SIAM J. NUMER. ANAL.
Vol. 51, No. 6, pp. 3163–3185

A WEIGHTED REDUCED BASIS METHOD FOR ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS WITH RANDOM INPUT DATA*

PENG CHEN†, ALFIO QUARTERONI‡, AND GIANLUIGI ROZZA§

Abstract. In this work we propose and analyze a weighted reduced basis method to solve elliptic partial differential equations (PDEs) with random input data. The PDEs are first transformed into a weighted parametric elliptic problem depending on a finite number of parameters. Distinctive importance of the solution at different values of the parameters is taken into account by assigning different weights to the samples in the greedy sampling procedure. A priori convergence analysis is carried out by constructive approximation of the exact solution with respect to the weighted parameters. Numerical examples are provided for the assessment of the advantages of the proposed method over the reduced basis method and the stochastic collocation method in both univariate and multivariate stochastic problems.

Key words. weighted reduced basis method, stochastic partial differential equation, uncertainty quantification, stochastic collocation method, Kolmogorov N-width, exponential convergence

AMS subject classifications. 35J15, 65C30, 65N35, 60H15, 60H35

DOI. 10.1137/130905253

1. Introduction.

When modeling a complex physical system, uncertainties inevitably arise from various sources, e.g., computational geometries, physical parameters, external forces, and initial or boundary conditions, and may significantly impact the computational results. When these uncertainties are incorporated into the underlying physical system, we are facing stochastic problems or uncertainty quantification. Various computational methods have been developed depending on the structure of the stochastic problem, including perturbation, Monte Carlo, stochastic Galerkin, stochastic collocation, reduced basis, and generalized spectral decomposition methods [21, 40, 1, 33, 7].

The perturbation method [25] based on Taylor expansion was developed for the random functions with only small fluctuations around a deterministic expectation. This method is applicable only when dealing with small uncertainties and suffers from inevitable errors and an extremely complicated structure for high order expansions. The most commonly used "brute-force" Monte Carlo method [20] as well as its multiple versions, e.g., quasi Monte Carlo [30] and multilevel Monte Carlo [23], converge very slowly and become prohibitive for achieving accurate results.

The stochastic Galerkin method, originated from spectral expansion of the random functions on some polynomial chaos, for instance, Hermite polynomials of independent

*Received by the editors January 8, 2013; accepted for publication (in revised form) September 18, 2013; published electronically November 19, 2013. This work is partially supported by the Swiss National Science Foundation under grant N.200021_141034.
http://www.siam.org/journals/sinum/51-6/90525.html

†Ecole Polytechnique Fédérale de Lausanne, Modelling and Scientific Computing, CMCS MATHICSE, EPFL, CH-1015 Lausanne, Switzerland (peng.chen@epfl.ch).

‡Ecole Polytechnique Fédérale de Lausanne, Modelling and Scientific Computing, CMCS MATHICSE, EPFL, CH-1015 Lausanne, Switzerland, and Dipartimento di Matematica F. Brioschi, Modellistica e Calcolo Scientifico, MOX, Politecnico di Milano, I-20133 Milano, Italy (alfio.quarteroni@epfl.ch).

§SISSA MathLab, International School for Advanced Studies, 34136 Trieste, Italy (gianluigi.rozza@sissa.it).

random variables, applies the Galerkin approaches to approximate the solution in both stochastic and deterministic space [21, 2]. It enjoys fast convergence provided the solution is regular [14, 13]. However, it yields a very large algebraic system, leading to the challenge of designing efficient solvers with appropriate preconditioners [19].

The stochastic collocation method was developed from the nonintrusive deterministic collocation method [36, 39, 1]. In principle, it employs multivariate polynomial interpolations for the integral in the variational formulation of the stochastic system with respect to probability space rather than the Galerkin approximation in the spectral polynomial space. Due to the heavy computation of a deterministic system at each collocation point in high-dimensional space, isotropic or anisotropic sparse grids with suitable cubature rules [31, 32] were analyzed and applied to reduce the computation load. This method is preferred for more practical applications because it features the advantages of both direct computation as a Monte Carlo method and fast convergence as a stochastic Galerkin method [3].

In principle, to solve a stochastic problem we need to solve one deterministic problem at many different realizations of the random inputs in order to evaluate the quantity of interest depending on the stochastic solutions. However, the solutions are "not far from" each other in practice. Therefore, instead of projecting the solutions on some prescribed bases, such as polynomial chaos for the stochastic Galerkin method [40], we can project the solution on some space generated by a few precomputed solutions, which leads to the development of reduced basis method. The reduced basis method has been proposed to solve primarily parametric systems [37, 34] and has been applied to stochastic problems lately [7, 6, 10]. In the latter context, it regards the random variables as parameters and selects the most representative points in the parameter space by greedy sampling based on a posteriori error estimation. The essential idea for the deterministic and stochastic reduced basis method is to separate the whole procedure into an offline stage and an online stage. During the former, the large computational ingredients are computed and stored once and for all, including sampling parameters, assembling matrices and vectors, solving and collecting snapshots of solutions, etc. In the online stage, only the parameter-related elements are left to be computed and a small Galerkin approximation problem needs to be solved [34]. Both the reduced basis method and the stochastic collocation method use precomputed solutions as approximation/construction bases. However, the former employs a posteriori error estimation for the construction and thus is more efficient provided that a posteriori error estimation is easy to compute. Comparison of the convergence property as well as the computational cost for offline construction and online evaluation between the reduced basis method and the stochastic collocation method was investigated in [10].

To our knowledge, the reduced basis method is currently used only for stochastic problems with uniformly distributed random inputs or parameter space with Lebesgue measure [6, 10]. In order to deal with more general stochastic problems with other distributed random inputs, we propose and analyze a new version of reduced basis method and name it the "weighted reduced basis method." The basic idea is to suitably assign a larger weight to samples that are more important or have a higher probability of occuring than the others according to either the probability distribution function or some other available weight function depending on the specific application at hand. The benefit is to lighten the reduced space construction using a smaller number of bases without lowering the numerical accuracy.

A priori convergence analysis for the reduced basis method by greedy algorithm has been carried out in previous works [28, 8, 5, 26] under various assumptions. More

specifically, an exponential convergence rate for a single-parameter elliptic PDE was obtained in [28] by exploring an eigenvalue problem; an algebraic or exponential convergence rate for greedy algorithm in multidimensional problem was achieved implicitly depending on the convergence rate of Kolmogorov N-width in [8] and improved in [5]; an exponential convergence rate was also recently obtained in [26] through direct expansion of the solution on a series of invertible elliptic operators. In this work, we carry out a priori convergence analysis of our weighted reduced basis method based on constructive spectral approximation for analytic functions, which is different from [28, 8, 5, 26].

The paper is organized as follows. An elliptic PDE with random input data is set up with appropriate assumptions on both the random coefficient and the forcing term in section 2. Section 3 is devoted to the development of the weighted reduced basis method consisting of a greedy algorithm, an a posteriori error estimate, as well as offline-online computational decomposition, which is followed by regularity analysis and a priori convergence analysis in section 4. Numerical examples for both the one-dimensional problem and the multiple-dimensional problem are presented as verification of the efficiency and convergence properties in section 5. Some brief concluding remarks are made in the last section 6.

2. Problem setting. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, where Ω is a set of outcomes $\omega \in \Omega$, $\mathcal{F}$ is σ-algebra of events, and $P : \mathcal{F} \to [0,1]$ with $P(\Omega) = 1$ assigns probability to the events. Let D be a convex, open, and bounded physical domain in $\mathbb{R}^d$ $(d = 1, 2, 3)$ with Lipschitz continuous boundary ∂D. We consider the following stochastic elliptic problem: find $u : \bar{D} \times \Omega \to \mathbb{R}$ such that it holds almost surely that

$$\begin{aligned} -\nabla \cdot (a(\cdot,\omega)\nabla u(\cdot,\omega)) &= f(\cdot,\omega) \quad \text{in } D, \\ u(\cdot,\omega) &= 0 \quad \text{on } \partial D, \end{aligned} \tag{2.1}$$

where $f : D \times \Omega \to \mathbb{R}$ is a random force term and $a : D \times \Omega \to \mathbb{R}$ is a random coefficient; a homogeneous Dirichlet boundary condition is prescribed on the whole boundary ∂D for simplicity. We consider the following assumptions for the random functions $f(\cdot,\omega)$ and $a(\cdot,\omega)$.

Assumption 2.1. The random forcing term $f(\cdot,\omega)$ is square integrable with respect to P, i.e.,

$$||f||^2_{L^2_P(\Omega)\otimes L^2(D)} := \int_{\Omega\times D} f^2(x,\omega)dxdP(\omega) < \infty. \tag{2.2}$$

Assumption 2.2. The random coefficient $a(\cdot,\omega)$ is assumed to be uniformly bounded from below and from above, i.e., there exist constants $0 < a_{min} < a_{max} < \infty$ such that

$$P(\omega \in \Omega : a_{min} < a(x,\omega) < a_{max} \quad \forall x \in \bar{D}) = 1. \tag{2.3}$$

We introduce the Hilbert space $V := L^2_P(\Omega) \otimes H^1_0(D)$ and equip it with the following norm:

$$||v||_V = ||v||_{L^2_P(\Omega)\otimes H^1_0(D)} = \left(\int_{\Omega\times D} |\nabla v|^2 dxdP\right)^{1/2} < \infty. \tag{2.4}$$

The weak formulation of problem (2.1) is stated as follows: find $u \in V$ such that

$$\int_{\Omega\times D} a\nabla u \cdot \nabla v dxdP = \int_{\Omega\times D} fv dxdP \quad \forall v \in V. \tag{2.5}$$

The existence of a unique solution to problem (2.5) is guaranteed by the Lax–Milgram theorem [36] under Assumptions 2.1 and 2.2 and the stability inequality holds for the solution straightforwardly,

$$||u||_V \leq \frac{C_P}{a_{min}} ||f||_{L^2_P(\Omega)\otimes L^2(D)}, \tag{2.6}$$

where the constant C_P comes from the Poincaré inequality $||v||_{L^2(D)} \leq C_P ||\nabla v||_{L^2(D)}$ $\forall v \in H^1_0(D)$.

The uncertainty of the random functions $a(\cdot,\omega)$ and $f(\cdot,\omega)$, in many practical applications, can be approximately projected to a series of finite dimensional random variables via statistical techniques. For instance, finite linear regression models are widely used to approximate various random fields [15]; under the assumption that the second moment of $a(\cdot,\omega)$ exists, we can apply Karhunen–Loève expansion [38] to the covariance kernel and truncate it up to a finite number of linear terms, etc. For this consideration, we make a further assumption to the random functions $a(\cdot,\omega)$ and $f(\cdot,\omega)$ as follows.

Assumption 2.3. The random coefficient $a(\cdot,\omega)$ and forcing term $f(\cdot,\omega)$ are linear combinations of a number of random variables $Y(\omega) = (Y_1(\omega),\dots,Y_K(\omega)) : \Omega \to \mathbb{R}^K$ as follows:

$$a(x,Y) = a_0(x) + \sum_{n=1}^{K} a_k(x) Y_k(\omega) \quad \text{and} \quad f(x,Y) = f_0(x) + \sum_{n=1}^{K} f_k(x) Y_k(\omega), \tag{2.7}$$

where $a_k \in L^\infty(D)$ and $f_k \in L^2(D)$ for $0 \leq k \leq K$. More specifically, $\{Y_k\}_{k=1}^K$ are real valued random variables with joint probability density function $\rho(y)$, being $y = Y(\omega) \in \mathbb{R}$. By denoting $\Gamma_k \equiv Y_k(\Omega), k = 1,\dots,K$, and $\Gamma = \Pi_{k=1}^K \Gamma_k$, we can also view y as a weighted parameter in the parametric domain Γ endowed with the measure $\rho(y)dy$. In particular, we assume that the random variables y are bounded in a continuous domain Γ for the sake of convergence analysis.

Remark 2.4. When the random variables $Y^a_k, 1 \leq k \leq K_a$, for a and $Y^f_k, 1 \leq k \leq K_f$, for f are not the same, we collect them as $Y = (Y^a_1,\dots,Y^a_{K_a},Y^f_1,\dots,Y^f_{K_f})$ and reorder them as $(Y_1,\dots,Y_K)$ with $K = K_a + K_f$.

Remark 2.5. In the more general case that the random function $a(x,Y)$ does not depend on Y linearly, for instance,

$$a(x,Y) = a_0(x) + \exp\left(\sum_{n=1}^{K} a_k(x) Y_k(\omega)\right), \tag{2.8}$$

one can employ the empirical interpolation method [4, 12] to approximate (2.8) with finite affine terms in the form

$$a(x,Y) \approx a_0(x) + \sum_{n=1}^{K'} a'_{k'}(x) \Theta_{k'}(Y(\omega)), \tag{2.9}$$

where $\Theta_{k'}(\cdot), 1 \leq k' \leq K'$, are functions of Y and can be transformed to random variables $Z_{k'} = \Theta_{k'}(Y(\omega)), 1 \leq k' \leq K$, resulting in a new random vector $Z = (Z_1,\dots,Z_{K'})$ and $a(x,Z)$ still satisfies Assumption 2.3.

Under the above assumptions, the weighted parametric weak formulation of the stochastic elliptic problem reads as follows: find $u(y) \in H^1_0(D)$ such that the following equation holds for all $y \in \Gamma$:

$$A(u,v;y) = F(v;y) \quad \forall v \in H_0^1(D), \tag{2.10}$$

where $A(\cdot,\cdot;y)$ and $F(\cdot;y)$ are parametrized bilinear and linear forms featuring the expansion

$$A(u,v;y) = A_0(u,v) + \sum_{k=1}^{K} A_k(u,v)y_k \quad \text{and} \quad F(v;y) = (f_0,v) + \sum_{k=1}^{K} (f_k,v)y_k \tag{2.11}$$

with the deterministic bilinear forms $A_k(u,v)$ given by $A_k(u,v) := (a_k\nabla u, \nabla v), k = 0,1,\ldots,K$. Because of assumption (2.3) the bilinear form is coercive and continuous, and thus the existence of a unique parametric solution $u(y) \in H_0^1(D)\ \forall y \in \Gamma$ to problem (2.10) is guaranteed by the Lax–Milgram theorem [36]. More often, we are interested in a linear functional $s(u;y)$ as output, e.g., $s(u;y) = F(u;y)$, as well as its statistics, e.g., the expectation $\mathbb{E}[s]$, which is defined as

$$\mathbb{E}[s] = \int_\Gamma s(u;y)\rho(y)dy. \tag{2.12}$$

Given any approximation space $X^{\mathcal{N}} \subset H_0^1(D)$ (e.g., finite element space) of dimension $\mathcal{N}$, we approximate the solution of (2.10) by solving the following problem: given any $y \in \Gamma$, find $u^{\mathcal{N}} \in X^{\mathcal{N}}$ such that

$$A(u,v;y) = F(v;y) \quad \forall v \in X^{\mathcal{N}}. \tag{2.13}$$

Consequently, the quantity of interest $s(u;y)$ and its statistics, e.g., $\mathbb{E}[s]$, can be approximated by $s(u;y) \approx s^{\mathcal{N}}(y) := s(u^{\mathcal{N}};y)$ and $\mathbb{E}[s] \approx \mathbb{E}[s^{\mathcal{N}}]$, respectively.

3. Weighted reduced basis method. The basic idea behind the weighted reduced basis method is to assign different weights in the construction of reduced basis space at different values of parameter $y \in \Gamma$ according to a prescribed weight function $w(y)$. The objective is that when the parameter y has distinctive weight $w(y)$ at different values $y \in \Gamma$, e.g., stochastic problems with random inputs obeying probability distribution far from uniform type, the weighted approach can considerably attenuate the computational effort for large-scale computational problems. The general paradigm of the weighted reduced basis method is formulated by following closely the reduced basis method in [34, 37, 10].

Given a training set of parameter samples $\Xi_{train} \subset \Gamma$ as well as a prescribed maximum dimension $N_{max} \ll \mathcal{N}$, we build the N-dimensional (Lagrange) reduced basis space $X_N^{\mathcal{N}} \subset X^{\mathcal{N}}$ for $N = 1,\ldots,N_{max}$ in a hierarchical way by taking into account the weight of the parameter at different values until we satisfy a certain tolerance requirement. The reduced basis space $X_N^{\mathcal{N}}$ is spanned by the "snapshots" (solutions $u^{\mathcal{N}} \in X^{\mathcal{N}}$ of problem (2.13)) based on suitably chosen samples $S_N = \{y^1,\ldots,y^N\}$ from the training set Ξ_{train}

$$X_N^{\mathcal{N}} = \text{span}\{u^{\mathcal{N}}(y^n), 1 \le n \le N\}. \tag{3.1}$$

Note that $X_1^{\mathcal{N}} \subset X_2^{\mathcal{N}} \subset \cdots \subset X_{N_{max}}^{\mathcal{N}}$. In order to evaluate $s(u;y)$ at any new parameter $y \in \Gamma$, we first seek the solution $u_N^{\mathcal{N}} \in X_N^{\mathcal{N}} \subset X^{\mathcal{N}}$ in the reduced basis space $X_N^{\mathcal{N}}$ by solving a reduced system

$$A(u_N^{\mathcal{N}},v;y) = F(v;y) \quad \forall v \in X_N^{\mathcal{N}} \tag{3.2}$$

and then approximate $s(u;y)$ by $s(u_N^{\mathcal{N}};y)$. Moreover, we can also compute the statistics of the output, e.g., expectation $\mathbb{E}[s_N^{\mathcal{N}}]$, by the numerical quadrature formula (Gauss or Clenshaw–Curtis quadrature [32]),

$$\mathbb{E}[s_N^{\mathcal{N}}] \approx \sum_{m=1}^{M} s(u_N^{\mathcal{N}};y^m)w(y^m), \tag{3.3}$$

where y^m and $w(y^m), m = 1,\dots,M$, are the K-dimensional quadrature abscissas and weights with respect to the probability density function, which can be chosen based on different schemes, e.g., full tensor product quadrature or sparse grid quadrature [32]. Note that the weights $w(y^m), m = 1,\dots,M$, may be distinct to each other depending on both the quadrature formula (e.g., Clenshaw–Curtis or Gaussian type) and the probability density function, so that the solution $u_N^{\mathcal{N}}(y^m)$ is expected to be more accurate where $w(y^m)$ is significantly larger than the other realization of the parameter $y \in \Gamma$.

Accurate computation of the solution $u_N^{\mathcal{N}}$ and the output $s_N^{\mathcal{N}}$ depends crucially on the construction of the reduced basis approximation space—more specifically, how to take different weight of the solution into consideration, how to cheaply and accurately select the most representative samples in order to hierarchically build the reduced basis space, as well as how to efficiently evaluate the solution and output based on the way of construction of the approximation space play a key role in the weighted reduced basis method. We address these issues in the following three aspects: the weighted greedy algorithm, the a posteriori error estimate, and the offline-online computational decomposition.

3.1. Weighted greedy algorithm. Let X be a Hilbert space equipped with the norm $||v||_X = \sqrt{A(v,v;\bar{y})}\ \forall v(y) \in H_0^1(D)$ at some reference value $\bar{y} \in \Gamma$ and let X_w be a weighted Hilbert space with norm $||v(y)||_{X_w} = w(y)||v(y)||_X\ \forall v \in X$ and $\forall y \in \Gamma$, being $w : \Gamma \to \mathbb{R}_+$ be a weight function taking positive real values. Note that both X and X_w are equivalent to $H_0^1(D)$. The weighted greedy algorithm essentially deals with the $L^\infty(\Gamma;X_w)$ optimization problem in a greedy way [37], seeking a new parameter $y^N \in \Gamma$ such that

$$y^N = \arg\sup_{y\in\Gamma} ||u^{\mathcal{N}}(y) - P_N u^{\mathcal{N}}(y)||_{X_w}, \tag{3.4}$$

where $P_N : X^{\mathcal{N}} \to X_N^{\mathcal{N}}$ is the Galerkin projection operator (by solving the Galerkin projection problem (3.2)). By solving the infinite dimensional problem (3.4) we would locate the least matching point $y^N \in \Gamma$ in $||\cdot||_{X_w}$ norm. A computable (finite dimensional) greedy algorithm relies on the following: (i) replace the parameter domain Γ by a finite training set $\Xi_{train} \subset \Gamma$ with cardinality $|\Xi_{train}| = n_{train} < \infty$; (ii) replace the mismatching term $||u^{\mathcal{N}}(y) - P_N u^{\mathcal{N}}(y)||_{X_w}$ by a cheap weighted posteriori error bound $\triangle_N^w$ that should be as sharp as possible, i.e.,

$$c_N \triangle_N^w(y) \le ||u^{\mathcal{N}}(y) - P_N u^{\mathcal{N}}(y)||_{X_w} \le C_N \triangle_N^w(y), \tag{3.5}$$

where C_N/c_N is close to 1. We leave the computation of the a posteriori error bound to the next section and present the weighted greedy algorithm in the following procedure; see Algorithm 1.

We note that for efficient computation of Galerkin projection and offline-online decomposition in practice, we normalize the snapshots with the Gram–Schmidt process to get the orthonormal basis of $\{\zeta_1^{\mathcal{N}},\dots,\zeta_N^{\mathcal{N}}\}$ such that $(\zeta_m^{\mathcal{N}},\zeta_n^{\mathcal{N}})_X = \delta_{mn}$, $1 \le m,n \le N$, and construct $X_N^{\mathcal{N}} = \text{span}\{\zeta_1^{\mathcal{N}},\dots,\zeta_N^{\mathcal{N}}\}$.

ALGORITHM 1. A weighted greedy algorithm for the construction of reduced basis approximation space.

procedure INITIALIZATION:
  sample training set $\Xi_{train} \subset \Gamma$ according to probability density function ρ;
  specify a tolerance ε_{tol} as stopping criteria of the algorithm;
  define the maximum number of reduced bases N_{max};
  choose the first sample $y^1 \in \Xi_{train}$ and build the sample space $S_1 = \{y^1\}$;
  solve the problem (2.13) at y^1, construct $X_1^{\mathcal{N}} = \text{span}\{u^{\mathcal{N}}(y^1)\}$;
end procedure
procedure CONSTRUCTION:
  for $N = 2, \ldots, N_{max}$ **do**
    compute a weighted posteriori error bound $\triangle_{N-1}^w(y)\ \forall y \in \Xi_{train}$;
    choose y^N to maximize $\triangle_{N-1}^w$, i.e., $y^N = \arg\max_{y \in \Xi_{train}} \triangle_{N-1}^w(y)$;
    if $\triangle_{N-1}^w(y^N) \le \varepsilon_{tol}$ **then**
      $N_{max} = N - 1$;
    end if
    solve problem (2.13) at y^N to obtain $u^{\mathcal{N}}(y^N)$;
    augment the sample space $S_N = S_{N-1} \cup \{y^N\}$;
    augment the reduced basis space $X_N^{\mathcal{N}} = X_{N-1}^{\mathcal{N}} \oplus \text{span}\{u^{\mathcal{N}}(y^N)\}$;
  end for
end procedure

Another algorithm that might be used for the sampling procedure is proper orthogonal decomposition (POD) [37], which is rather expensive in dealing with $L^2(\Xi_{train}; X)$ optimization and thus more suitable for low-dimensional problems. We remark that for both the greedy algorithm and the POD algorithm, an original training set Ξ_{train} is needed. Two criteria should be followed for its choice: (1) it should be cheap without too many ineffectual samples in order to avoid too much computation with little gain; (2) it should be sufficient to capture the most representative snapshots so as to build an accurate reduced basis space.

Adaptive approaches for building the training set have also been well explored by moving from a small number of samples to more samples in the space Γ adaptively; see [41] for details.

3.2. A posteriori error bound. The efficiency and reliability of the reduced basis approximation by weighted greedy algorithm relies critically on the availability of an inexpensive, sharp, and weighted a posteriori error bound $\triangle_N^w$. For every $y \in \Gamma$, let $R(v; y) \in (X^{\mathcal{N}})'$ be the residual in the dual space of $X^{\mathcal{N}}$, which is defined as

$$R(v; y) := F(v; y) - A(u_N^{\mathcal{N}}(y), v; y) \quad \forall v \in X^{\mathcal{N}}. \tag{3.6}$$

By the Riesz representation theorem [36], we have a unique function $\hat{e}(y) \in X^{\mathcal{N}}$ such that

$$(\hat{e}(y), v)_{X^{\mathcal{N}}} = R(v; y) \quad \forall v \in X^{\mathcal{N}} \text{ and } ||\hat{e}(y)||_{X^{\mathcal{N}}} = ||R(\cdot; y)||_{(X^{\mathcal{N}})'}, \tag{3.7}$$

where the $X^{\mathcal{N}}$ norm is specified as $||v||_{X^{\mathcal{N}}} = \sqrt{A(v, v; \bar{y})}$ at some reference value $\bar{y} \in \Gamma$. Define the error between the "truth" solution and the reduced basis solution as $e(y) := u^{\mathcal{N}}(y) - u_N^{\mathcal{N}}(y)$; by (2.13), (3.2), and (3.6) we have the equation

$$A(e(y), v; y) = R(v; y) \quad \forall v \in X^{\mathcal{N}}. \tag{3.8}$$

For every $y \in \Gamma$, we define the coercivity constant $\alpha(y)$ such that $\alpha(y)||v||^2_{X^{\mathcal{N}}} \leq A(v, v; y)\ \forall v \in X^{\mathcal{N}}$ and denote by $\alpha_{LB}(y)$ its lower bound, i.e., $\alpha_{LB}(y) \leq \alpha(y)$. By choosing $v = e(y)$ in (3.8) and using the Cauchy–Schwarz inequality, we have

$$\begin{aligned} \alpha_{LB}(y)||e(y)||^2_{X^{\mathcal{N}}} &\leq A(e(y), e(y); y) \\ &= R(e(y); y) \\ &\leq ||R(\cdot, y)||_{(X^{\mathcal{N}})'}||e(y)||_{X^{\mathcal{N}}} \\ &= ||\hat{e}(y)||_{X^{\mathcal{N}}}||e(y)||_{X^{\mathcal{N}}}, \end{aligned} \tag{3.9}$$

so that we can define a weighted posteriori error bound $\triangle^w_N(y)$ for the solution $u^{\mathcal{N}}_N(y), y \in \Gamma$, as

$$\triangle^w_N(y) := ||\hat{e}(y)||_{X_w}/\alpha_{LB}(y) \tag{3.10}$$

and obtain immediately the relation $||u^{\mathcal{N}}(y) - u^{\mathcal{N}}_N(y)||_{X_w} \leq \triangle^w_N(y)$ from (3.9). As for output $s(u)$,

$$|s(u^{\mathcal{N}}) - s(u^{\mathcal{N}}_N)|w(y) \leq ||s||_{(X^{\mathcal{N}})'}||u^{\mathcal{N}}(y) - u^{\mathcal{N}}_N(y)||_{X_w} \leq ||s||_{(X^{\mathcal{N}})'}\triangle^w_N(y), \tag{3.11}$$

where $||s||_{(X^{\mathcal{N}})'}$ is a constant independent of y, the same error bound can also be used in the greedy algorithm when considering the output $s^{\mathcal{N}}_N$. The efficient computation of a sharp and accurate a posteriori error bound thus relies on the computation of a lower bound of the coercivity constant $\alpha_{LB}(y)$ as well as the value $||\hat{e}(y)||_{X_w}$ for any given $y \in \Gamma$. For the former, we apply the successive constraint linear optimization method [24] to compute a lower bound $\alpha_{LB}(y)$ close to the "truth" value $\alpha(y)$. For the latter, we turn to an offline-online computational decomposition procedure.

3.3. Offline-online computational decomposition. The evaluation of the expectation $\mathbb{E}[s^{\mathcal{N}}_N]$ and the weighted a posteriori error estimator $\triangle^w_N$ requires us to compute the output $s^{\mathcal{N}}_N$ and the solution $u^{\mathcal{N}}_N$ many times. Similar situations can be encountered for other applications in the context of many query (optimal design, control) and real-time computational problems. One of the key ingredients that makes the reduced basis method stand out in this ground is the offline-online computational decomposition, which becomes possible due to the affine or linear assumption such as that made in (2.7). To start, we express the reduced basis solution in the form

$$u^{\mathcal{N}}_N(y) = \sum_{m=1}^{N} u^{\mathcal{N}}_{Nm}(y)\zeta^{\mathcal{N}}_m, \tag{3.12}$$

where we recall that $\zeta^{\mathcal{N}}_m, 1 \leq m \leq N$, are the orthonormal bases of the reduced basis space $X^{\mathcal{N}}_N$. Upon replacing the reduced basis solution in (3.2) and choosing $v = \zeta^{\mathcal{N}}_n, 1 \leq n \leq N$, we obtain for $n = 1, \ldots, N$

$$\sum_{m=1}^{N}\left(A_0(\zeta^{\mathcal{N}}_m, \zeta^{\mathcal{N}}_n) + \sum_{k=1}^{K} y_k A_k(\zeta^{\mathcal{N}}_m, \zeta^{\mathcal{N}}_n)\right) u^{\mathcal{N}}_{Nm}(y) = (f_0, \zeta^{\mathcal{N}}_n) + \sum_{k=1}^{K}(f_k, \zeta^{\mathcal{N}}_n)y_k. \tag{3.13}$$

From (3.13) we can see that the quantities $A_k(\zeta^{\mathcal{N}}_m, \zeta^{\mathcal{N}}_n), 0 \leq k \leq K, 1 \leq m, n \leq N_{max}$, and $(f_k, \zeta^{\mathcal{N}}_n), 0 \leq k \leq K, 1 \leq n \leq N_{max}$, are independent of y, and we may thus

precompute and store them in the offline procedure. In the online procedure, we only need to assemble the stiffness matrix in (3.13) and solve the resulting $N \times N$ stiffness system with much less computational effort compared to solving a full $\mathcal{N} \times \mathcal{N}$ stiffness system. As for the computation of the error bound $\triangle_N(y)$, we need to evaluate $||\hat{e}(y)||_{X^\mathcal{N}}$ at y chosen in the course of sampling procedure. We expand the residual (3.6) as

$$R(v;y) = F(v;y) - A(u_N^\mathcal{N}, v; y) = \sum_{k=0}^{K}(f_k, v)y_k - \sum_{n=1}^{N} u_{Nn}^\mathcal{N}\left(\sum_{k=0}^{K} A_k(\zeta_n^\mathcal{N}, v)y_k\right), \tag{3.14}$$

where $y_0 = 1$. Set $(\mathcal{C}_k, v)_{X^\mathcal{N}} = (f_k, v)$ and $(\mathcal{L}_n^k, v)_{X^\mathcal{N}} = -A_k(\zeta_n^\mathcal{N}, v)\, \forall v \in X_N^\mathcal{N}, 1 \le n \le N, 0 \le k \le K$, where $\mathcal{C}_k$ and $\mathcal{L}_n^k$ are the representatives in $X^\mathcal{N}$ of f_k and $\zeta_n^\mathcal{N}$, respectively, whose existence is secured by the Riesz representation theorem. By recalling $(\hat{e}(y), v)_{X^\mathcal{N}} = R(v;y)$, we obtain

(3.15)

$$\begin{aligned}||\hat{e}(y)||_{X^\mathcal{N}}^2 = &\sum_{k=0}^{K} y_k\left(\sum_{k'=0}^{K} y_{k'}(\mathcal{C}_k, \mathcal{C}_{k'})_{X^\mathcal{N}}\right) + \sum_{k=0}^{K}\sum_{n=1}^{N} y_k u_{Nn}^\mathcal{N}(y)\left(\sum_{k'=0}^{K} y_{k'} 2(\mathcal{C}_{k'}, \mathcal{L}_n^k)_{X^\mathcal{N}}\right)\\ &+ \sum_{k=0}^{K}\sum_{n=1}^{N} y_k u_{Nn}^\mathcal{N}(y)\left(\sum_{k'=0}^{K}\sum_{n'=1}^{N} y_{k'} u_{Nn'}^\mathcal{N}(y)(\mathcal{L}_n^k, \mathcal{L}_{n'}^{k'})_{X^\mathcal{N}}\right).\end{aligned}$$

Therefore, we can compute and store $(\mathcal{C}_k, \mathcal{C}_{k'})_{X^\mathcal{N}}, (\mathcal{C}_{k'}, \mathcal{L}_n^k)_{X^\mathcal{N}}, (\mathcal{L}_n^k, \mathcal{L}_{n'}^{k'})_{X^\mathcal{N}}, 1 \le n, n' \le N_{max}, 0 \le k, k' \le K$, in the offline procedure and evaluate $||\hat{e}(y)||_{X^\mathcal{N}}$ in the online procedure by assembling (3.15) with $O((K+1)^2N^2)$ scalar products, which is far more efficient provided that $O((K+1)^2N^2) \ll \mathcal{N}$.

4. Regularity and a priori convergence analysis. Without loss of generality, we work in the space X rather than in the discretization space $X^\mathcal{N}$ for regularity and a priori error estimates for the weighted reduced basis method; the regularity with respect to random variables $y \in \Gamma$ and convergence results of the weighted reduced basis approximation hold the same in the discretization space $X^\mathcal{N}$.

4.1. Regularity results.

LEMMA 4.1. *Under Assumptions* 2.1–2.3*, the solution to problem* (2.10) *satisfies* $u \in C^0(\Gamma; H_0^1(D))$*. Moreover, if* u *and* $\tilde{u}$ *are two weak solutions of problem* (2.10) *associated with data* a, f *and* $\tilde{a}, \tilde{f}$*, respectively, we have the stability estimate*

$$\begin{aligned}||u - \tilde{u}||_{C^0(\Gamma;H_0^1(D))} \le &\frac{C_P}{a_{min}}||f - \tilde{f}||_{C^0(\Gamma;L^2(D))}\\ &+ \frac{C_P}{a_{min}^2}||\tilde{f}||_{C^0(\Gamma;L^2(D))}||a - \tilde{a}||_{C^0(\Gamma;L^\infty(D))}.\end{aligned} \tag{4.1}$$

Proof. We rewrite (2.10) explicitly as $\forall y \in \Gamma$

$$\int_D a(x,y)\nabla u(x,y)\cdot\nabla v(x)dx = \int_D f(x,y)v(x)dx \quad \forall v \in H_0^1(D). \tag{4.2}$$

A similar problem holds for $\tilde{f}$ and $\tilde{a}$. By subtraction we obtain the difference equation:

$$\int_D a\nabla(u - \tilde{u})\cdot\nabla v dx = \int_D (f - \tilde{f})v dx + \int_D (\tilde{a} - a)\nabla\tilde{u}\cdot\nabla v dx. \tag{4.3}$$

By taking $v = u - \tilde{u}$, applying the Cauchy–Schwarz and Poincaré inequalities, and using Assumption 2.2 we have

$$a_{min}||u-\tilde{u}||^2_{H^1_0(D)} \le C_P||f-\tilde{f}||_{L^2(D)}||u-\tilde{u}||_{H^1_0(D)} + ||\tilde{u}||_{H^1_0(D)}||u-\tilde{u}||_{H^1_0(D)}||a-\tilde{a}||_{L^\infty(D)}, \tag{4.4}$$

so that the following stability estimate holds for $\forall y \in \Gamma$ by the fact $||\tilde{u}||_{H^1_0(D)} \le (C_P/a_{min})||\tilde{f}||_{L^2(D)}$ (due to the Lax–Milgram theorem and Assumption 2.2 for $\tilde{a}$):

$$||u(y)-\tilde{u}(y)||_{H^1_0(D)} \le \frac{C_P}{a_{min}}||f(y)-\tilde{f}(y)||_{L^2(D)} + \frac{C_P}{a^2_{min}}||\tilde{f}(y)||_{L^2(D)}||a(y)-\tilde{a}(y)||_{L^\infty(D)}. \tag{4.5}$$

Setting $\tilde{a}(y) = a(y+\delta y)$ and $\tilde{f}(y) = f(y+\delta y)$ such that $y+\delta y \in \Gamma$, we have by Assumption 2.3 that $\tilde{a}(y) \to a(y)$ in $L^\infty(D)$ and $\tilde{f}(y) \to f(y)$ in $L^2(D)$ so that $\tilde{u}(y) = u(y+\delta y) \to u(y)$ in $H^1_0(D)$ when $\delta y \to 0$. Therefore, the solution is continuous with respect to the parameter $y \in \Gamma$, i.e., $u \in C^0(\Gamma; H^1_0(D))$. ∎

A direct application of Lemma 4.1 leads to the following lemma for the existence of partial derivatives of the solution with respect to the parameter $y \in \Gamma$ as well as their bound in $H^1_0(D)$.

LEMMA 4.2. *For any $y \in \Gamma$, there exists a unique $\partial^\nu_y u(y)$ in $H^1_0(D)$ provided that Assumptions* 2.1–2.3 *are satisfied for any $y \in \Gamma$ and $\nu = (\nu_1, \dots, \nu_K) \in \Lambda$, where $\Lambda \subset \mathbb{N}^K$ is a multiple index set. Moreover, we have the following estimate:*

$$||\partial^\nu_y u(y)||_{H^1_0(D)} \le B(y)|\nu|!\eta^\nu + \frac{C_P}{a_{min}}|\nu|! \sum_{k:\nu_k \neq 0} \left(\eta^{\nu-e_k}||f_k||_{L^2(D)}\right), \tag{4.6}$$

where

$$B(y) = \frac{C_P}{a_{min}}||f(y)||_{L^2(D)}, |\nu|! = (\nu_1 + \cdots + \nu_K)!, \eta^\nu = \prod_{k=1}^K \eta_k^{\nu_k}, \eta_k = \frac{||a_k||_{L^\infty(D)}}{a_{min}}. \tag{4.7}$$

Proof. We use an induction argument for the proof in the following few steps.

Step 1. First, when $|\nu| = 0$, there exists a unique solution $u \in H^1_0(D)$ of problem (2.10) for every $y \in \Gamma$ thanks to the Lax–Milgram theorem. Moreover, the estimate

$$||\partial^\nu_y u(y)||_{H^1_0(D)} = ||u(y)||_{H^1_0(D)} \le \frac{C_P}{a_{min}}||f(y)||_{L^2(D)} = B(y) \tag{4.8}$$

holds, which verifies (4.6) for $|\nu| = 0$.

Step 2. For $|\nu| \ge 1$, we are about to prove that there exists a unique function $\partial^\nu_y u(y)$ satisfying the following general recursive equation (write $a(y)$ in short for $a(x,y)$, etc.):

$$\int_D a(y)\nabla\partial^\nu_y u(y)\cdot\nabla v = -\sum_{k:\nu_k\neq 0}\nu_k\int_D a_k\nabla\partial^{\nu-e_k}_y u(y)\cdot\nabla v + \sum_{k:\nu=e_k}\int_D f_k v \quad \forall v \in H^1_0(D), \tag{4.9}$$

where e_k is a K-dimensional vector with the kth element as 1 and all the other elements as 0. To see this, let us first show that for $|\nu| = 1$, i.e., $\nu = e_k, 1 \le k \le K$, there exists

a unique solution $\partial_y^\nu u(y)$ to (4.9). We take the perturbation $\tilde{a}(y) = a(y - he_k)$, $\tilde{f}(y) = f(y - he_k)$, and $\tilde{u}(y) = u(y - he_k)$ in (4.3) and set $D_h^k u = (u(y) - u(y - he_k))/h$; then (4.3) becomes

$$\int_D a(y)\nabla D_h^k u(y)\nabla v(y) = \int_D f_k v - \int_D a_k \nabla u(y - he_k)\cdot\nabla v \quad \forall v \in H_0^1(D), \tag{4.10}$$

which results in a unique solution $D_h^k u(y) \in H_0^1(D)$ by the Lax–Milgram theorem. Taking the limit $h \to 0$, we have by the continuity result in Lemma 4.1 that $u(y - he_k) \to u(y)$ so that $D_h^k u(y) \to \partial_y^\nu u(y)$ exists. Therefore, $\partial_y^\nu u(y)$ is a unique solution of (4.9) for $\nu = e_k, 1 \le k \le K$. By induction we suppose that there exists a unique function $\partial_y^{\tilde{\nu}} u(y)$ satisfying (4.9) for $|\tilde{\nu}| = |\nu| - 1$, i.e., $\tilde{\nu} = \nu - e_j$ for some $j = 1, \ldots, K$; then we claim that there exists a unique function $\partial_y^\nu u(y)$ satisfying (4.9) for each ν such that $|\nu| > 1$. By the same argument of perturbation and continuity property, we are able to take the derivative of (4.9) with respect to y_j, where ν is replaced by $\tilde{\nu} = \nu - e_j$ in (4.9), yielding

$$\begin{aligned}\int_D a(y)\nabla\partial_y^\nu u(y)\cdot\nabla v + \int_D a_j \nabla\partial_y^{\nu-e_j} u(y)\cdot\nabla v \\ = -\sum_{k\neq j:\nu_k\neq 0} \nu_k \int_D a_k \nabla\partial_y^{\nu-e_k} u(y)\cdot\nabla v \\ - (\nu_j - 1)\int_D a_j\nabla\partial_y^{\nu-e_j} u(y)\cdot\nabla v + \sum_{k:\nu=e_k}\int_D f_k v,\end{aligned} \tag{4.11}$$

which can be simplified by summing up the same terms to end up with (4.9). By the Lax–Milgram theorem, we have that there exists a unique solution $\partial_y^\nu u(y) \in H_0^1(D)$ to (4.9).

Step 3. We are going to show that the estimate (4.6) holds for $|\nu| \ge 1$ in this step. Upon replacing v by $\partial_y^\nu u(y)$ in (4.9), we have by Assumption 2.2 as well as the Cauchy–Schwarz and Poincaré inequalities the following estimate:

$$||\partial_y^\nu u(y)||_{H_0^1(D)} \le \sum_{k:\nu_k\neq 0} \nu_k\eta_k ||\partial_y^{\nu-e_k} u(y)||_{H_0^1(D)} + \frac{C_P}{a_{min}} \sum_{k:\nu=e_k} ||f_k||_{L^2(D)}. \tag{4.12}$$

Observe that when $|\nu| = 1$, i.e., $\nu = e_k, 1 \le k \le K$, estimate (4.12) becomes

$$||\partial_y^\nu u(y)||_{H_0^1(D)} = ||\partial_{y_k} u(y)||_{H_0^1(D)} \le B(y)\eta_k + \frac{C_P}{a_{\min}}||f_k||_{L^2(D)}, \tag{4.13}$$

which is the same as in (4.6). If $|\nu| > 1$, estimate (4.12) becomes

$$||\partial_y^\nu u(y)||_{H_0^1(D)} \le \sum_{k:\nu_k\neq 0} \nu_k\eta_k ||\partial_y^{\nu-e_k} u(y)||_{H_0^1(D)}. \tag{4.14}$$

Suppose estimate (4.6) holds for any $|\tilde{\nu}| < |\nu|$ with $|\nu| > 1$; then we have

(4.15)

$$
\begin{aligned}
&||\partial_y^\nu u(y)||_{H_0^1(D)} \\
&\quad \le \sum_{j:\nu_j \neq 0} \nu_j \eta_j ||\partial_y^{\nu - e_j} u(y)||_{H_0^1(D)} \\
&\quad \le \sum_{j:\nu_j \neq 0} \nu_j \eta_j \left(B(y)(|\nu|-1)!\eta^{\nu-e_j} + \frac{C_P}{a_{\min}}(|\nu|-1)! \sum_{k:\nu_k \neq 0} \left(\eta^{\nu - e_j - e_k} ||f_k||_{L^2(D)}\right) \right) \\
&\quad = B(y) \left(\sum_{j:\nu_j \neq 0} \nu_j \right) (|\nu|-1)!\eta^\nu + \frac{C_P}{a_{\min}} \left(\sum_{j:\nu_j \neq 0} \nu_j \right) (|\nu|-1)! \sum_{k:\nu_k \neq 0} \left(\eta^{\nu - e_k} ||f_k||_{L^2(D)}\right) \\
&\quad = B(y)|\nu|!\eta^\nu + \frac{C_P}{a_{min}}|\nu|! \sum_{k:\nu_k \neq 0} \left(\eta^{\nu - e_k} ||f_k||_{L^2(D)}\right) \equiv C_{a,f}(y)|\nu|!\eta^\nu,
\end{aligned}
$$

where

$$
C_{a,f}(y) = B(y) + C_P \sum_{k:\nu_k \neq 0, ||a_k||_{L^\infty(D)} \neq 0} \frac{||f_k||_{L^2(D)}}{||a_k||_{L^\infty(D)}}, \tag{4.16}
$$

so that estimate (4.6) also holds for ν with $|\nu| > 1$. ∎

An analytic extension of the solution u in a certain region Σ such that $\Gamma \subset \Sigma$ is a consequence of the regularity result in Lemma 4.2 provided conditions are suitable, as stated in the following lemma.

LEMMA 4.3. *Holding all the assumptions in Lemma* 4.2*, and defining*

$$
\Sigma = \left\{ z \in \mathbb{C}^K : \exists\, y \in \Gamma \ s.t.\ |(\eta \cdot |z - y|)| = \sum_{k=1}^K \eta_k |z_k - y_k| < 1 \right\}, \tag{4.17}
$$

we have the existence of an analytic extension of the stochastic solution u in the complex region Σ and we define $\Sigma(\Gamma;\tau) := \{z \in \Sigma : dist(z,\Gamma) \le \tau\} \subset \Sigma$ for the largest possible vector $\tau = (\tau_1, \ldots, \tau_K)$.

Proof. By the Taylor expansion of $u(z)$ about $y \in \Gamma$ in the complex domain we obtain

$$
u(z) = \sum_\nu \frac{\partial_y^\nu u(y)}{\nu!}(z-y)^\nu \tag{4.18}
$$

with $\nu! = \nu_1! \cdots \nu_K!$. Thanks to the regularity result in Lemma 4.2, we obtain

$$
\begin{aligned}
\left|\left| \sum_\nu \frac{\partial_y^\nu u(y)}{\nu!}(z-y)^\nu \right|\right|_{H_0^1(D)} &\le \sum_\nu \frac{|z-y|^\nu}{\nu!} ||\partial_y^\nu u(y)||_{H_0^1(D)} \\
&\le C_{a,f}(y) \sum_{n \ge 0 : |\nu| = n} \frac{|\nu|!}{\nu!} (\eta \cdot |z-y|)^\nu \\
&= C_{a,f}(y) \sum_{n \ge 0} \left(\sum_{k=1}^K \eta_k |z_k - y_k| \right)^n \\
&= \frac{C_{a,f}(y)}{1 - \sum_{k=1}^K \eta_k |z_k - y_k|},
\end{aligned} \tag{4.19}
$$

where the second inequality is due to Lemma 4.2 and the first equality comes from the generalized Newton binomial formula. In the complex region defined in (4.17), we obtain that the function $u(z)$ admits a Taylor expansion around $y \in \Gamma$ so that the solution u can be analytically extended to the complex region (4.17). ∎

4.2. A priori convergence analysis. To prove the exponential convergence of the weighted reduced basis method for problem (2.10) for the case of one random variable, i.e., $\Gamma \subset \mathbb{R}$, we bound the error by another type of constructive spectral approximation or, more specifically, extension of the Chebyshev polynomial approximation for analytic functions (see [17, Chapter 7]). The idea has also been used in the proof of the exponential convergence property of the stochastic collocation method [1]. Based on this idea we also obtain the a priori error estimate of the reduced basis approximation for multidimensional problems, e.g., $\Gamma \subset \mathbb{R}^K, K > 1$.

We define the weighted space $C^0_w(\Gamma; X)$ equipped with the norm

$$||v||_{C^0_w(\Gamma;X)} = \max_{y\in\Gamma}(w(y)||v(y)||_X) \tag{4.20}$$

for any positive continuous bounded weight function $w : \Gamma \to \mathbb{R}_+$. Because of Assumption 2.3, the linear coefficient a and forcing term f satisfy $a \in C^0(\Gamma; L^\infty(D))$ and $f \in C^0_w(\Gamma; L^2(D))$.

THEOREM 4.4. *Under Assumptions* 2.1–2.3 *with bounded* $\Gamma \subset \mathbb{R}$*, the error between the reduced basis solution* $P_N u$ *of problem* (3.2) *(recall that* $P_N : u \to u_N$ *represents the Galerkin projection operator) and the true solution* u *of problem* (2.10) *enjoys the exponential convergence*

$$||u - P_N u||_{C^0_w(\Gamma;X)} \le C^w e^{-rN} \max_{z\in\Sigma(\Gamma;\tau)} ||u(z)||_X, \tag{4.21}$$

where the constant C^w *depends on the weight* w*, and the rate* r *is defined as*

$$1 < r = log\left(\frac{2\tau}{|\Gamma|} + \sqrt{1 + \frac{4\tau^2}{|\Gamma|^2}}\right). \tag{4.22}$$

Remark 4.5. The convergence rate stated above does not depend on the specific problem (2.1). In fact, as long as $u = u(y)$ is an analytic function, the exponential convergence rate (4.21) holds for reduced basis approximation as demonstrated in the proof of this theorem later, which provides the same a priori convergence property for problems other than the elliptic problem (2.1) under linear or affine assumptions (2.7) as studied in [28, 26].

Proof. First, we note that the results obtained in the above lemmas in $H^1_0(D)$ norm are still valid in the equivalent X-norm. Given a bounded and continuous one-dimensional domain $\Gamma \subset \mathbb{R}$, we introduce the change of variables $y(t) = \bar{y} + \frac{|\Gamma|}{2}t$ with $t \in [-1, 1]$ and $\bar{y}$ the center of domain Γ, so that $y : [-1, 1] \to \Gamma$ is bijective. Let the solution of problem (2.10) be set as $\hat{u}(t) = u(y(t))$ for $t \in [-1, 1]$; then we have that $\hat{u} : [-1, 1] \to X$ can be analytically extended to $\Sigma([-1, 1], 2\tau/|\Gamma|)$ by Lemma 4.3. Consequently, their exists a spectral expansion of $\hat{u}$ on the standard Chebyshev polynomials $c_k : [-1, 1] \to \mathbb{R}$ and $|c_n| \le 1, n = 0, 1, \ldots$, in the form

$$\hat{u}(t) = \frac{u_0}{2} + \sum_{n=1}^{\infty} \hat{u}_n c_n(t). \tag{4.23}$$

The nth ($n = 0, 1, \ldots$) Chebyshev coefficient satisfies [17]

$$\hat{u}_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \hat{u}(\cos(t))\cos(nt)dt, \quad ||\hat{u}_n||_X \le 2\varrho^{-n}\max_{z\in D_\varrho}||\hat{u}(z)||_X, \tag{4.24}$$

where the elliptic disc D_ϱ is bounded by the ellipse E_ϱ with foci ± 1 and the sum of the half-axes $\varrho = 2\tau/|\Gamma| + \sqrt{1 + (4\tau^2/|\Gamma|^2)}$. Define the Nth order Chebyshev polynomial approximation of $\hat{u}$ as the truncation of (4.23) up to N terms, written as

$$\Pi_N \hat{u} = \frac{u_0}{2} + \sum_{n=1}^{N} \hat{u}_n c_n(t); \tag{4.25}$$

then the truncation error is bounded by using $|c_n| \le 1, n = N+1, \ldots,$ and (4.24) as follows:

$$||\hat{u} - \Pi_N\hat{u}||_{C^0([-1,1];X)} \le \sum_{n\ge N+1}||\hat{u}_n||_X \le \frac{2}{\varrho - 1}e^{-\log(\varrho)N}\max_{z\in D_\varrho}||\hat{u}(z)||_X. \tag{4.26}$$

Therefore, by the identity $\hat{u}(t) = u(y(t)), t \in [-1, 1]$, we have

$$||u - \Pi_N u||_{C^0(\Gamma;X)} \le \frac{2}{\varrho - 1}e^{-rN}\max_{z\in D_\varrho}||\hat{u}(z)||_X \le \frac{2}{\varrho - 1}e^{-rN}\max_{z\in\Sigma(\Gamma;\tau)}||u(z)||_X, \tag{4.27}$$

where we define $r := \log(\varrho)$, as given in (4.22). It's left to prove that the reduced basis approximation error can be bounded by the above truncation error. In fact, for any function $v \in \mathcal{P}_N(\Gamma) \otimes X$, a tensor product of polynomial space of polynomials with total degree at most N and X, we have that $\mathcal{I}_N v = v$ [9, 1], where $\mathcal{I}_N$ is the Lagrange interpolation operator based on the interpolation points $y^n, n = 1, \ldots, N+1$; see [1]. We have the following estimate with the help of the Lagrange interpolation operator:

$$\begin{aligned} ||u - P_{N+1}u||_X &\le C_0 \inf_{v\in X_{N+1}}||u - v||_X \\ &\le C_0||u - \mathcal{I}_N u||_X \\ &\le C_0 \inf_{v\in\mathcal{P}_N(\Gamma)\otimes X}(||u - v||_X + ||v - \mathcal{I}_N u||_X) \\ &= C_0 \inf_{v\in\mathcal{P}_N(\Gamma)\otimes X}(||u - v||_X + ||\mathcal{I}_N v - \mathcal{I}_N u)||_X) \\ &\le (C_0 + C_1)\inf_{v\in\mathcal{P}_N(\Gamma)\otimes X}||u - v||_X, \end{aligned} \tag{4.28}$$

where the first inequality is due to Cea's lemma [36] with constant $C_0 < \infty$ and the second due to the fact $\inf_{v\in X_{N+1}}||u - v||_X \le ||u - \mathcal{I}_N u||_X$; as for the last inequality, we have used the property that the Lagrange interpolation operator $\mathcal{I}_N$ is linear and $||\mathcal{I}_N v||_X \le C_1||v||_X \ \forall v \in C^0(\Gamma, X)$ for a constant $C_1 < \infty$ (see [1]). Moreover, because the Chebyshev polynomials $c_k \in \mathcal{P}_N([-1, 1]), k = 0, 1, \ldots, N$, we have

$$\inf_{v\in\mathcal{P}_N(\Gamma)\otimes X}||u - v||_X = \inf_{\hat{v}\in\mathcal{P}_N([-1,1])\otimes X}||\hat{u} - \hat{v}||_X \le ||\hat{u} - \Pi_N\hat{u}||_X = ||u - \Pi_N u||_X. \tag{4.29}$$

A combination of (4.27), (4.28), and (4.29) leads to the following bound for the reduced basis approximation error with $C = 2(C_0 + C_1)e^r/(\varrho - 1)$:

$$||u - P_N u||_X \le Ce^{-rN}\max_{z\in\Sigma(\Gamma;\tau)}||u(z)||_X. \tag{4.30}$$

Since the reduced basis approximation $P_N u$ satisfies the linear system (3.13), which can be written in the compact form as

$$A(P_N u, v; y) = F(v; y) \quad \forall v \in X_N, \tag{4.31}$$

we obtain the same regularity for $P_N u$ as for the solution u to system (2.10) with respect to the parameter y. In particular, $P_N u \in C_w^0(\Gamma; X)$, so that $u - P_N u \in C_w^0(\Gamma; X)$. Multiplying both sides of (4.30) by the weight function w and taking the maximum value over the parameter domain Γ, we have obtain the exponential convergence result (4.21) with the constant $C^w = C \max_{y\in\Gamma} w(y)$. ∎

Remark 4.6. The exponential convergence result (4.21) holds for the case of a single parameter in a bounded parameter domain $|\Gamma| < \infty$. Extension to a single parameter in the unbounded domain, e.g., a normal distributed random variable, requires that the data a and f feature a fast decrease at the parameter far away from the origin, and the constructive approximation by spectral expansion on Chebyshev polynomials (4.23) is replaced by that on Hermite polynomials [1]. The proof follows the same procedure as for Theorem 4.4 and we skip it for simplicity.

As for the reduced basis approximation in the multidimensional case, we have the following a priori error estimate.

THEOREM 4.7. *Under Assumptions* 2.1–2.3 *with* $\Gamma \subset \mathbb{R}^K, K > 1$*, the approximation error of the reduced basis solution can be bounded by*

$$||u - P_N u||_{C_w^0(\Gamma;X)} \le \max_{z\in\Sigma(\Gamma;\tau)} ||u(z)||_X \sum_{k=1}^{K} C_k^w e^{-r_k N_k}, \tag{4.32}$$

where the constants $C_k^w, 1 \le k \le K$*, depend on the weight* w *and dimension* k*,* $N = \prod_{k=1}^K N_k$*, and the rate* r_k *is defined as*

$$1 < r_k = log\left(\frac{2\tau_k}{|\Gamma_k|} + \sqrt{1 + \frac{4\tau_k^2}{|\Gamma_k|^2}}\right), 1 \le k \le K. \tag{4.33}$$

Proof. Let us choose the training set as all the nodes of a tensor product grid, i.e., $\Xi_{train} := \{(y_1^{n_1}, \ldots, y_K^{n_K}), 1 \le n_k \le N_k, 1 \le k \le K\}$, for instance, the Gauss quadrature nodes corresponding to the probability density function of the random vector y. We define the reduced basis space $X_N^k, 1 \le k \le K$, as a linear combination of the snapshots $u(y)$ at $y = (y_k^1, y_k^*), \ldots, (y_k^N, y_k^*)$, where $y_k^n \in \Gamma_k, 1 \le n \le N$, and y_k^* is any point in the rest of the $K-1$ dimensional domain denoted as Γ_k^*. Correspondingly, we define the Galerkin projection operator $P_N^k : X \to X_N^k, 1 \le k \le K$, such that $P_N^k u$ is the solution of the reduced problem (4.31) in X_N^k whenever u is the solution of the original problem (2.10) in X at any $y = (y_k, y_k^*) \in \Gamma_k \times \Gamma_k^*$. Let X_N be the reduced basis space spanned by the snapshots at all the $N = \prod_{k=1}^K N_k$ samples and $P_N : X \to X_N$ be the associated Galerkin projection operator; then we have for the solution $u \in X$ of problem (2.10) at any $y = (y_1, y_1^*) \in \Gamma_1 \times \Gamma_1^*$,

$$P_N u = P_N^1 \circ \cdots \circ P_N^K u, \tag{4.34}$$

the symbol $\circ$ being the composition of the projection operators. By triangular inequality, we have

$$||u - P_N u||_X \le ||u - P_N^1 u||_X + ||P_N^1(u - P_N^2 \circ \cdots \circ P_N^K u)||_X, \tag{4.35}$$

where we can bound the first term as in (4.30) by

(4.36)
$$||u - P_N^1 u||_X \le C_1 e^{-r_1 N_1} \max_{(z_1, z_1^*) \in \Sigma(\Gamma_1 \times \Gamma_1^*; \tau)} ||u(z)||_X \le C_1 e^{-r_1 N_1} \max_{z \in \Sigma(\Gamma; \tau)} ||u(z)||_X,$$

where the constant C_1 has similar definition as C in (4.30) and r_k is defined in (4.33). As for the second term, thanks to the fact that $||P_N^1 v||_X \le ||v||_X$ we have

$$||P_N^1(u - P_N^2 \circ \cdots \circ P_N^K u)||_X \le ||u - P_N^2 \circ \cdots \circ P_N^K u||_X. \tag{4.37}$$

By iteration, we obtain the error bound

$$||u - P_N u||_X \le \max_{z \in \Sigma(\Gamma; \tau)} ||u(z)||_X \sum_{k=1}^{K} C_k e^{-r_k N_k}, \tag{4.38}$$

which leads to the a priori error estimate (4.32) by multiplying by the weight function w on both sides and noting that $P_N u \in C_w^0(\Gamma; X)$, where the constants $C_k^w := C_k \max_{y \in \Gamma} w(y), 1 \le k \le K$. ∎

Remark 4.8. In practice, the training set Ξ_{train} can be chosen in a more general way, e.g., by sampling according to the probability density function, and the cardinality of the reduced basis space X_N is much lower than $\prod_{k=1}^{K} N_k$ given in the theorem. In fact the error estimate obtained in this theorem is rather crude. An improved convergence rate $e^{-r' N^{\beta/(\beta+1)}}$ was achieved in [5] provided that the Kolmogorov N-width by the optimal N dimensional approximation decays as e^{-rN^β} in a more general setting, e.g., if Γ is not bounded. However, the Kolmogorov N-width is not available in general.

A direct consequence of Theorems 4.4 and 4.7 for the convergence of quantity of interest and its statistical moments is as follows.

COROLLARY 4.9. *Suppose that the assumptions in Theorem* 4.4 *are satisfied. We have*

$$||s(u) - s(P_N u)||_{C_w^0(\Gamma)} \le ||s||_{X'} ||u - P_N u||_{C_w^0(\Gamma; X)}, \tag{4.39}$$

and for the kth order statistical moment, where $k = 1, 2, \ldots$, *we have by* (3.3)

$$\begin{aligned} &|\mathbb{E}[s^k(u)] - \mathbb{E}[s^k(P_N u)]| \\ &\quad \approx \left| \sum_{m=1}^{M} w(y^m)(s(u; y^m) - s(P_N; y^m)) \left(\sum_{l=0}^{k-1} s^{k-1-l}(u; y^m) s^l(P_N u; y^m) \right) \right| \\ &\quad \le M ||s(u) - s(P_N u)||_{C_w^0(\Gamma)} C_s^k, \end{aligned} \tag{4.40}$$

where C_s^k *is a constant depending on the output* s *and the statistical moment* k *with* $C_s^1 = 1$.

5. Numerical examples. In this section, we present several numerical examples to illustrate the efficiency of the weighted reduced basis method compared to the reduced basis method and the stochastic collocation method. The output of interest is defined as the integral of the solution over the physical domain D

$$s(y) = \int_D u(x, y) dx. \tag{5.1}$$

We define the following two errors as criteria of different numerical methods:

$$||s - s_N||_{C^0_w(\Gamma)} \quad \text{and} \quad |\mathbb{E}[s] - \mathbb{E}[s_N]|, \tag{5.2}$$

where s_N is the approximated value of s obtained using N bases for (weighted) reduced basis method or N collocation points for the stochastic collocation method. In particular, we use the weight function in one dimension as the probability density function of the random variable obeying Beta(α, β) distribution with shape parameter α and β providing distinctive property of the weight, defined as

$$w(y; \alpha, \beta) = \frac{1}{2\text{Beta}(\alpha, \beta)}(1+y)^{\alpha-1}(1-y)^{\beta-1}, \quad y \in [-1, 1], \tag{5.3}$$

where Beta(α, β) is a constant (beta function) chosen so that $w(\cdot; \alpha, \beta)$ is a probability density function. In our numerical experiments, we use the Gauss–Jacobi quadrature formula to compute expectation (5.2) with the solution at the abscissas evaluated by the reduced basis methods. As for the stochastic collocation method, we use the Gauss–Jacobi abscissas as the collocation points, which is more accurate than other choices, especially when the weight function is more concentrated. We specify the detailed setting of the weighted reduced basis method in the following subsections. The physical domain is a square $D = (-1, 1)^2$ and homogeneous Dirichlet boundary conditions are prescribed on the entire boundary ∂D.

5.1. One-dimensional problem. We set the stochastic coefficient $a(x, \omega)$, $x = (x_1, x_2) \in D$, in problem (2.1) as

$$a(x, \omega) = \frac{1}{10}(1.1 + \sin(2\pi x_1)Y(\omega)) \tag{5.4}$$

with random variable $Y \sim$ Beta(α, β) with $(\alpha, \beta) = (1, 1)$, $(10, 10)$, and $(100, 100)$, respectively. We remark that when $(\alpha, \beta) = (1, 1)$ the weighted reduced basis method becomes a reduced basis method with uniformly distributed random variable, which has been examined in [10]. The left of Figure 5.1 depicts the shape of weight at different locations. The forcing term is the deterministic value $f = 1$ for simplicity. We use a tolerance at the same value $\varepsilon = 1 \times 10^{-15}$ for three different weight functions to

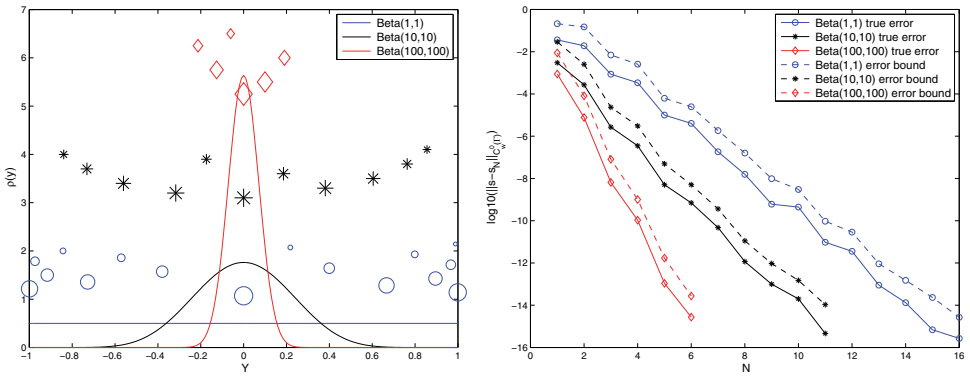

FIG. 5.1. *Left: probability density function of Beta(α, β) distribution with different α, β and samples selected by weighted reduced basis approximation in order; the bigger the size the earlier it has been selected. Right: convergence of the error* $\log_{10}\left(||s - s_N||_{C^0_w(\Gamma)}\right)$ *by weighted RBM.*

stop the greedy algorithm. $n_{train} = 1000$ samples are uniformly selected to construct the reduced basis space. Another 1000 samples are used to test the accuracy of different methods. The exponential convergence of the error $||s - s_N||_{C^0_w(\Gamma)}$ and the error bound (3.11) in logarithmic scale for three different weight functions are displayed on the right side of Figure 5.1 for the weighted reduced basis method. The maximum number of bases $N_{max} = 16, 11, 6$ built at the training samples with selection order are visualized by the marker size on the left side of Figure 5.1; they are quite different for different weight functions. From the location and selecting order of the samples on the left of Figure 5.1, we can tell that the weight function plays an important role in choosing the most representative bases.

In the comparison of the convergence property of the reduced basis method, the weighted reduced basis method as well as the stochastic collocation method, we select the weight function of Beta(10, 10) and compute the two errors defined in (5.2) with the results shown in Figure 5.2. It's evident that the weighted reduced basis method outperforms the reduced basis method in both norms, and these two methods are more accurate than the stochastic collocation method in the $||\cdot||_{C^0_w(\Gamma)}$ norm. As for the expectation, the weighted reduced basis method is the best and the reduced basis method does not beat the stochastic collocation method because it doesn't take the weight into account.

However, as demonstrated in [10], the computation of both reduced basis methods for the one-dimensional stochastic problem is more expensive than that of the stochastic collocation method because of the offline construction with a large number of training samples, especially for the problem requiring low computational effort in one deterministic solving. Similar numerical examples for some other weight functions are presented in the appendix for expository convenience.

5.2. Multiple-dimensional problem. For the test of a multiple-dimensional problem, we specify the coefficient $a(x, \omega), x = (x_1, x_2) \in D$, as

$$a(x,\omega) = \frac{1}{10}\left(4 + \left(\frac{\sqrt{\pi}L}{2}\right)^{1/2} y_1(\omega)\right) + \frac{1}{10}\left(\sum_{n=1}^{2} \sqrt{\lambda_n}\left(\sin(n\pi x_1) y_{2n}(\omega) + \cos(n\pi x_1) y_{2n+1}(\omega)\right)\right), \tag{5.5}$$

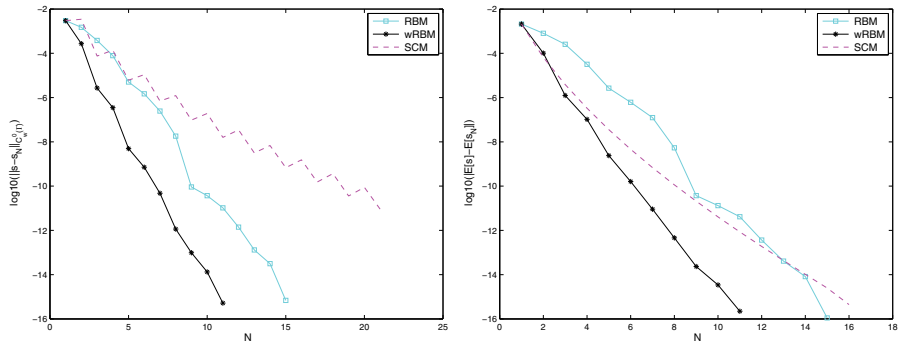

FIG. 5.2. *Left: convergence of the error* $\log_{10}\left(||s - s_N||_{C^0_w(\Gamma)}\right)$ *by the reduced basis method (RBM), the weighted reduced basis method (wRBM), and the stochastic collocation method (SCM). Right: convergence of the error* $\log_{10}\left(|E[s] - E[s_N]|\right)$ *by RBM, wRBM, and SCM, both with* $K = 1$, *Beta*$(10, 10)$.

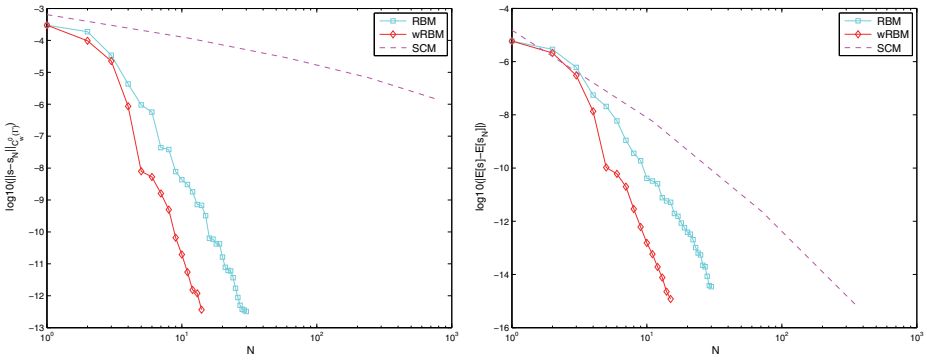

FIG. 5.3. *Left: convergence of the error* $\log_{10}\left(||s-s_N||_{C^0_w(\Gamma)}\right)$. *Right: convergence of the error* $\log_{10}\left(|E[s]-E[s_N]|\right)$, *computed by RBM, wRBM, and SCM, both with* $K=5$, *Beta*$(100,100)$.

where $y_k, 1 \le k \le 5$, obeying Beta$(100,100)$, $L = 1/4$, and $\lambda_1 = 0.3798, \lambda_2 = 0.2391$. A sufficient number of $n_{train} = 10000$ samples (in fact $n_{train} = 1000$ provides almost the same result in this example) obeying independent and identically distributed $y_k \sim$ Beta$(100,100), 1 \le k \le 5$, are taken within the parameter domain $\Gamma = [-1,1]^5$ to construct the reduced basis space and another 1000 samples following the same distribution are taken independently to test different methods. We compare the performance of the weighted reduced basis method, the reduced basis method, and a sparse grid collocation method, with results displayed in Figure 5.3. The two reduced basis methods are obviously more efficient in both norms (5.2) with the weighted type providing faster convergence: the number of bases constructed for the weighted reduced basis method ($N_{max} = 15$) is half that necessary for the reduced basis method ($N_{max} = 30$).

As for the computational effort, the stochastic collocation method with sparse grid depends critically on the dimension [32], while the reduced basis methods are near the best approximation in the sense that they considerably alleviate the "curse of dimensionality" for the analytic problem and save the computational effort significantly for high-dimensional problems, especially those with a cost for one deterministic solving. The weighted reduced basis method uses fewer bases than the conventional reduced basis method in both offline construction and online evaluation and thus costs less computational effort, particularly for high concentrated weight function as shown in the above examples. For a detailed comparison of computational cost for the reduced basis method and the stochastic collocation method in various conditions, notably for large-scale and high-dimensional problems, see [10].

6. Concluding remarks. We proposed a weighted reduced basis method to deal with parametric elliptic problems with distinctive weight or importance at different values of the parameters. This method is particularly useful in solving stochastic problems with random variables obeying various probability distributions. Analytic regularity of the stochastic solution with respect to random variables was obtained under certain assumptions for the random input data, based on which an exponential convergence property of this method was studied by constructive approximation of general functions with analytic dependence on the parameters. The computational efficiency of the proposed method in comparison with the reduced basis method as well as the (sparse grid) stochastic collocation method was demonstrated numerically for both univariate and multivariate stochastic elliptic problems.

There are a few potential limitations we would like to warn the reader about: first, the performance of the weighted reduced basis method for low regularity problems is to be investigated, possibly improved by combination with the hp-adaptive reduced basis method [18]. Second, the efficient empirical interpolation method [4, 12] needs to be applied in order to use the weighted reduced basis method to solve nonlinear stochastic problems or linear stochastic problems with nonaffine random inputs exhibiting various probability structure. Finally, we would like to mention that application of the weighted reduced basis method to more general problems, e.g., parabolic problems [22], fluid dynamics [35], multiphysical problems [27], stochastic optimization problems [11], and inverse problems [29], as well as more general stochastic problems with various probability structures is ongoing research.

7. Appendix. To illustrate more about the efficiency of the weighted reduced basis method, we present the following numerical examples with some widely used weight functions other than those considered in section 5:

1. weight function as truncated probability density function of normal distributed random variable:

$$a(x,\omega) = \frac{1}{10}(3.1 + \sin(2\pi x_1)Y(\omega)\mathbb{I}(|Y| \le 3)),$$

where

$$Y \sim \text{Normal}(\mu,\sigma), w(y) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right);$$

2. weight function as truncated probability density function of gamma distributed random variable:

$$a(x,\omega) = \frac{1}{10}(10.1 + \sin(2\pi x_1)Y(\omega)\mathbb{I}(Y \le 10)),$$

where

$$Y \sim \text{Gamma}(k,\gamma), w(y) = \frac{1}{\gamma^k\Gamma(k)}y^{k-1}\exp\left(-\frac{y}{\gamma}\right);$$

3. weight function as truncated probability density function of Poisson distributed random variable:

$$a(x,\omega) = \frac{1}{10}(100.1 + \sin(2\pi x_1)Y(\omega)\mathbb{I}(Y \le 100)),$$

where

$$Y \sim \text{Poisson}(\lambda), w(y) = \frac{\lambda^y e^{-\lambda}}{y!}.$$

The selected samples for different weight functions and error of $\log_{10}(||s - s_N||_{C^0_w(\Gamma)})$ are displayed in Figures 7.1, 7.2, and 7.3, respectively, from which we can observe that the samples are effectively chosen according to the weight functions. Consequently, both the offline construction and the online evaluation become more efficient by the weighted reduced basis method than the conventional one.

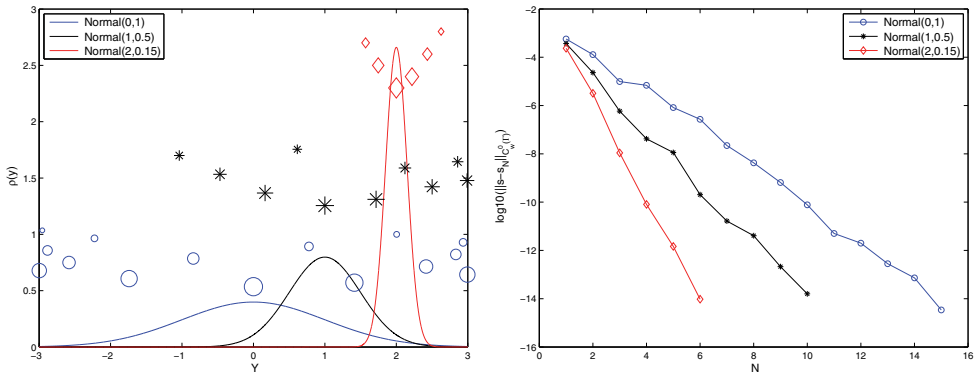

FIG. 7.1. *Left: probability density function of* $Y \sim Normal(\mu, \sigma)$ *with different* μ, σ *and samples selected by weighted reduced basis approximation in order; the bigger the size the earlier it has been selected. Right: convergence of the error* $\log_{10}\left(||s-s_N||_{C^0_w(\Gamma)}\right)$ *by the weighted reduced basis method.*

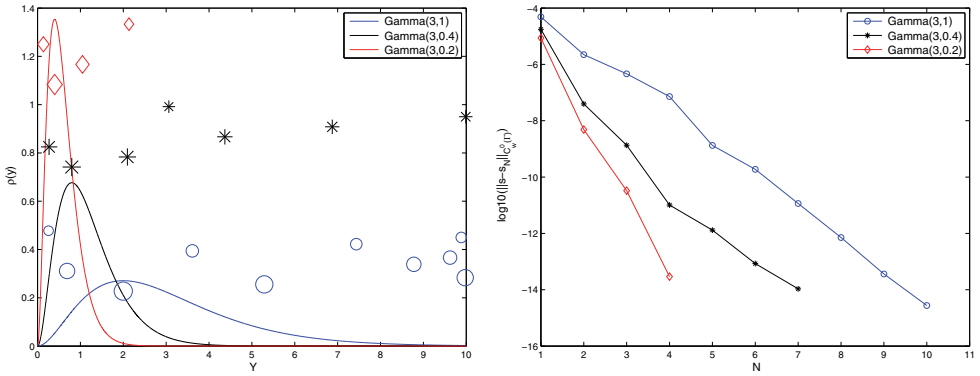

FIG. 7.2. *Left: probability density function of* $Y \sim Gamma(k, \gamma)$ *with different* γ *and samples selected by weighted reduced basis approximation in order; the bigger the size the earlier it has been selected. Right: convergence of the error* $\log_{10}\left(||s-s_N||_{C^0_w(\Gamma)}\right)$ *by the weighted reduced basis method.*

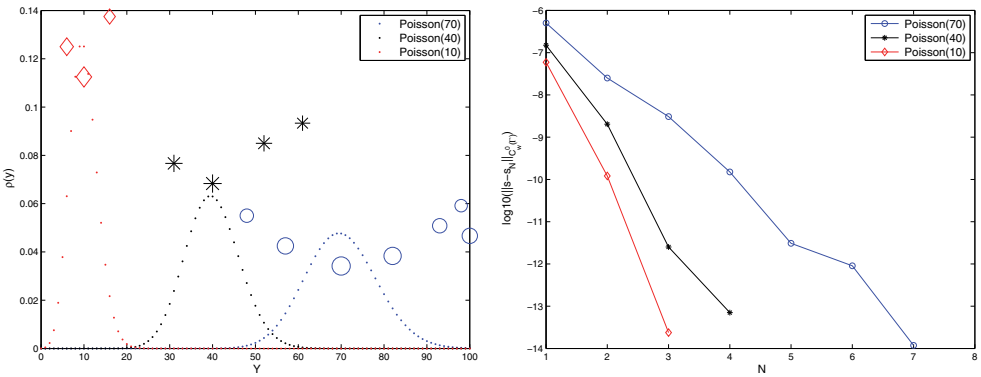

FIG. 7.3. *Left: probability density function of* $Y \sim Poisson(\lambda)$ *with different* λ *and samples selected by weighted reduced basis approximation in order; the bigger the size the earlier it has been selected. Right: convergence of the error* $\log_{10}\left(||s-s_N||_{C^0_w(\Gamma)}\right)$ *by the weighted reduced basis method.*

Acknowledgments. We acknowledge the use of the MATLAB package MLife previously developed by Prof. Fausto Saleri of MOX, Politecnico di Milano. The third author acknowledges the support provided by the program NOFYSAS (New Opportunities for Young Scientists) at SISSA, International School for Advanced Studies, Trieste.

REFERENCES

[1] I. Babuška, F. Nobile, and R. Tempone, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM J. Numer. Anal., 45 (2007), pp. 1005–1034.

[2] I. Babuška, R. Tempone, and G. E. Zouraris, *Galerkin finite element approximations of stochastic elliptic partial differential equations*, SIAM J. Numer. Anal., 42 (2005), pp. 800–825.

[3] J. Bäck, F. Nobile, L. Tamellini, and R. Tempone, *Stochastic spectral Galerkin and collocation methods for PDEs with random coefficients: A numerical comparison*, in Spectral and High Order Methods for Partial Differential Equations, Lecture Notes in Comput. Sci. Engrg. 76, Springer, New York, 2011, pp. 43–62.

[4] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera, *An empirical interpolation method: Application to efficient reduced-basis discretization of partial differential equations*, C. R. Math. Anal. Numér., 339 (2004), pp. 667–672.

[5] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk, *Convergence rates for greedy algorithms in reduced basis methods*, SIAM J. Math. Anal., 43 (2011), pp. 1457–1472.

[6] S. Boyaval, C. Le Bris, T. Lelièvre, Y. Maday, N. C. Nguyen, and A. T. Patera, *Reduced basis techniques for stochastic problems*, Arch. Comput. Methods Engrg., 17 (2010), pp. 435–454.

[7] S. Boyaval, C. LeBris, Y. Maday, N. C. Nguyen, and A. T. Patera, *A reduced basis approach for variational problems with stochastic parameters: Application to heat conduction with variable Robin coefficient*, Comput. Methods Appl. Mech. Engrg., 198 (2009), pp. 3187–3206.

[8] A. Buffa, Y. Maday, A. Patera, C. Prudhomme, and G. Turinici, *A priori convergence of the greedy algorithm for the parametrized reduced basis*, ESAIM Math. Model. Numer. Anal., 46 (2012), pp. 595–603.

[9] C. Canuto, Y. Maday, and A. Quarteroni, *Analysis of the combined finite element and Fourier interpolation*, Numer. Math., 39 (1982), pp. 205–220.

[10] P. Chen, A. Quarteroni, and G. Rozza, *Comparison of reduced basis method and collocation method for stochastic elliptic problems*, J. Sci. Comput., in press.

[11] P. Chen, A. Quarteroni, and G. Rozza, *Stochastic optimal Robin boundary control problems of advection-dominated elliptic equations*, SIAM J. Numer. Anal., 51 (2013), pp. 2700–2722.

[12] P. Chen, A. Quarteroni, and G. Rozza, *A weighted empirical interpolation method: A priori convergence analysis and applications*, accepted in ESAIM: M2AN.

[13] A. Cohen, R. DeVore, and C. Schwab, Convergence rates of best N-term Galerkin approximations for a class of elliptic SPDEs, Found. Comput. Math., 10 (2010), pp. 615–646.

[14] A. Cohen, R. Devore, and C. Schwab, *Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDE's*, Anal. Appl., 9 (2011), pp. 11–47.

[15] A. C. Davison, *Statistical Models*, Cambridge University Press, Cambridge, UK, 2003.

[16] R. DeVore, G. Petrova, and P. Wojtaszczyk, *Greedy Algorithms for Reduced Bases in Banach Spaces*, preprint, arXiv:1204.2290, 2012.

[17] R. A. DeVore and G. G. Lorentz, *Constructive Approximation*, Springer, New York, 1993.

[18] J. L. Eftang, A. T. Patera, and E. M. Rønquist, *An "hp" certified reduced basis method for parametrized elliptic partial differential equations*, SIAM J. Sci. Comput., 32 (2010), pp. 3170–3200.

[19] O. G. Ernst, C. E. Powell, D. J. Silvester, and E. Ullmann, *Efficient solvers for a linear stochastic Galerkin mixed formulation of diffusion problems with random data*, SIAM J. Sci. Comput., 31 (2009), pp. 1424–1447.

[20] G. S. Fishman, *Monte Carlo: Concepts, Algorithms, and Applications*, Springer, New York, 1996.

[21] R. G. Ghanem and P. D. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Dover Civil and Mechanical Engineering, New York, 2003.

[22] M. A. Grepl and A. T. Patera, *A posteriori error bounds for reduced-basis approximations of parametrized parabolic partial differential equations*, ESAIM Math. Model. Numer. Anal., 39 (2005), pp. 157–181.
[23] S. Heinrich, *Multilevel Monte Carlo methods*, Large-Scale Sci. Comput., 2179 (2001), pp. 58–67.
[24] D. B. P Huynh, G. Rozza, S. Sen, and A. T. Patera, *A successive constraint linear optimization method for lower bounds of parametric coercivity and inf-sup stability constants*, C. R. Math. Anal. Numér., 345 (2007), pp. 473–478.
[25] M. Kleiber and T. D. Hien, *The stochastic finite element method*, Wiley, New York, 1992.
[26] T. Lassila, A. Manzoni, A. Quarteroni, and G. Rozza, *Generalized reduced basis methods and n-width estimates for the approximation of the solution manifold of parametric PDEs*, in Analysis and Numerics of Partial Differential Equations Series, Springer INdAM Series 4, F. Brezzi, P. Colli Franzone, U. Gianazza, and G. Gilardi, eds., Springer, New York, 2013.
[27] T. Lassila, A. Quarteroni, and G. Rozza, *A reduced basis model with parametric coupling for fluid-structure interaction problems*, SIAM J. Sci. Comput., 34 (2012), pp. 1187–1213.
[28] Y. Maday, A. T. Patera, and G. Turinici, *A priori convergence theory for reduced-basis approximations of single-parameter elliptic partial differential equations*, J. Sci. Comput., 17 (2002), pp. 437–446.
[29] A. Manzoni, T. Lassila, A. Quarteroni, and G. Rozza, *A reduced-order strategy for solving inverse bayesian shape identification problems in physiological flows*, in High-Performance Scientific Computing, Springer, New York, 2012.
[30] H. Niederreiter, *Quasi-Monte Carlo Methods*, Wiley, New York, 1992.
[31] F. Nobile, R. Tempone, and C. G. Webster, An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. SIAM J. Numer. Anal., 46 (2008), pp. 2411–2442.
[32] F. Nobile, R. Tempone, and C. G. Webster, *A sparse grid stochastic collocation method for partial differential equations with random input data*, SIAM J. Numer. Anal., 46 (2008), pp. 2309–2345.
[33] A. Nouy, *Recent developments in spectral stochastic methods for the numerical solution of stochastic partial differential equations*, Arch. Comput. Methods Engrg., 16 (2009), pp. 251–285.
[34] A. T. Patera and G. Rozza, *Reduced basis approximation and a posteriori error estimation for parametrized partial differential equations Version* 1.0, http://augustine.mit.edu (2007).
[35] A. Quarteroni and G. Rozza, *Numerical solution of parametrized Navier–Stokes equations by reduced basis methods*, Numer. Methods Partial Differential Equations, 23 (2007), pp. 923–948.
[36] A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer, New York, 1994.
[37] G. Rozza, D. B. P. Huynh, and A. T. Patera, *Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations*, Arch. Comput. Methods Engrg., 15 (2008), pp. 229–275.
[38] C. Schwab and R. A. Todor, *Karhunen-Loève approximation of random fields by generalized fast multipole methods*, J. Comput. Phys., 217 (2006), pp. 100–122.
[39] D. Xiu and J. S. Hesthaven, *High-order collocation methods for differential equations with random inputs*, SIAM J. Sci. Comput., 27 (2005), pp. 1118–1139.
[40] D. Xiu and G. E. Karniadakis, *The Wiener-Askey polynomial chaos for stochastic differential equations*, SIAM J. Sci. Comput., 24 (2003), pp. 619–644.
[41] S. Zhang, *Efficient Greedy Algorithms for Successive Constraints Methods with High-Dimensional Parameters*, Tech report 23, Brown University, Providence, RI, 2011.