{\mathrm{R}}} u_{N,K}(\mu_K). \tag{56}$$

The RB output will be evaluated in the Online stage, by the procedure described in Section 3.4, with a computational cost depending on N and K but *not* on $\mathcal{N}$: hence, for small N and K, the RB approximation can be significantly less expensive than the FE approximation.

We shall use this RB approximation to approximate the expected value and variance of the output of interest, for sufficiently large integer $M > 0$, through the empirical estimations

$$E_M[s_{N,K}(\mathscr{M}_K)] = \frac{1}{M} \sum_{m=1}^{M} s_{N,K}(\mathscr{M}^m_K), \tag{57}$$

$$V_M[s_{N,K}(\mathscr{M}_K)] = \frac{1}{M-1} \sum_{m=1}^{M} \left(E_M[s_{N,K}(\mathscr{M}_{(K)})] - s_{N,K}(\mathscr{M}^m_K)\right)^2. \tag{58}$$

In the next section we develop rigorous *a posteriori* bounds for these quantities relative to $E_M[s_{(\mathcal{N}),(\mathscr{K})}(\mathscr{M}_{(K)})]$ and $V_M[s_{(\mathcal{N}),(\mathscr{K})}(\mathscr{M}_{(K)})]$, respectively.

3.3. A posteriori error estimation

3.3.1. Error bounds for the RB output

We note from (55) that, for any $\mu \in \Lambda^{\mu}$, the residual $r(v; \mu_K)$ associated with $u_{N,K}(\mu_K)$ reads

$$r(v; \mu_K) = f(v) - a(u_{N,K}(\mu_K), v; \mu_K), \quad \forall v \in X_{\mathcal{N}}; \tag{59}$$

the dual norm of the residual (defined over the FE "truth" space) is given by

$$\|r(\cdot; \mu_K)\|_{X'_{\mathcal{N}}} = \sup_{v \in X_{\mathcal{N}}} \frac{r(v; \mu_K)}{\|v\|_X}. \tag{60}$$

We next introduce a bilinear form parametrized by the deterministic parameter $\varrho = (\kappa, \overline{\mathrm{Bi}})$ but independent of the parameter y,

$$a_C(w, v; (\kappa, \overline{\mathrm{Bi}})) = \int_{\mathscr{D}_1} \nabla w \cdot \nabla v + \kappa \int_{\mathscr{D}_2} \nabla w \cdot \nabla v + \frac{\overline{\mathrm{Bi}}}{2} \int_{\Gamma_{\mathrm{B}}} G(x) w v, \quad \forall w, v \in X_{\mathcal{N}}, \tag{61}$$

such that, since $\mathrm{Bi}_K(x, y^K) \geqslant \overline{\mathrm{Bi}} G(x)/2, \forall x \in \Gamma_B$, by (27) (assumption H5)

$$a_C(v, v; (\kappa, \overline{\mathrm{Bi}})) \leqslant a(v, v; \mu_K), \quad \forall \mu \in \Lambda^{\mu},\ \forall v \in X_{\mathcal{N}},\ \forall 1 \leqslant K \leqslant \mathscr{K}.$$

Denoting $\alpha(\mu_K)$ the coercivity constant associated with $a(\cdot,\cdot;\mu_K)$, it follows

$$\alpha_C(\kappa, \overline{\mathrm{Bi}}) = \inf_{v \in X_{\mathcal{N}}} \frac{a_C(v, v; (\kappa, \overline{\mathrm{Bi}}))}{\|v\|^2_X} \leqslant \alpha(\mu_K) := \inf_{v \in X_{\mathcal{N}}} \frac{a(v, v; \mu_K)}{\|v\|^2_X}, \quad \forall \mu \in \Lambda^{\mu}. \tag{62}$$

It should be noted that $\alpha_C(\kappa, \overline{\mathrm{Bi}})$ depends only on the deterministic parameters κ and $\overline{\mathrm{Bi}}$, *not* on the (ultimately mapped to a random) parameter y^K! The following result is standard [6,32,43].

Proposition 4. *Given a computable lower bound $\alpha_{\rm LB}$ for $\alpha_C(\kappa,\overline{\rm Bi})$, thus also for $\alpha(\mu_K),\forall\mu\in\mathcal{A}^\mu$, the following a posteriori estimates hold for all positive integers $N,\mathcal{N},K$*

$$\|u_{\mathcal{N},K}(\mu_K)-u_{N,K}(\mu_K)\|_X\leqslant\Delta_{N,K}(\mu_K)\equiv\frac{\|r(\cdot;\mu_K)\|_{X'_\mathcal{N}}}{\alpha_{\rm LB}},\tag{63}$$

$$|s_{\mathcal{N},K}(\mu_K)-s_{N,K}(\mu_K)|\leqslant\Delta^s_{N,K}(\mu_K)\equiv\frac{\|r(\cdot;\mu_K)\|^2_{X'_\mathcal{N}}}{\alpha_{\rm LB}}.\tag{64}$$

3.3.2. Error bounds for the KL truncation effect

We now bound the error $|s_\mathcal{N}(\mu)-s_{\mathcal{N},K}(\mu_K)|$ due to the truncation of the KL expansion for any $\mu\in\mathcal{A}^\mu$, where μ_K is the truncated version that retains the $K+2$ first entries of μ.

Proposition 5. *With the same lower bound $\alpha_{\rm LB}$ as in Proposition 4, $\forall\mu\in\mathcal{A}^\mu$, holds for all positive integer $N,\mathcal{N},K$*

$$|s_\mathcal{N}(\mu)-s_{\mathcal{N},K}(\mu_K)|\leqslant\Delta^t_{N,K}(\mu_K)\equiv\frac{\overline{\rm Bi}\tau_K\gamma_\mathcal{N}}{\alpha_{\rm LB}}\|f\|_{X'_\mathcal{N}}\left(\|u_{N,K}(\mu_K)\|_X+\Delta_{N,K}(\mu_K)\right),\tag{65}$$

where $\Delta_{N,K}(\mu)$ is the error bound defined above in (63) for $\|u_{\mathcal{N},K}(\mu_K)-u_{N,K}(\mu_K)\|_{X_\mathcal{N}}$ and τ_K is the bound introduced in (28).

Proof. First note that

$$\begin{aligned}|s_\mathcal{N}(\mu)-s_{\mathcal{N},K}(\mu_K)|&=|f(u_\mathcal{N}(\mu)-u_{\mathcal{N},K}(\mu_K))|\\&\leqslant\|f\|_{X'_\mathcal{N}}\,\|u_\mathcal{N}(\mu)-u_{\mathcal{N},K}(\mu_K)\|_X.\end{aligned}\tag{66}$$

Then, to get (65), we now show that the last term is bounded by

$$\|u_\mathcal{N}(\mu)-u_{\mathcal{N},K}(\mu_K)\|_X\leqslant\frac{\overline{\rm Bi}\;\tau_K\gamma_\mathcal{N}}{\alpha_{\rm LB}}\left(\|u_{N,K}(\mu_K)\|_X+\Delta_{N,K}(\mu_K)\right),\tag{67}$$

where $\overline{\rm Bi}\;\tau_K$ is the error bound for $\|{\rm bi}(\cdot;\overline{\rm Bi},y)-{\rm bi}(\cdot;\overline{\rm Bi},y^K)\|_{L^\infty(\Gamma_{\rm B})}$ introduced in (28) and $\gamma_\mathcal{N}$ is the continuity constant for the trace application $X_\mathcal{N}\to\Gamma_{\rm B}$ already defined in (31).

To prove (67), we substract the truncated and full problems (48) after FE discretization, and choose $v=e_{\mathcal{N},K}(\mu)=u_\mathcal{N}(\mu)-u_{\mathcal{N},K}(\mu_K)$ as test function. We obtain

$$\begin{aligned}&a(e_{\mathcal{N},K}(\mu),e_{\mathcal{N},K}(\mu);\mu)\\&\quad=-\int_{\Gamma_{\rm B}}({\rm bi}(\cdot;\overline{\rm Bi},y)-{\rm bi}_K(\cdot;\overline{\rm Bi},y^K))\;u_{\mathcal{N},K}(\mu_K)\;e_{\mathcal{N},K}(\mu).\end{aligned}\tag{68}$$

Furthermore, the left-hand side (LHS) of (68) is bounded below by

$${\rm LHS}\geqslant a_C(e_{\mathcal{N},K}(\mu),e_{\mathcal{N},K}(\mu);(\kappa,\overline{\rm Bi}))\geqslant\alpha_{\rm LB}\|e_{\mathcal{N},K}(\mu)\|^2_X,\tag{69}$$

and the right-hand side (RHS) of (68) is bounded above by

$$\begin{aligned}|{\rm RHS}|&\leqslant\overline{\rm Bi}\tau_K\|u_{\mathcal{N},K}(\mu_K)\|_{L^2(\Gamma_{\rm B})}\|e_{\mathcal{N},K}\|_{L^2(\Gamma_{\rm B})}\\&\leqslant\overline{\rm Bi}\;\tau_K\gamma_\mathcal{N}\|u_{\mathcal{N},K}(\mu_K)\|_X\|e_{\mathcal{N},K}(\mu)\|_X\\&\leqslant\overline{\rm Bi}\tau_K\gamma_\mathcal{N}(\|u_{N,K}(\mu_K)\|_X+\Delta_{N,K}(\mu_K))\|e_{\mathcal{N},K}(\mu)\|_X.\end{aligned}\tag{70}$$

The desired result (67) follows directly from (68)–(70). □

3.3.3. Error bounds for the expected value and variance

Using the notations introduced in (64) and (65) we have, from the triangle inequality,

$$|s_\mathcal{N}(\mu)-s_{N,K}(\mu_K)|\leqslant\Delta^o_{N,K}(\mu_K):=\Delta^s_{N,K}(\mu_K)+\Delta^t_{N,K}(\mu_K).\tag{71}$$

Thus we obtain the error bound for the error in the expected value **P**-a.s. as

$$\begin{aligned}&\left|E_M[s_\mathcal{N}(\mathcal{M})]-E_M[s_{N,K}(\mathcal{M}_K)]\right|\\&\quad\leqslant\Delta^o_E[s_{N,K}(\mathcal{M}_K)]:=\Delta^s_E[s_{N,K}(\mathcal{M}_K)]+\Delta^t_E[s_{N,K}(\mathcal{M}_K)],\end{aligned}\tag{72}$$

using M i.i.d. (truncated) copies $(\mathcal{M}^m)_{1\leqslant m\leqslant M}$ of $\mathcal{M}$, and the following random variables:

$$\begin{aligned}\Delta^s_E[s_{N,K}(\mathcal{M}_K)]&\equiv\frac{1}{M}\sum_{m=1}^M\Delta^s_{N,K}(\mathcal{M}^m_K),\\\Delta^t_E[s_{N,K}(\mathcal{M}_K)]&\equiv\frac{1}{M}\sum_{m=1}^M\Delta^t_{N,K}(\mathcal{M}^m_K).\end{aligned}\tag{73}$$

The error bound (72) consists of the RB estimate (64) and the KL truncation estimate (65). The two estimates depend on both N and K but in different ways: the former will decrease rapidly with increasing N and typically increase with increasing K, while the latter will decrease rapidly with increasing K.

For the error bound in the variance, we introduce a function of $\mu\in\mathcal{A}^\mu$ (recall (71))

$$s^\pm_{N,K}(\mu_K):=s_{N,K}(\mu_K)\pm\Delta^o_{N,K}(\mu_K),\tag{74}$$

a random variable that is a sum of MC estimators:

$$E^\pm_M[s_{N,K}(\mathcal{M}_K)]:=E_M[s_{N,K}(\mathcal{M}_K)]\pm\Delta^o_E[s_{N,K}(\mathcal{M}_K)],\tag{75}$$

and random variables parametrized by $\mu_K\in\mathcal{A}^\mu$

$$\begin{aligned}&A_{N,K}(\mathcal{M}_K;\mu_K):=E^+_M[s_{N,K}(\mathcal{M}_K)]-s^-_{N,K}(\mu_K),\\&B_{N,K}(\mathcal{M}_K;\mu_K):=E^-_M[s_{N,K}(\mathcal{M}_K)]-s^+_{N,K}(\mu_K),\\&C_{N,K}(\mathcal{M}_K;\mu_K):=\begin{cases}0\quad\text{if }[s^-_{N,K}(\mu_K),s^+_{N,K}(\mu_K)]\\\cap[E^-_M[s_{N,K}(\mathcal{M}_K)],\\E^+_M[s_{N,K}(\mathcal{M}_K)]]\neq\emptyset\\\min\{|A_{N,K}(\mathcal{M}_K;\mu_K)|,|B_{N,K}(\mathcal{M}_K;\mu_K)|\}\quad\text{otherwise}\end{cases},\\&D_{N,K}(\mathcal{M}_K;\mu_K):=\max\{|A_{N,K}(\mathcal{M}_K;\mu_K)|,|B_{N,K}(\mathcal{M}_K;\mu_K)|\}.\end{aligned}\tag{76}$$

We thus have **P**-a.s.

$$C^2_{N,K}(\mathcal{M}_K;\mu_K)\leqslant(E_M[s_\mathcal{N}(\mathcal{M}_K)]-s_\mathcal{N}(\mu_K))^2\leqslant D^2_{N,K}(\mathcal{M}_K;\mu_K),\tag{77}$$

and hence after summation, also **P**-a.s.

$$V^{\rm LB}_M[s_{N,K}(\mathcal{M}_K)]\leqslant V_M[s_\mathcal{N}(\mathcal{M}_K)]\leqslant V^{\rm UB}_M[s_{N,K}(\mathcal{M}_K)],\tag{78}$$

where we have used the MC estimators

$$\begin{aligned}V^{\rm LB}_M[s_{N,K}(\mathcal{M}_K)]&:=\frac{1}{M-1}\sum_{m=1}^M C^2_{N,K}(\mathcal{M}_K;\mathcal{M}^m_K),\\V^{\rm UB}_M[s_{N,K}]&:=\frac{1}{M-1}\sum_{m=1}^M D^2_{N,K}(\mathcal{M}_K;\mathcal{M}^m_K),\end{aligned}\tag{79}$$

with the same collection $\{\mathcal{M}^m_K\}$ as in the MC estimators (76).

Thus we obtain **P**-a.s. a bound for the error in the variance as

$$\left|V_M[s_\mathcal{N}(\mathcal{M}_K)]-V_M[s_{N,K}](\mathcal{M}_K)\right|\leqslant\Delta^o_V[s_{N,K}(\mathcal{M}_K)]\tag{80}$$

with

$$\begin{aligned}\Delta^o_V[s_{N,K}(\mathcal{M}_K)]\equiv\max\Big\{&|V_M[s_{N,K}(\mathcal{M}_K)]-V^{\rm UB}_M[s_{N,K}(\mathcal{M}_K)]|,|V_M[s_{N,K}(\mathcal{M}_K)]\\&-V^{\rm LB}_M[s_{N,K}(\mathcal{M}_K)]|\Big\}.\end{aligned}\tag{81}$$

This variance error bound also includes both an RB contribution and a KL truncation contribution.

Finally, although it is not our main goal, we point out that without consideration of the KL truncation effect we may also obtain the error bounds (at fixed K)

$$\begin{aligned}\left|E_M[s_{\mathcal{N},K}(\mathcal{M}_K)]-E_M[s_{N,K}(\mathcal{M}_K)]\right|&\leqslant\Delta^s_E[s_{N,K}(\mathcal{M}_K)],\\\left|V_M[s_{\mathcal{N},K}(\mathcal{M}_K)]-V_M[s_{N,K}(\mathcal{M}_K)]\right|&\leqslant\Delta^s_V[s_{N,K}(\mathcal{M}_K)].\end{aligned}\tag{82}$$

Here $\Delta^s_E[s_{N,K}(\mathcal{M}_K)]$ is given by (73), and $\Delta^s_V[s_{N,K}(\mathcal{M}_K)]$ is defined in the same way as $\Delta^o_V[s_{N,K}(\mathcal{M}_K)]$ but replacing $\Delta^o_{N,K}(\mu_K)$ with $\Delta^s_{N,K}(\mu_K)$ in

(74) and $\Delta_E^o[s_{N,K}]$ with $\Delta_E^s[s_{N,K}]$ in (75). We introduce the contribution due to the KL truncation to the variance error bound (80) as

$$\Delta_V^t[s_{N,K}(\mathscr{M}_K)] \equiv \Delta_V^o[s_{N,K}(\mathscr{M}_K)] - \Delta_V^s[s_{N,K}(\mathscr{M}_K)]. \tag{83}$$

3.4. Offline–Online computational approach

3.4.1. Construction–Evaluation decomposition

The system (55) comprises N linear algebraic equations in N unknowns. However, its formation involves entities $\zeta_n, 1 \leqslant n \leqslant N$, associated with the $\mathscr{N}$-dimensional FE approximation space. If we must invoke FE fields in order to form the system *for each new value of* μ, the marginal cost per input–output evaluation $\mu \to s_{N,K}(\mu_K)$ will remain unacceptably large. Fortunately, we can compute this output very efficiently by constructing Offline–Online procedures [32,42,43], as we now discuss.

First, we note that the bilinear form a_K as introduced in (50) can be expressed as the following "affine" decomposition

$$a(w, v; \mu_K) = \sum_{k=1}^{K+3} \Theta_k(\mu_K) a_k(w, v), \quad \forall w, v \in X. \tag{84}$$

Here $\Theta_1(\mu_K) = 1, \Theta_2(\mu_K) = \kappa, \Theta_3(\mu_K) = \overline{\mathrm{Bi}}$, and $\Theta_{3+k}(\mu_K) = \overline{\mathrm{Bi}}\, y_k$, $1 \leqslant k \leqslant K$, are parameter-*dependent* functions, and $a_1(w, v) = \int_{\mathscr{D}_1} \nabla w \cdot \nabla v, a_2(w, v) = \int_{\mathscr{D}_2} \nabla w \cdot \nabla v, a_3(w, v) = \int_{\Gamma_B} G(x) w v$, and $a_{3+k}(w, v) = \int_{\Gamma_B} \Phi_k(\cdot) w v, 1 \leqslant k \leqslant K$, are parameter-*independent* bilinear forms. Note the crucial role of the "separable" (in ω and x) form of the KL expansion is ensuring an affine representation; the affine representation is, in turn, crucial to the Offline–Online strategy.

We next express $u_{N,K}(\mu_K) = \sum_{m=1}^N c_{N,K,m}(\mu_K)\zeta_m$, choose $v = \zeta_n$, $1 \leqslant n \leqslant N$, and invoke the affine representation (84) to write the system (55) as

$$\sum_{m=1}^N \left(\sum_{k=1}^{K+3} \Theta_k(\mu_K) a_k(\zeta_m, \zeta_n) \right) c_{N,K,m}(\mu_K) = f(\zeta_n), \quad 1 \leqslant n \leqslant N, \tag{85}$$

and subsequently evaluate our RB output as

$$s_{N,K}(\mu_K) = \sum_{n=1}^N c_{N,K,n}(\mu_K) f(\zeta_n). \tag{86}$$

We observe that the quantities $a_k(\zeta_m, \zeta_n)$ and $f(\zeta_n)$ are independent of μ and thus can be pre-computed in a Construction–Evaluation decomposition.

In the Construction phase, we form and store the $f(\zeta_n)$ and $a_k(\zeta_m, \zeta_n)$, $1 \leqslant n, m \leqslant N_{\max}, 1 \leqslant k \leqslant K+3$. In the Evaluation phase, we first perform the sum $\sum_{k=1}^{K+3} \Theta_k(\mu_K) a_k(\zeta_m, \zeta_n)$, we next solve the resulting $N \times N$ system (85) to obtain the $c_{N,K,n}(\mu_K), 1 \leqslant n \leqslant N$, and finally we evaluate the output (86). The operation count for the Evaluation phase is $O((K+3)N^2)$ to perform the sum, $O(N^3)$ to invert (85), and finally $O(N)$ to effect the inner product (86); the storage for the Evaluation phase (the data archived in the Construction phase) is only $O(N_{\max} + (K+3)N_{\max}^2)$. The Evaluation cost (operation cost and storage) – and hence marginal cost and also asymptotic average cost – to evaluate $\mu \to s_{N,K}(\mu_K)$ is thus independent of $\mathscr{N}$. The implications are twofold: first, if N and K are indeed small, we shall achieve very fast response in many-query contexts (in which the initial Offline investment is eventually "forgotten"); second, we may choose $\mathscr{N}$ very conservatively – to effectively eliminate the error between the exact and FE predictions – without adversely affecting the Evaluation (marginal) cost.

The Construction–Evaluation for the error bounds is a bit more involved. To begin, we note from standard duality arguments that $\|r(\cdot; \mu_K)\|_{X'_\mathscr{N}} = \|\mathscr{R}_{N,K}(\mu_K)\|_X$; here $\mathscr{R}_{N,K}(\mu_K) \in X_\mathscr{N}$ satisfies $(\mathscr{R}_{N,K}(\mu_K), v)_X = r(v; \mu_K), \forall v \in X_\mathscr{N}$, where $r(v; \mu_K) \equiv f(v) - a(u_N(\mu), v; \mu_K), \forall v \in X_\mathscr{N}$, is the residual introduced earlier. We can thus express (63) and (64) as

$$\Delta_{N,K}(\mu_K) = \frac{\|\mathscr{R}_{N,K}(\mu_K)\|_X}{\alpha_{\mathrm{LB}}}, \quad \text{and} \quad \Delta_{N,K}^s(\mu_K) = \frac{\|\mathscr{R}_{N,K}(\mu_K)\|_X^2}{\alpha_{\mathrm{LB}}}. \tag{87}$$

There are two components to the error bounds: the dual norm of the residual, $\|\mathscr{R}_{N,K}(\mu_K)\|_X$, and our lower bound for the coercivity constant, α_{LB}. The Construction–Evaluation decomposition for the coercivity constant lower bound is based on the Successive Constraint Method (SCM) described in detail in [7,21,43]. We focus here on the Construction–Evaluation decomposition for the dual norm of the residual and express our residual $r(v; \mu_K)$ in terms of (84)

$$(\mathscr{R}_{N,K}(\mu), v)_X = f(v) - \sum_{k=1}^{K+3} \sum_{n=1}^N \Theta_k(\mu) c_{N,K\ n}(\mu) a_k(\zeta_n, v), \quad \forall v \in X,$$

and hence obtain by linear superposition

$$\mathscr{R}_{N,K}(\mu_K) = z_0 + \sum_{k=1}^{K+3} \sum_{n=1}^N \Theta_k(\mu_K) c_{N,K,n}(\mu_K) z_n^k,$$

where $(z_0, v)_X = f(v)$, and $(z_n^k, v)_X = -a_k(\zeta_n, v), \forall v \in X_\mathscr{N}, 1 \leqslant n \leqslant N$, $1 \leqslant k \leqslant K+3$, thus

$$\|\mathscr{R}_{N,K}\|_X^2 = (z_0, z_0)_X + 2 \sum_{k,n=1}^{K+3,N} \Theta_k(\mu_K) c_{N,K\ n}(\mu_K)(z_n^k, z_0)_X + \sum_{k,k',n,n'=1}^{K+3,K+3,N,N} \Theta_k(\mu_K) c_{N,K,n}(\mu_K) \Theta_{k'}(\mu_K) c_{N,K,n'}(\mu_K)(z_n^k, z_{n'}^{k'})_X.$$

Since the $(\cdot,\cdot)_X$ inner products are independent of μ, we can pre-compute these quantities in the Construction–Evaluation decomposition.

In the Construction phase – parameter independent, and performed only once – we find $z_0, z_n^k, 1 \leqslant k \leqslant K+3, 1 \leqslant n \leqslant N$, and then form and store the inner products $(z_0, z_0)_X, (z_n^k, z_0)_X$, $1 \leqslant k \leqslant K+3, 1 \leqslant n \leqslant N$, and $(z_n^k, z_{n'}^{k'})_X$, $1 \leqslant k, k' \leqslant K+3, 1 \leqslant n, n' \leqslant N$. Then, in the Evaluation phase – given any desired value of μ_K – we simply evaluate (87) from the summation (88) and the SCM evaluation for α_{LB} at cost $O((K+3)^2 N^2)$. The crucial point, again, is that the cost and storage in the Evaluation phase – the *marginal* cost for each new value of μ – is independent of $\mathscr{N}$: thus we cannot only evaluate our output prediction but also our rigorous output error bound very rapidly in the many-query (or real-time) context.

Finally, the error bound $\Delta_{N,K}^t(\mu_K)$ of (65) requires additional quantities: $\tau_K, \gamma_\mathscr{N}, \|f\|_{X'_\mathscr{N}}$, and $\|u_{N,K}(\mu_K)\|_X$. Note the first three quantities are independent of μ: τ_K can be pre-computed for any $1 \leqslant K \leqslant \mathscr{K}$ from the expansion (28); $\gamma_\mathscr{N}$ can be pre-computed from the eigenvalue problem (32); and finally $\|f\|_{X'_\mathscr{N}}$ can be pre-computed (by duality) as a standard FE Poisson problem. We note further that

$$\|u_{N,K}(\mu_K)\|_X^2 = \sum_{n,n'=1}^{N,N} c_{N,K,n}(\mu_K) c_{N,K,n'}(\mu_K)(\zeta_n, \zeta_{n'})_X, \tag{89}$$

which readily admits a Construction–Evaluation decomposition; clearly, the Evaluation-phase summation (89) requires only $O(N^2)$ operations. In summary, in the Evaluation phase, we can evaluate $s_{N,K}(\mu_K), \Delta_{N,K}^s(\mu_K), \Delta_{N,K}^t(\mu_K)$, and $\Delta_{N,K}^o(\mu_K)$ at total cost $O(N^3 + (K+3)^2 N^2)$ operations.

3.4.2. Greedy sampling

Finally, we turn to the construction of our reduced basis $\zeta_n, 1 \leqslant n \leqslant N_{\max}$: we pursue a very simple but also very effective Greedy procedure [43]. To initiate the Greedy procedure we specify a very large (exhaustive) "train" sample of n_{train} points in $\mathscr{A}^\mu, \Xi_{\mathrm{train}}$, a maximum RB dimension $N_{\max}$, and an initial (say, random) sample $S_1 = \{\mu^1\}$ and associated RB space X_1. (In actual practice, we typically specify an error tolerance-*cum*-stopping criterion which

then implicitly determines $N_{\max}$.) We specify $K = \mathscr{K}$ (in practice, finite) for the Greedy algorithm described below.

Then, for $N = 1, \ldots, N_{\max}$: Step (1) find $\mu^{N+1} = \arg\max_{\mu \in \Xi_{\text{train}}} \Delta_{N,K}(\mu)$; Step (2) update $S_{N+1} = S_N \cup \mu^{N+1}$ and $X_{N+1} = X_N + \text{span}\{u_{\mathscr{N},K}(\mu^{N+1})\}$. The heuristic is simple: we append to our sample the point μ^{N+1} which is least well represented by the space X_N (as predicted by the error bound associated with our RB Galerkin approximation). In practice, the basis must be orthogonalized with respect to the $(\cdot,\cdot)_X$ inner product; the algebraic system then inherits the conditioning properties of the underlying partial differential equation. Note that the Greedy automatically generates *hierarchical* spaces $X_N, 1 \leqslant N \leqslant N_{\max}$, which is computationally very advantageous.

The important point to note from the computational perspective is that the operation count for a few $N_{\max} \ll \mathscr{N}^k$ steps of the Greedy algorithm (using truncations at order $K = \mathscr{K} \ll \mathscr{N}^k$) is $O(\mathscr{N}^k + n_{\text{train}})$ and *not* $O(\mathscr{N}^k n_{\text{train}})$ (where $O(\mathscr{N}^k)$ is the complexity for numerically solving *one* system of size $\mathscr{N} \times \mathscr{N}$) – and hence much less expensive than classical approaches such as the KL (here Proper Orthogonal Decomposition, or POD) expansion for the sample $(u_{\mathscr{N},K}(\mu))_{\mu \in \Xi_{\text{train}}}$. The reason is simple: in Step (1), to calculate $\Delta_{N,K}(\mu)$ over Ξ_{train}, we invoke the Construction–Evaluation decomposition to obtain (per Greedy cycle) an operation count of $O(NK\mathscr{N}^k) + n_{\text{train}} O(K^2 N^2)$. (Of course, much of the computational economies are due not to the Greedy itself, but rather to the accommodation within the Greedy of the inexpensive error bounds.) As a result, we can take n_{train} very large – often 10^4 or larger – particularly important for the high – $K + P_\varrho$ – dimensional parameter domains encountered in the SPDE context (here P_ϱ is dimension of the deterministic parameter ϱ). Furthermore, extensive numerical results for a wide variety of problems indicate that the Greedy RB space X_N is typically as good as more global (and provably optimal) approaches such as the POD [43]. (Of course, the latter result is norm dependent: the Greedy prefers $L^\infty(\Xi_{\text{train}})$, whereas the POD expansion is optimal in $L^2(\Xi_{\text{train}})$.)

3.4.3. Offline–Online stages

Finally, we delineate Offline and Online stages. The Offline stage comprises the Greedy sampling strategy, and thus appeals to both the Construction and Evaluation phases. The Online stage includes all subsequent evaluations of the RB output and output error bound for many-query computations: it involves only the Evaluation phase, and hence will be extremely rapid.

We now discuss the implications for the MC sums required for the evaluation of our statistical outputs – the focus of the current paper. In particular, it is clear the *total* operation count – Offline and Online – to evaluate $E_M[s_{N,K}(\mathscr{M}_K)], V_M[s_{N,K}(\mathscr{M}_K)], \Delta_E^o[s_{N,K}(\mathscr{M}_K)]$ and $\Delta_V^o[s_{N,K}(\mathscr{M}_K)]$ for J different values of $\varrho = (\kappa, \overline{\text{Bi}})$ scales as $W_{\text{Offline}}(N_{\max}, \mathscr{K}, \mathscr{N}) + W_{\text{Online}}(J, M, N, K)$ where

$$W_{\text{Offline}}(N, K, \mathscr{N}) = O(NK\mathscr{N}^k) + n_{\text{train}} O(K^2 N^2) \quad \text{and}$$
$$W_{\text{Online}}(J, M, N, K) = JM \times O(N^3 + K^2 N^2).$$

Thus as either $M \to \infty$ or $J \to \infty$ and in particular as $J, M \to \infty$ – many evaluations of our statistical output – $W_{\text{Offline}} \ll W_{\text{Online}}$. We further note that if $N, K \ll \mathscr{N}$ then $W_{\text{Online}} \ll W_{\text{FE}} \equiv JM(O(\mathscr{N}^k))$, where W_{FE} is the operation count for standard FE evaluation of the MC sums. Hence the interest in the RB approach. In addition, here are two final observations. First, a "con": as we consider less smooth covariance functions with less rapidly decaying spectra not only – for a fixed desired accuracy – will K increase, but also N will increase (due to the more extended domain Λ_K^y). Clearly for sufficiently non-smooth covariances the RB approach will no longer be competitive. Second, a "pro": the *a posteriori* error bounds will permit us to choose N and K minimally – for minimum computational effort – without sacrificing accuracy and certainty.

3.5. Numerical results

In this section, we present numerical results for the model problem described in Section 3.1. We consider a *homogeneous random input field* with:

- a uniform mean, thus $G(x) \equiv 1$,
- and a finite-rank covariance kernel $\mathbf{Cov}_{\mathbf{P}}(\text{Bi})(x,y)$ that coincides with the first $\boxed{\mathscr{K} = 25}$ terms in the KL expansion of $(\overline{\text{Bi}}\Upsilon)^2 e^{-\frac{(x-y)^2}{\delta^2}}$.

The "additional" deterministic parameter $\varrho = (\kappa, \overline{\text{Bi}})$ shall take value in the range $\Lambda^\varrho = [0.1, 10] \times [0.1, 1]$. For the "truth" FE approximation, we use a regular mesh with quadratic elements and $\mathscr{N} = 6882$ degrees of freedom.

First, we choose $\boxed{\delta = 0.5}$ (recall that the length of Γ_{B} is 4, and hence δ is reasonably "small") – we shall subsequently consider even smaller δ.

We calculate the eigenvalues and eigenvectors of $\mathbf{Cov}_{\mathbf{P}}(\text{Bi})(x,y)$ using the standard (Matlab®) Arpack routines. We present in Fig. 2 the eigenvalues λ_k as a function of k; we observe that the eigenvalues decay exponentially with respect to k^2, which is in good agreement with theoretical bounds [45]. Then, to satisfy our assumption (27), we set $\boxed{\tau_0 = \frac{1}{2}}$ which yields the requirement $\Upsilon \leqslant \Upsilon_{\max} \equiv 0.058$. In the following numerical example, we choose $\boxed{\Upsilon = \Upsilon_{\max} \equiv 0.058}$.

We first report results for the case $\boxed{\kappa = 2.0}$ and $\boxed{\overline{\text{Bi}} = 0.5}$. We show in Fig. 3 four realizations $(\text{bi}(x; \overline{\text{Bi}}, y_i^K))_{1 \leqslant i \leqslant 4}$ of the Biot number, and in Fig. 4 the corresponding temperature fields $u_{\mathscr{N},K}(\mu_K^i)$ (where $K = \mathscr{K}$).

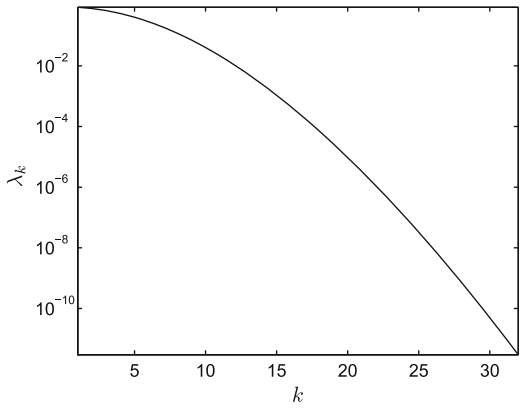


Fig. 2. Eigenvalues λ_k as functions of k.

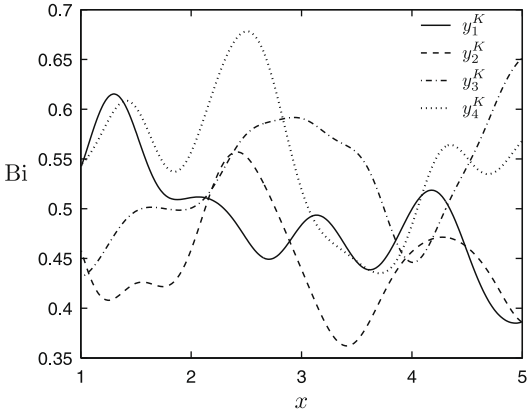


Fig. 3. Four realizations of the Biot number $x \to \text{bi}(x; \overline{\text{Bi}} = 0.5, y_i^K) - 1 \leqslant i \leqslant 4$.

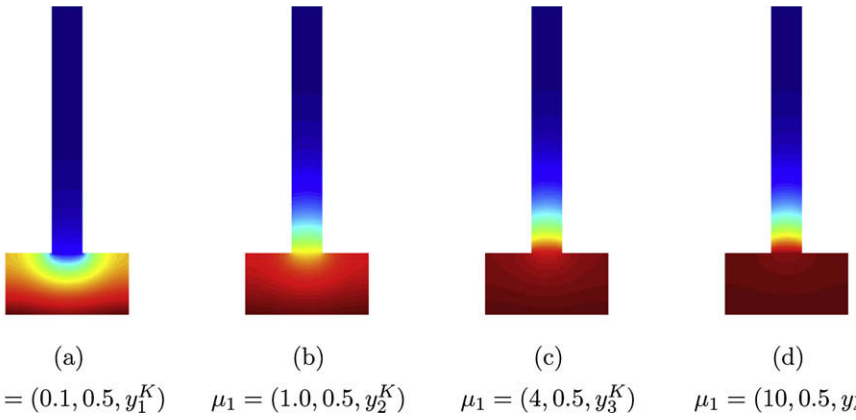


Fig. 4. The temperature field $u_{\mathcal{N},K}(\mu_i)$ for four different realizations $\mu_i = (\kappa_i, \overline{\mathrm{Bi}} = 0.5, y_i)$ – $1 \leqslant i \leqslant 4$ – when $K = \mathcal{K}$, corresponding to the four realizations of $\mathrm{bi}(\cdot; \overline{\mathrm{Bi}} = 0.5, y_i^K)$ in Fig. 3.

RB approximation: we present in Fig. 5 the five leading basis functions $(\zeta_n)_{n=1,2,\ldots,5}$ obtained by pursuing the Greedy sampling procedure over a training set Ξ_{train} of $n_{\mathrm{train}} = 10,000$ parameter points randomly selected with uniform law in the parameter space $\mathcal{A}^\mu$. Note $n_{\mathrm{train}} = 10,000$ is arguably adequate given the rapid decay of the eigenvalues. In any event, our *a posteriori* error bounds will certify (in the Online stage) the accuracy of our RB predictions. The Greedy procedure terminates when a maximum number of basis functions $N_{\max} = 18$ is reached, while the maximum error bound $\Delta_{N,K,\max} = \max_{\mu \in \Xi_{\mathrm{train}}} \Delta_{N,K}(\mu)$ is less than 5×10^{-3}.

Statistical outputs: we present in Fig. 6 the expected value and variance as a function of M, obtained for $N = 10$ and $K = 20$ (note that we do not need to repeat the Offline stage for different M.) We next choose $M = 10,000$ for our Monte-Carlo sums. We show in Table 1 the expected value and associated error bound for the integrated temperature at the bottom surface of the fin as a function of $N(\leqslant N_{\max})$ and $K(\leqslant \mathcal{K})$. Table 2 displays the corresponding variance and associated error bound. Fig. 7a and b show the error bounds for the expected value and variance, respectively.

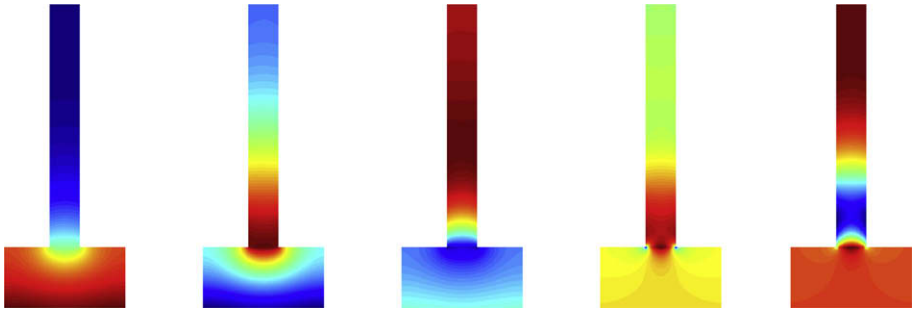
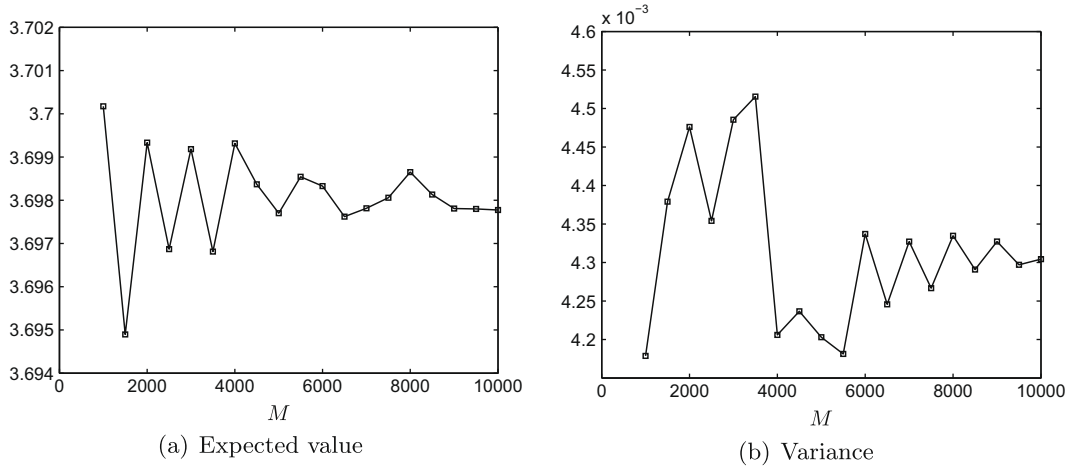


Fig. 5. The five leading RB basis functions $(\zeta_n)_{n=1,2,\ldots,5}$, ordered from left to right as successively chosen (and orthonormalized) by the Greedy sampling procedure.



(a) Expected value (b) Variance

Fig. 6. Outputs $E_M[s_{N,K}(\mathcal{M}_K)], V_M[s_{N,K}(\mathcal{M}_K)]$ as functions of M, with $\varrho = (2.0, 0.5)$.

Table 1
Expected value $E_M[s_{N,K}(\mathscr{M}_K)]$ and error bound $\Delta_E^o[s_{N,K}(\mathscr{M}_K)]$ for different values of the RB dimension N and of the KL truncation order K with $\varrho = (\kappa = 2.0, \overline{\mathrm{Bi}} = 0.5)$.

N	$K = 5$		$K = 10$		$K = 15$		$K = 20$	
	$E_M[s_{N,K}]$	$\Delta_E^o[s_{N,K}]$	$E_M[s_{N,K}]$	$\Delta_E^o[s_{N,K}]$	$E_M[s_{N,K}]$	$\Delta_E^o[s_{N,K}]$	$E_M[s_{N,K}]$	$\Delta_E^o[s_{N,K}]$
2	3.2602	4.74×10^{0}	3.2599	2.23×10^{0}	3.2600	1.59×10^{0}	3.2600	1.51×10^{0}
4	3.6920	2.20×10^{0}	3.6947	5.08×10^{-1}	3.6941	7.18×10^{-2}	3.6942	1.60×10^{-2}
6	3.6972	2.09×10^{0}	3.6974	4.76×10^{-1}	3.6979	5.80×10^{-2}	3.6966	4.54×10^{-3}
8	3.6981	2.09×10^{0}	3.6975	4.74×10^{-1}	3.6969	5.77×10^{-2}	3.6986	4.33×10^{-3}
10	3.6974	2.08×10^{0}	3.6977	4.71×10^{-1}	3.6976	5.69×10^{-2}	3.6978	3.94×10^{-3}
12	3.6973	2.07×10^{0}	3.6976	4.70×10^{-1}	3.6981	5.68×10^{-2}	3.6976	3.90×10^{-3}
14	3.6975	2.07×10^{0}	3.6974	4.70×10^{-1}	3.6977	5.68×10^{-2}	3.6978	3.89×10^{-3}

Table 2
Variance $V_M[s_{N,K}(\mathscr{M}_K)]$ and error bound $\Delta_V^o[s_{N,K}(\mathscr{M}_K)]$ for different values of the RB dimension N and of the KL truncation order K with $\varrho = (\kappa = 2.0, \overline{\mathrm{Bi}} = 0.5)$.

N	$K = 5$		$K = 10$		$K = 15$		$K = 20$	
	$V_M[s_{N,K}]$	$\Delta_V^o[s_{N,K}]$	$V_M[s_{N,K}]$	$\Delta_V^o[s_{N,K}]$	$v_M[s_{N,K}]$	$\Delta_V^o[s_{N,K}]$	$V_M[s_{N,K}]$	$\Delta_V^o[s_{N,K}]$
2	0.0039	9.38×10^{-1}	0.0041	4.38×10^{-1}	0.0041	3.23×10^{-1}	0.0041	3.00×10^{-1}
4	0.0039	4.54×10^{-1}	0.0045	1.11×10^{-1}	0.0045	1.56×10^{-2}	0.0045	3.52×10^{-3}
6	0.0037	4.05×10^{-1}	0.0043	1.02×10^{-1}	0.0043	1.23×10^{-2}	0.0043	9.89×10^{-4}
8	0.0037	4.05×10^{-1}	0.0043	1.08×10^{-1}	0.0043	1.26×10^{-2}	0.0043	9.09×10^{-4}
10	0.0038	4.16×10^{-1}	0.0043	9.72×10^{-2}	0.0043	1.24×10^{-2}	0.0043	8.32×10^{-4}
12	0.0038	4.16×10^{-1}	0.0043	9.72×10^{-2}	0.0043	1.24×10^{-2}	0.0043	8.36×10^{-4}
14	0.0038	4.12×10^{-1}	0.0043	9.72×10^{-2}	0.0043	1.23×10^{-2}	0.0043	8.46×10^{-4}

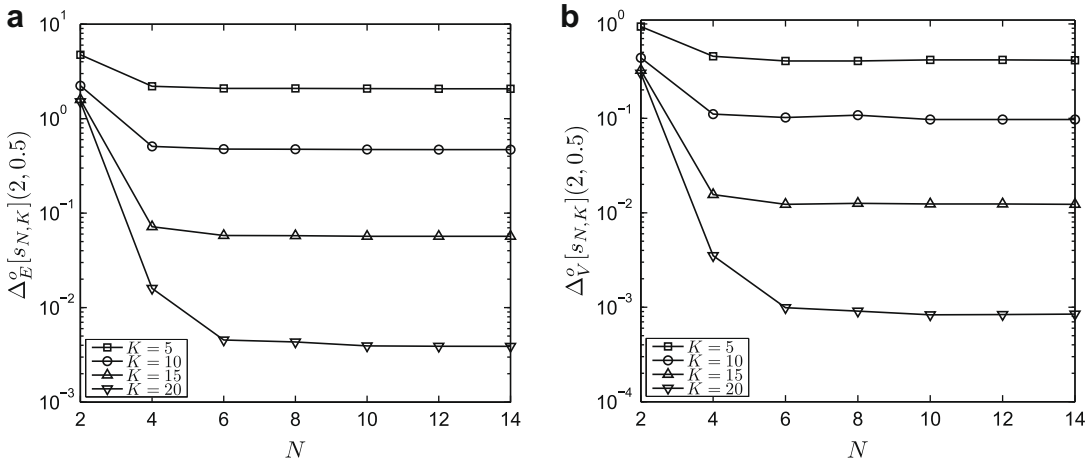


Fig. 7. (a) $\Delta_E^o[s_{N,K}(\mathscr{M}_K)]$ and (b) $\Delta_V^o[s_{N,K}(\mathscr{M}_K)]$ as functions of N and K; $\varrho = (2.0, 0.5)$.

(N, K)**-variations:** we observe that the error bounds $\Delta_E^o[s_{N,K}(\mathscr{M}_K)]$ and $\Delta_V^o[s_{N,K}(\mathscr{M}_K)]$ depend on N and K in a strongly coupled manner: for a fixed value of K the error bounds initially decrease with increasing N and then level off for N large; when the error bounds no longer improve with increasing N, increasing K further reduces the error. This behavior of the error bounds is expected since the accuracy of our predictions is limited by both the RB error bound $\Delta_{N,K}^s(\mu)$ and the KL truncation error bound $\Delta_{N,K}^t(\mu)$: the former decreases rapidly with increasing N only while the latter decreases rapidly with increasing K only. We note that the KL truncation error bounds, $\Delta_E^t[s_{N,K}(\mathscr{M}_K)]$ and $\Delta_V^t[s_{N,K}(\mathscr{M}_K)]$, dominate the RB error bounds $\Delta_E^s[s_{N,K}(\mathscr{M}_K)]$ and $\Delta_V^s[s_{N,K}(\mathscr{M}_K)]$, respectively, as shown in Figs. 8 and 9.

Reduction efficiency: the expectation and variance error bounds (and the actual errors) decrease very rapidly as both N and K increase (such a rapid convergence is expected because the solution is very smooth with respect to the Biot number Bi and also because the eigenvalues decay rapidly). For $N = 10$ and $K = 20$ the error bounds for the expected value and variance are 3.94×10^{-3} (corresponding to a relative error of 0.1%) and 8.32×10^{-4} (corresponding to a relative error of 20%), respectively, while the RB computational savings (including both Offline and Online effort) relative to the FE method is more than a factor of $\frac{1}{45}$. In the limit $J \to \infty$ of many $(\kappa, \overline{\mathrm{Bi}})$-queries, or $M \to \infty$ for better accuracy in the MC evaluations, the RB savings will approach $\frac{1}{200}$ – which reflects just the Online effort. The $(N = 10, K = 20)$-statistical results can be obtained Online in only 70 s (for a given $(\kappa, \overline{\mathrm{Bi}})$) on a Pentium IV 1.73 GHz; it would take roughly 4 h for the FE method to perform the same calculation.

We see that for $\kappa = 2.0$ and $\overline{\mathrm{Bi}} = 0.5$, the standard deviation of the integrated temperature is less than 2% of the expected integrated temperature; we can conclude that, for this value of κ and $\overline{\mathrm{Bi}}$, uncertainties in Bi are not too important to "device performance". However, for larger κ and small $\overline{\mathrm{Bi}}$ we expect more sensitivity: we find that for $\kappa = 10$ and $\overline{\mathrm{Bi}} = 0.1$ the standard deviation of the integrated temperature is now 6% of the expected integrated

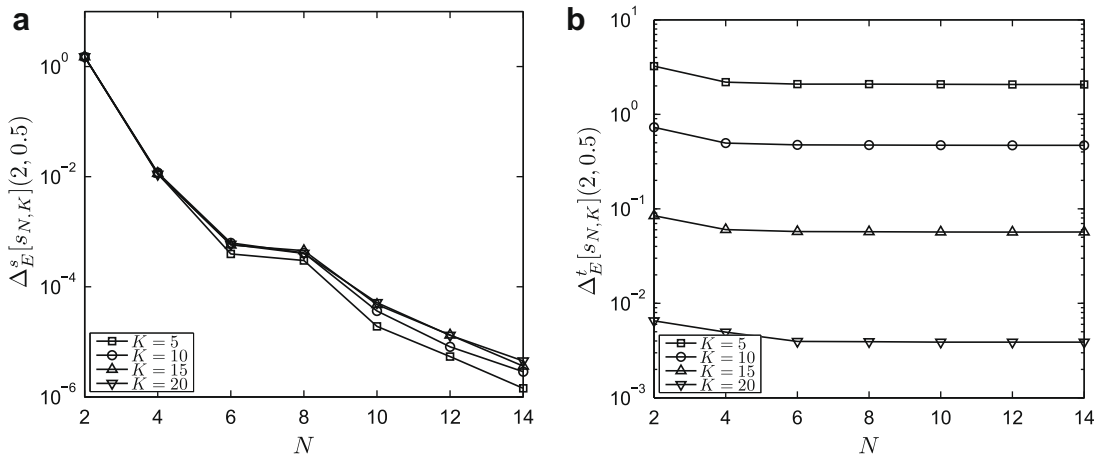


Fig. 8. (a) $\Delta_E^s[s_{N,K}(\mathscr{M}_K)]$ and (b) $\Delta_E^t[s_{N,K}(\mathscr{M}_K)]$ as functions of N and K; $\varrho = (2.0, 0.5)$.

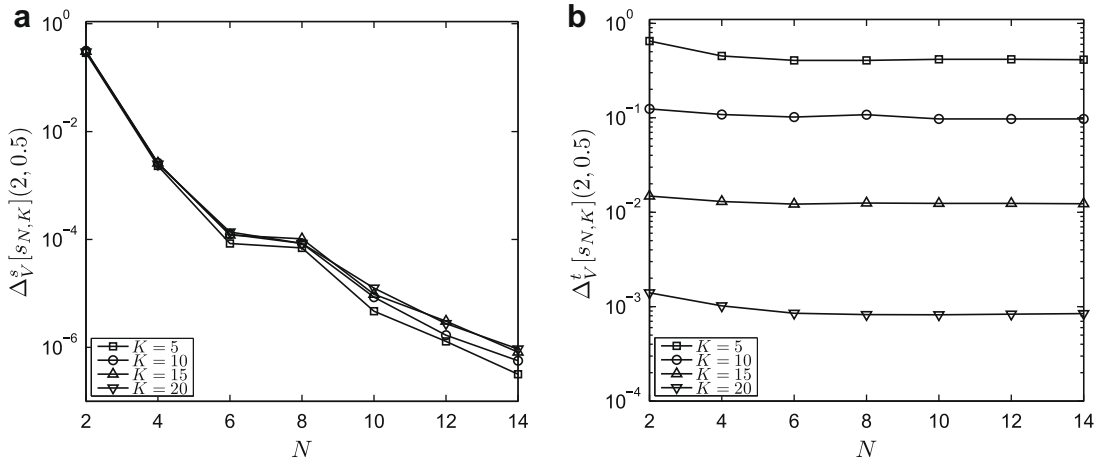
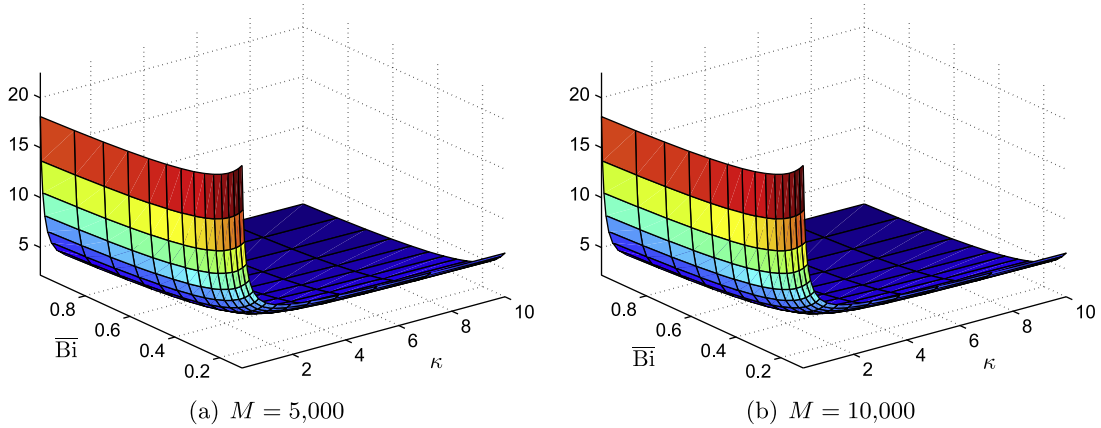


Fig. 9. (a) $\Delta_V^s[s_{N,K}(\mathscr{M}_K)]$ and (b) $\Delta_V^t[s_{N,K}(\mathscr{M}_K)]$ as functions of N and K; $\varrho = (2.0, 0.5)$.

temperature – and hence of engineering relevance. It is also possible to calculate the empirical cumulative distribution function to both assess the range and likelihood of "tails".

$(\kappa, \overline{\mathrm{Bi}})$-**variations:** we show in Fig. 10 the expected value of the integrated temperature at the bottom surface of the heat sink as a function of κ and $\overline{\mathrm{Bi}}$. The statistical outputs, which are obtained for $N = 10, K = 20$ and $J = 15 \times 15 = 225$ grid points in the parameter space, are plotted in Fig. 10a for $M = 5000$ and in Fig. 10b for $M = 10,000$. The maximum relative error in the expectation over the 225 parameter grid points is 9.4×10^{-4}. (The results in



(a) $M = 5{,}000$ (b) $M = 10{,}000$

Fig. 10. Expected value of the integrated temperature at the bottom surface of the fin as a function of κ and $\overline{\mathrm{Bi}}$ over $\Lambda^\varrho \equiv [0.1, 10] \times [0.1, 1]$.

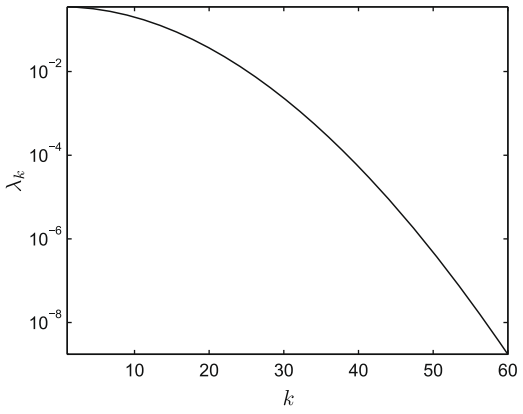


Fig. 11. Eigenvalues λ_k as functions of k for the correlation length $\delta = 0.2$.

Fig. 10a and b each require $J = 225$ evaluations of the empirical estimations for the expectation and the variance.)

Next, we consider another finite-rank covariance kernel $\mathbf{Cov_P}(\mathrm{Bi})(x,y)$ that coincides with the first $\boxed{\mathscr{K} = 60}$ terms in the KL expansion of $(\overline{\mathrm{Bi}}\Upsilon)^2 e^{-\frac{(x-y)^2}{\delta^2}}$ for a smaller correlation length $\boxed{\delta = 0.2}$. We present in Fig. 11 the eigenvalues λ_k as a function of k. We see that the eigenvalues decay at a slower rate than the previous case (shown in Fig. 2). We then obtain from (27) the requirement $\Upsilon_{\max} \equiv 0.074$; in our numerical examples we choose $\boxed{\Upsilon = \Upsilon_{\max} = 0.074}$. Fig. 12a shows four random realizations of the Biot number $\mathrm{Bi}(x,y)$ (these four random realizations vary more rapidly in space than the earlier instances of Fig. 3). We then pursue the greedy sampling procedure which yields $N_{\max} = 32$ for the same accuracy of 5×10^{-3} in the maximal error bound as in the case $\delta = 0.5$. It is not surprising from the Figs. 11 and 12a that the RB method needs larger $N_{\max}$ as the correlation length δ decreases.

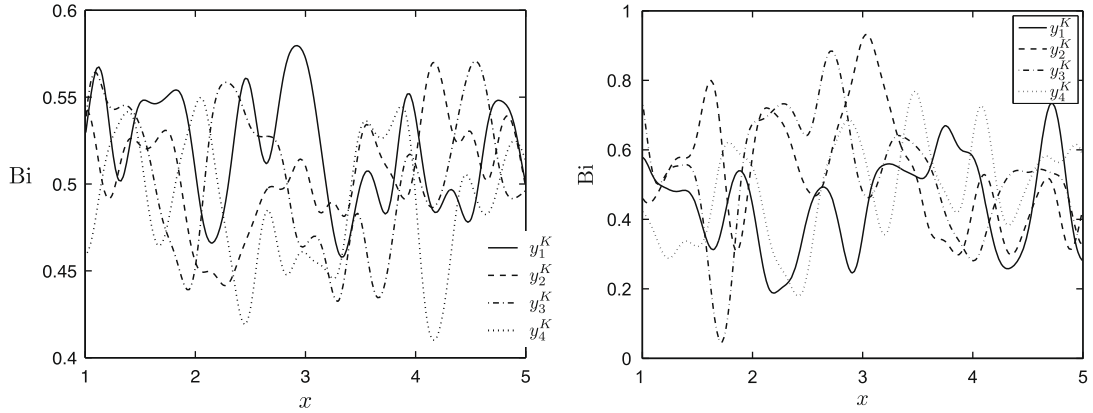


Fig. 12. Four realizations of the Biot number $x \to \mathrm{bi}(x;\overline{\mathrm{Bi}},y_i^K) - 1 \leqslant i \leqslant 4$ – for $\delta = 0.2$.

Table 3
Expected value $E_M[s_{N,K}(\mathscr{M}_K)]$ and error bound $\Delta_E^o[s_{N,K}(\mathscr{M}_K)]$ for different values of N and K with $\delta = 0.2, \Upsilon = 0.074$ and $\varrho = (2.0, 0.5)$.

N	$K = 15$		$K = 30$		$K = 45$		$K = 60$	
	$E_M[s_{N,K}]$	$\Delta_E^o[s_{N,K}]$	$E_M[s_{N,K}]$	$\Delta_E^o[s_{N,K}]$	$E_M[s_{N,K}]$	$\Delta_E^o[s_{N,K}]$	$E_M[s_{N,K}]$	$\Delta_E^o[s_{N,K}]$
5	3.6975	4.09×10^{0}	3.6970	4.80×10^{-1}	3.6960	1.55×10^{-2}	3.6960	2.68×10^{-3}
10	3.6975	4.03×10^{0}	3.6973	4.71×10^{-1}	3.6979	1.34×10^{-2}	3.6963	7.62×10^{-4}
15	3.6973	4.02×10^{0}	3.6978	4.70×10^{-1}	3.6970	1.32×10^{-2}	3.6977	6.05×10^{-4}
20	3.6980	4.00×10^{0}	3.6980	4.67×10^{-1}	3.6973	1.29×10^{-2}	3.6980	3.65×10^{-4}
25	3.6969	3.99×10^{0}	3.6977	4.66×10^{-1}	3.6972	1.28×10^{-2}	3.6981	3.36×10^{-4}
30	3.6968	3.99×10^{0}	3.6975	4.66×10^{-1}	3.6972	1.28×10^{-2}	3.6975	3.30×10^{-4}

Table 4
Variance $V_M[s_{N,K}(\mathscr{M}_K)]$ and error bound $\Delta_V^o[s_{N,K}(\mathscr{M}_K)]$ for different values of N and K with $\delta = 0.2, \Upsilon = 0.074$ and $\varrho = (2.0, 0.5)$.

N	$K = 15$		$K = 30$		$K = 45$		$K = 60$	
	$V_M[s_{N,K}]$	$\Delta_V^o[s_{N,K}]$	$V_M[s_{N,K}]$	$\Delta_V^o[s_{N,K}]$	$v_M[s_{N,K}]$	$\Delta_V^o[s_{N,K}]$	$V_M[s_{N,K}]$	$\Delta_V^o[s_{N,K}]$
5	0.0038	8.09×10^{-1}	0.0039	9.64×10^{-2}	0.0039	3.15×10^{-3}	0.0038	5.41×10^{-4}
10	0.0039	8.04×10^{-1}	0.0039	9.36×10^{-2}	0.0039	2.68×10^{-3}	0.0039	1.53×10^{-4}
15	0.0040	8.07×10^{-1}	0.0039	9.50×10^{-2}	0.0040	2.67×10^{-3}	0.0039	1.21×10^{-4}
20	0.0039	7.99×10^{-1}	0.0039	9.39×10^{-2}	0.0040	2.57×10^{-3}	0.0039	7.28×10^{-5}
25	0.0039	8.02×10^{-1}	0.0039	9.28×10^{-2}	0.0040	2.62×10^{-3}	0.0040	6.76×10^{-5}
30	0.0039	7.84×10^{-1}	0.0040	9.39×10^{-2}	0.0040	2.58×10^{-3}	0.0040	6.71×10^{-5}

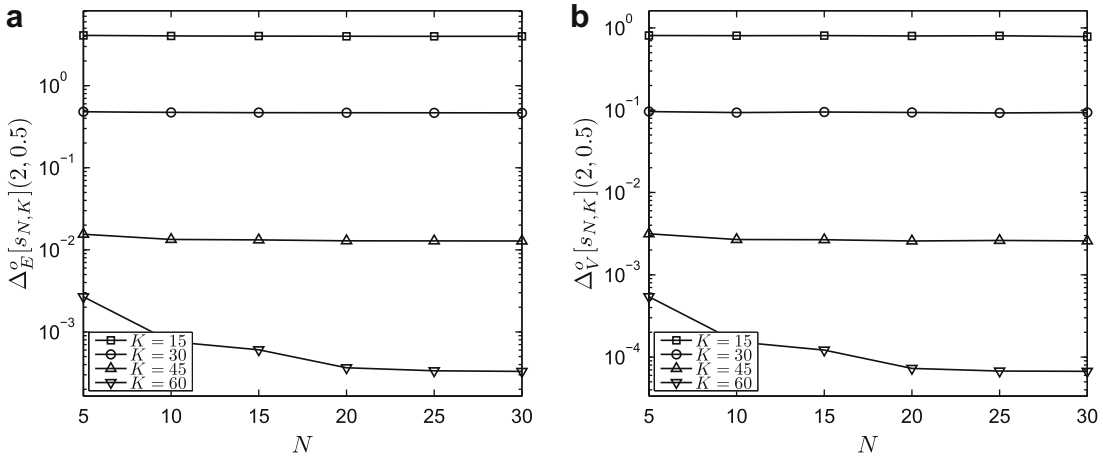


Fig. 13. (a) $\Delta_E^o[s_{N,K}(\mathscr{M}_K)]$ and (b) $\Delta_V^o[s_{N,K}(\mathscr{M}_K)]$ as functions of N and K with $\delta = 0.2, \Upsilon = 0.074$ and $\varrho = (2.0, 0.5)$.

Table 5
Expected value $E_M[s_{N,K}(\mathscr{M}_K)]$ and error bound $\Delta_E^o[s_{N,K}(\mathscr{M}_K)]$ for different values of N and K with $\delta = 0.2, \Upsilon = 0.3$ and $\varrho = (2.0, 0.5)$.

N	$K = 15$		$K = 30$		$K = 45$		$K = 60$	
	$E_M[s_{N,K}]$	$\Delta_E^o[s_{N,K}]$	$E_M[s_{N,K}]$	$\Delta_E^o[s_{N,K}]$	$E_M[s_{N,K}]$	$\Delta_E^o[s_{N,K}]$	$E_M[s_{N,K}]$	$\Delta_E^o[s_{N,K}]$
5	3.7230	1.82×10^{1}	3.7229	2.17×10^{0}	3.7215	1.02×10^{-1}	3.7239	4.25×10^{-2}
10	3.7312	1.70×10^{1}	3.7389	2.00×10^{0}	3.7273	6.00×10^{-2}	3.7299	5.65×10^{-3}
15	3.7341	1.67×10^{1}	3.7345	1.97×10^{0}	3.7287	5.59×10^{-2}	3.7311	2.53×10^{-3}
20	3.7327	1.66×10^{1}	3.7338	1.94×10^{0}	3.7328	5.40×10^{-2}	3.7351	1.08×10^{-3}
25	3.7323	1.65×10^{1}	3.7342	1.93×10^{0}	3.7350	5.33×10^{-2}	3.7364	6.73×10^{-4}
30	3.7322	1.64×10^{1}	3.7399	1.93×10^{0}	3.7385	5.30×10^{-2}	3.7370	5.20×10^{-4}

Table 6
Variance $V_M[s_{N,K}(\mathscr{M}_K)]$ and error bound $\Delta_V^o[s_{N,K}(\mathscr{M}_K)]$ for different values of N and K with $\delta = 0.2, \Upsilon = 0.3$ and $\varrho = (2.0, 0.5)$.

N	$K = 15$		$K = 30$		$K = 45$		$K = 60$	
	$V_M[s_{N,K}]$	$\Delta_V^o[s_{N,K}]$	$V_M[s_{N,K}]$	$\Delta_V^o[s_{N,K}]$	$v_M[s_{N,K}]$	$\Delta_V^o[s_{N,K}]$	$V_M[s_{N,K}]$	$\Delta_V^o[s_{N,K}]$
5	0.0721	1.54×10^{1}	0.0716	1.85×10^{0}	0.0744	8.90×10^{-2}	0.0718	3.72×10^{-2}
10	0.0738	1.46×10^{1}	0.0764	1.78×10^{0}	0.0743	5.25×10^{-2}	0.0738	5.03×10^{-3}
15	0.0717	1.43×10^{1}	0.0734	1.68×10^{0}	0.0735	4.81×10^{-2}	0.0744	2.25×10^{-3}
20	0.0705	1.41×10^{1}	0.0737	1.69×10^{0}	0.0725	4.61×10^{-2}	0.0728	9.48×10^{-4}
25	0.0699	1.38×10^{1}	0.0699	1.62×10^{0}	0.0723	4.56×10^{-2}	0.0732	5.83×10^{-4}
30	0.0755	1.44×10^{1}	0.0757	1.68×10^{0}	0.0722	4.64×10^{-2}	0.0723	4.43×10^{-4}

For $\kappa = 2.0$ and $\overline{\mathrm{Bi}} = 0.5$ again, we show in Table 3 the expected value and associated error bound for the integrated temperature at the bottom surface of the heat sink as a function of N and K.[7] Table 4 displays the corresponding variance and associated error bound. Fig. 13 shows the error bounds for the expected value and variance. We see that while the convergence pattern is similar to that of the previous case ($\delta = 0.5$), we need to use larger N and K to obtain the same accuracy for $\delta = 0.2$.

Nevertheless, the reduction in computational time is still quite significant: for $N = 10$ and $K = 45$ (for which the ratio $\Delta_E[s_{N,K}(\mathscr{M}_K)]/E_M[s_{N,K}(\mathscr{M}_K)]$ is **P**-a.s. less than 0.01 at $\varrho = (2.0, 0.5)$) the Online RB evaluation is still more than 50 times faster than the FE evaluation. Obviously, when the correlation length decreases further and further, the RB approach will no longer offer significant economies or may even become more expensive than the FE method; note however that, in three spatial dimensions, the RB method can "afford" a smaller correlation length since the FE truth will be considerably more expensive.

[7] The values for $\delta = 0.2$ are very similar to the values for $\delta = 0.5$ for the same reason that the variance is in general small: the output is relatively insensitive to Bi fluctuations.

Finally, in the latter case of a correlation length $\delta = 0.2$, we also consider $\boxed{\Upsilon = 0.3 > \Upsilon_{\max}}$ which yields a much larger domain Λ^y for the random parameter y^K. We note however that $\Upsilon = 0.3$ does not satisfy the well-posedness requirement (27). As a result, $\mathrm{bi}(\cdot; \overline{\mathrm{Bi}}, y^K)$ might be negative over the physical boundary $x \in [1, 5]$ for some y^K. In such case, we simply ignore all possible values of y^K at which $\min_{x\in[1,5]} \mathrm{bi}(x; \overline{\mathrm{Bi}}, y^K) \leqslant 0$, in the Offline stage as well as Online computation. (This should not introduce a significant bias in the SLLN limit providing only very few realizations are rejected, which is indeed the case here with a rejection rate of approximately 1/100). We show in Fig. 12b four random realizations of the Biot number $\mathrm{Bi}(x, Y)$; these four random realizations have much larger amplitudes than the instances of Fig. 12a. We then pursue the greedy sampling procedure for $K = 60$ (*a priori* determined) to construct the nested basis sets $X_N, 1 \leqslant N \leqslant N_{\max}$; we obtain $N_{\max} = 45$ – it is not surprising from Fig. 12b that the RB method needs larger $N_{\max}$ as Υ increases.

For $\kappa = 2.0$ and $\overline{\mathrm{Bi}} = 0.5$ again, we further present in Table 5 the expected value and associated error bound for the integrated temperature at the bottom surface of the heat sink as a function of N and K. The expected values are now slightly larger than those shown in Table 3. Table 6 displays the corresponding variance

and associated error bound. As expected, the variances are much larger than those shown in Table 4. More specifically, the standard deviation of the integrated temperature is approximately 7.2% of the expected integrated temperature, while the standard deviation of the integrated temperature is only 1.7% of the expected integrated temperature in the earlier results (see Tables 3 and 4).

For $\boxed{\kappa = 10}$ and $\boxed{\overline{\mathrm{Bi}} = 0.1}$, we find that the standard deviation of the integrated temperature is 14.6% of the expected integrated temperature, which consequently defines much more stringent conditions. The reduction in computational time is still significant: for $N = 10$ and $K = 45$ (for which the ratio $\Delta_E[s_{N,K}(\mathscr{M}_K)]/E_M[s_{N,K}(\mathscr{M}_K)]$ is **P**-a.s. less than 0.02 at $\varrho = (2.0, 0.5)$ and thus slightly larger than that of the previous case) the Online RB evaluation is more than 50 times faster than the FE evaluation. These results demonstrate that the RB error bound is inexpensive and accurate even for a significant variation in the random variables y^K.

4. Conclusions

In this article we have developed the theoretical framework (error bounds) for, and numerically demonstrated the attractiveness of, an RB approach for the rapid and reliable computation of expectations of linear functionals of variational solutions to a BVP with ω-x "separable" random parameter fields. The *a posteriori* error bounds certify the quality of the approximation and quantify the effects of both the FE $\rightarrow$ RB reduction for the BVP and the KL truncation in the random field expansion. The method also permits the study of the parametric dependence of the outputs with respect to other (deterministic) parameters entering the problem.

Future developments may include:

(a) test problems in which the random input field multiplies the solution field not only on the boundary but also over the entire domain (*e.g.* random diffusivity coefficient κ),
(b) more general variates (and sampling procedures) in the KL expansion of the input field,
(c) inputs developed with expansions other than KL (not necessarily decoupling $\mathscr{D}$ and Ω, and thus requiring empirical interpolation [4,17]),
(d) more general statistical outputs (that remain sufficiently smooth functionals of the random solution field – continuous in $L^p_{\mathbf{P}}(\Omega, H^1(\mathscr{D}))$), and
(e) application of the RB approach to Ω-weak/$\mathscr{D}$-weak collocation formulations [2,33].

But from our first results, it is arguably already interesting to apply an RB approach within many of the Ω-strong/$\mathscr{D}$-weak formulations in view of the simplicity of the implementation, the considerable reduction in computational time, and the availability of rigorous error bounds (suitably generalized, in particular as regards the contribution of the KL truncation and associated continuity constants).

We end this paper by pointing out that the RB methods and associated *a posteriori* error estimation have been developed for several classes of parametrized PDEs including linear coercive/non-coercive elliptic problems [4,7,21,42,43,46], linear elasticity [20], eigenvalue problems [27], linear parabolic problems [18,19], Boltzmann equations [39], nonlinear elliptic and parabolic problems [17], and incompressible Navier–Stokes equations [31,32,52]. It appears that the extension to other classes of SPDEs beyond the particular linear elliptic SPDE discussed in this paper can be achieved by combining the current RB approach with those of the previous work. We consider to pursue this line of development in future work.

Acknowledgements

We thank Dr. Gianluigi Rozza of EPFL for helpful discussions. This work was supported by AFOSR Grant FA9550-07-1-0425, the Singapore-MIT Alliance, and the Chaire d'excellence ACI "Application des polynômes de chaosà la modélisation des incertitudes en mécaniques des fluides".

References

[1] B.O. Almroth, P. Stern, F.A. Brogan, Automatic choice of global shape functions in structural analysis, AIAA J. 16 (1978) 525–528.
[2] I. Babuška, F. Nobile, R. Tempone, A stochastic collocation method for elliptic partial differential equations with random input data, SIAM J. Numer. Anal. 45 (3) (2007) 1005–1034.
[3] I. Babuška, R. Tempone, G. Zouraris, Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation, Comput. Methods Appl. Mech. Engrg. 194 (2005) 1251–1294.
[4] M. Barrault, Y. Maday, N.C. Nguyen, A.T. Patera, An empirical interpolation method; application to efficient reduced-basis discretization of partial differential equations, C.R. Acad. Sci. Paris 339 (2004) 667–672.
[5] X. Blanc, C. Le Bris, P.L. Lions, Une variante de la théorie de l'homogénéisation stochastique des opérateurs elliptiques, C.R. Acad. Sci. Paris 343 (11–12) (2006) 717–724.
[6] S. Boyaval, Reduced-basis approach for homogenization beyond the periodic setting, SIAM Multiscale Model. Simulat. 7 (1) (2008) 466–494.
[7] Y. Chen, J. Hesthaven, Y. Maday, J. Rodríguez, A monotonic evaluation of lower bounds for inf–sup stability constants in the frame of reduced basis approximations, C. R. Acad. Sci. Paris, Ser. I 346 (23–24) (2008) 1295–1300.
[8] D.B. Chung, M.A. Gutiérrez, J.J.C. Remmers, R. de Borst, Stochastic finite element modelling of fibre-metal laminates, in: 45th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics& Materials Conference, April 2004.
[9] M.K. Deb, I.M. Babuška, J.T. Oden, Solution of stochastic partial differential equations using Galerkin finite element techniques, Comput. Methods Appl. Mech. Engrg. 190 (48) (2001) 6359–6372.
[10] B.J. Debusschere, H.N. Najm, P.P. Pebay, O.M. Knio, R.G. Ghanem, O.P.L. Maitre, Numerical challenges in the use of polynomial chaos representations for stochastic processes, SIAM J. Sci. Comput. 26 (2) (2004) 698–719.
[11] A. Doostan, R.G. Ghanem, J. Red-Horse, Stochastic model reduction for chaos representations, Comput. Methods Appl. Mech. Engrg. 196 (37–40) (2007) 3951–3966.
[12] J.P. Fink, W.C. Rheinboldt, On the error behavior of the reduced basis technique for nonlinear finite element approximations, Z. Angew. Math. Mech. 63 (1) (1983) 21–28.
[13] P. Frauenfelder, C. Schwab, R.A. Todor, Finite elements for elliptic problems with stochastic coefficients, Comput. Methods Appl. Mech. Engrg. 194 (2005) 205–228.
[14] R. Ghanem, S. Dham, Stochastic finite element analysis for multiphase flow in heterogeneous porous media, Transport Porous Med. 32 (239) (1998).
[15] R. Ghanem, G. Saad, A. Doostan, Efficient solution of stochastic systems: application to the embankment dam problem, Struct. Safety 29 (3) (2007) 138–251.
[16] R.G. Ghanem, P.D. Spanos, Stochastic Finite Elements: A Spectral Approach, second ed., Springer-Verlag, New York, 1991 (Revised First Dover Edition, 2003).
[17] M.A. Grepl, Y. Maday, N.C. Nguyen, A.T. Patera, Efficient approximation for reduced-basis treatment of nonaffine and nonlinear partial differential equations, Math. Model. Numer. Anal. (M2AN) 41 (2007) 575–605.
[18] M.A. Grepl, A.T. Patera, A posteriori error bounds for reduced-basis approximations of parametrized parabolic partial differential equations, Math. Model. Numer. Anal. (M2AN) 39 (2005) 157–181.
[19] B. Haasdonk, M. Ohlberger, Reduced basis method for finite volume approximations of parametrized linear evolution equations, Math. Model. Numer. Anal. (M2AN) 42 (2008) 277–302.
[20] D.B.P. Huynh, A.T. Patera, Reduced-basis approximation and a posteriori error estimation for stress intensity factors, Int. J. Numer. Methods Engrg. 72 (10) (2007) 1219–1259. doi:10.1002/nme.2090.
[21] D.B.P. Huynh, G. Rozza, S. Sen, A.T. Patera, A successive constraint linear optimization method for lower bounds of parametric coercivity inf–sup stability constants, C.R. Acad. Sci. Paris Anal. Numer. 345 (8) (2007) 473–478, doi:10.1016/j.crma.2007.09.019).
[22] J.H. Lienhard IV, J.H. Lienhard V, A Heat Transfer Textbook, Phlogiston Press, Cambridge, Mass, 2002.
[23] M. Jardak, R. Ghanem, Spectral stochastic homogenization for of divergence-type PDEs, Comput. Methods Appl. Mech. Engrg. 193 (2004) 429–447.
[24] B. Jin, J. Zou, Inversion of Robin coefficient by a spectral stochastic finite element approach, J. Comput. Phys. 227 (6) (2008) 3282–3306.
[25] A. Keese, H.G. Matthies, Hierarchical parallelisation for the solution of stochastic finite element equations, Comput. Struct. 83 (14) (2005) 1033–1047.
[26] M. Loève, Probability Theory, vols. I–II, Springer, New York, 1978.

[27] L. Machiels, Y. Maday, I.B. Oliveira, A.T. Patera, D.V. Rovas, Output bounds for reduced-basis approximations of symmetric positive definite eigenvalue problems, C.R. Acad. Sci. Paris Série I 331 (2) (2000) 153–158.
[28] L. Mathelin, M.Y. Hussaini, T.A. Zang, Stochastic approaches to uncertainty quantification in CFD simulations, Numer. Algorithms 38 (2005) 209–236.
[29] H.G. Matthies, A. Keese, Galerkin methods for linear and nonlinear elliptic stochastic pdes, Comput. Methods Appl. Mech. Engrg. 194 (2005) 1295–1331.
[30] P.B. Nair, A.J. Keane, Stochastic reduced basis methods, AIAA J. 40 (8) (2002) 1653–1664.
[31] N.C. Nguyen, G. Rozza, D.B.P. Huynh, A. T Patera, Reduced basis approximation and a posteriori error estimation for parametrized parabolic PDEs; application to real-time Bayesian parameter estimation, in: Biegler, Biros, Ghattas, Heinkenschloss, Keyes, Mallick, Tenorio, van Bloemen Waanders, Willcox (Eds.), Computational Methods for Large Scale Inverse Problems and Uncertainty Quantification, John Wiley& Sons, UK, 2009.
[32] N.C. Nguyen, K. Veroy, A.T. Patera, Certified real-time solution of parametrized partial differential equations, in: S. Yip (Ed.), Handbook of Materials Modeling, Springer, 2005, pp. 1523–1558.
[33] F. Nobile, R. Tempone, C. Webster, A sparse grid stochastic collocation method for partial differential equations with random input data, SIAM J. Numer. Anal. 46 (2008) 2309–2345.
[34] A.K. Noor, J.M. Peters, Reduced basis technique for nonlinear analysis of structures, AIAA J. 18 (4) (1980) 455–462.
[35] A. Nouy, A generalized spectral decomposition technique to solve a class of linear stochastic partial differential equations, Comput. Methods Appl. Mech. Engrg. 196 (45–48) (2007) 4521–4537.
[36] A. Nouy, Generalized spectral decomposition method for solving stochastic finite element equations: invariant subspace problem and dedicated algorithms, Comput. Methods Appl. Mech. Engrg 197 (51–52) (2008) 4718–4736.
[37] B. Oksendal, Stochastic Differential Equations, sixth ed., An Introduction with Applications, Springer-Verlag, 2003.
[38] M. Ostoja-Starzewski, X. Wang, Stochastic finite elements as a bridge between random material microstructure and global response, Comput. Methods Appl. Mech. Engrg. 168 (1–4) (1999) 35–49.
[39] A.T. Patera, E.M. Rønquist, Reduced basis approximations and a posteriori error estimation for a Boltzmann model, Comput. Methods Appl. Mech. Engrg. 196 (29) (2007) 2925–2942.
[40] M.F. Pellissetti, R.G. Ghanem, Iterative solution of systems of linear equations arising in the context of stochastic finite elements, Adv. Engrg. Software 31 (2000) 607–616.
[41] T.A. Porsching, Estimation of the error in the reduced basis method solution of nonlinear equations, Math. Comput. 45 (172) (1985) 487–496.
[42] C. Prud'homme, D. Rovas, K. Veroy, Y. Maday, A.T. Patera, G. Turinici, Reliable real-time solution of parametrized partial differential equations: reduced-basis output bounds methods, J. Fluids Engrg. 124 (1) (2002) 70–80.
[43] G. Rozza, D.B.P. Huynh, A.T. Patera, Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations; application to transport and continuum mechanics, Arch. Comput. Methods Engrg., in press, doi:10.1007/s11831-008-9019-9.
[44] S.K. Sachdeva, P.B. Nair, A.J. Keane, Hybridization of stochastic reduced basis methods with polynomial chaos expansions, Probabilist. Engrg. Mech. 21 (2006) 182–192.
[45] C. Schwab, R.A. Todor, Karhunen–Loève approximation of random fields by generalized fast multipole methods, J. Comput. Phys. 217 (1) (2006) 100–122.
[46] S. Sen, K. Veroy, D.B.P. Huynh, S. Deparis, N.C. Nguyen, A.T. Patera, Natural norm a posteriori error estimators for reduced basis approximations, J. Comput. Phys. 217 (1) (2006) 37–62.
[47] C. Soize, R. Ghanem, Physical systems with random uncertainties: chaos representations with arbitrary probability measure, SIAM J. Sci. Comput. 26 (2) (2004) 395–410.
[48] B. Sudret, M. Berveiller, M. Lemaire, A stochastic finite element procedure for moment and reliability analysis, Rev. Eur. Méca. Numer. 15 (7–8) (2006) 825–866.
[49] P. Surya Mohan, Prasanth B. Nair, Andy J. Keane, Multi-element stochastic reduced basis methods, Comput. Methods Appl. Mech. Engrg. 197 (17–18) (2008) 1495–1506.
[50] R.A. Todor, C. Schwab, Convergence rates for sparse chaos approximations of elliptic problems with stochastic coefficients, IMA J. Numer. Anal. 27 (2007) 232–261.
[51] G. Venkiteswaran, M. Junk, Quasi-Monte-Carlo algorithms for diffusion equations in high dimensions, Math. Comput. Simulat. 68 (1) (2005) 23–41.
[52] K. Veroy, A.T. Patera, Certified real-time solution of the parametrized steady incompressible Navier–Stokes equations; rigorous reduced-basis a posteriori error bounds, Int. J. Numer. Methods Fluids 47 (2005) 773–788.
[53] X. Wan, G.E. Karniadakis, An adaptive multi-element generalized polynomial chaos method for stochastic differential equations, J. Comput. Phys. 209 (2005) 617–642.
[54] X. Wan, G.E. Karniadakis, Error control in multi-element generalized polynomial chaos method for elliptic problems with random coefficients, Commun. Comput. Phys. 5 (2–4) (2009) 793–820 (Available online 1 August 2008).
[55] N. Wiener, The homogeneous chaos, Am. J. Math. 60 (1938) 897–936.
[56] D. Xiu, G.E. Karniadakis, The Wiener–Askey polynomial chaos for stochastic differential equations, SIAM J. Sci. Comput. 24 (2002) 619–644.
[57] D. Xiu, J.S. Hesthaven, High-order collocation methods for differential equations with random inputs, SIAM J. Sci. Comput. 27 (2005) 1118–1139.
[58] X. Frank Xu, A multiscale stochastic finite element method on elliptic problems involving uncertainties, Comput. Methods Appl. Mech. Engrg. 196 (25–2) (2007) 2723–2736.