Uncertainty Quantification for Stochastic Approximation Limits Using Chaos Expansion

Stéphane Crépey, Gersende Fort, Emmanuel Gobet, Uladzislau Stazhynski

Uncertainty Quantification for Stochastic Approximation Limits Using Chaos Expansion*

S. Crépey †, G. Fort ‡, E. Gobet§, and U. Stazhynski ¶

Abstract. We analyze the uncertainty quantification for the limit of a Stochastic Approximation (SA) algorithm. In our setup, this limit $\phi^\star$ is defined as the zero of an intractable function and is modeled as uncertain through a parameter θ: we aim at deriving the probabilistic distribution of $\phi^\star(\theta)$, given a probability distribution π for θ. We introduce the so-called Uncertainty for SA (USA) algorithm, an SA algorithm in increasing dimension for computing the basis coefficients of a chaos expansion of $\phi^\star(\cdot)$ on an orthogonal basis of a suitable Hilbert space. USA returns a finite set of coefficients, providing an approximation $\widehat{\phi^\star}(\cdot)$ of $\phi^\star(\cdot)$; for a convenient choice of the basis, it also provides an approximation of the expectation, of the variance-covariance matrix and of higher order moments of $\{\widehat{\phi^\star}(\theta); \theta \sim \pi\}$. The almost-sure and L^p-convergences of USA, in the Hilbert space, are established under mild, tractable conditions and constitute original results, not covered by the existing literature for convergence analysis of infinite dimensional SA algorithms. Finally, USA is illustrated and the role of its design parameters is discussed through a numerical analysis.

Key words. Stochastic Approximation in Hilbert space, Uncertainty Quantification, Chaos expansion, Stochastic Programming, Almost-sure convergence.

AMS subject classifications. 62L20, 41A10, 41A25, 60B12, 60F15, 90C15.

1. Introduction.

Since the seminal work of Robbins and Monro [25], the Stochastic Approximation (SA) method has become mainstream for various applications - see e.g. [18, 5] for a general introduction and examples of applications. SA is used to find zeros of an intractable function $h : \mathbb{R}^q \to \mathbb{R}^q$ of the form $z \mapsto \mathbb{E}[H(z, V)]$, where V is some random variable and H can be computed explicitly (as opposed to its expectation h). Uncertainty Quantification (UQ) applied to SA limits corresponds to the situation where the distribution of the random variable V or/and the function H is not known exactly. In this work, this is studied in the form of *(i)* the distribution of V depends on an unknown parameter θ in Θ, i.e. $V \sim \mu(\theta, \mathrm{d}v)$, for which only some probability distribution π is available; *(ii)* the function H is modeled through a dependency in the parameter θ.

*This research benefited from the support of the Chair *Financial Risks* of the *Risk Foundation*, the Chair *Markets in Transition* of the *Fédération Bancaire Française*, the *Finance for Energy Market Research Centre*, the ANR project *CAESARS* (ANR-15-CE05-0024).

†Email: stephane.crepey@univ-evry.fr. LaMME, Univ. Evry, CNRS, Université Paris-Saclay, 91037, Evry, France.

‡Email: Gersende.Fort@math.univ-toulouse.fr. IMT, Université de Toulouse & CNRS, F-31062 Toulouse, France.

§Email: emmanuel.gobet@polytechnique.edu. Centre de Mathématiques Appliquées (CMAP), Ecole Polytechnique and CNRS, Université Paris-Saclay, Route de Saclay, 91128 Palaiseau Cedex, France.

¶Email: uladzislau.stazhynski@polytechnique.edu. Centre de Mathématiques Appliquées (CMAP), Ecole Polytechnique and CNRS, Université Paris-Saclay, Route de Saclay, 91128 Palaiseau Cedex, France. Corresponding author.

Therefore, our mathematical framework is the following: for any $(z, \theta) \in \mathbb{R}^q \times \Theta$, set

$$h(z, \theta) := \int_{\mathcal{V}} H(z, v, \theta)\mu(\theta, \mathrm{d}v), \tag{1.1}$$

and for any θ, denote by $\phi^\star(\theta)$ its zero with respect to (w.r.t.) the z-variable - for simplicity in this section and the next one, we assume uniqueness of such $\phi^\star$ (belonging to a suitable Hilbert space), but in Section 3 our main algorithm is proved to be convergent even in the case of multiple zeros. The UQ problem consists in determining the distribution of $\{\phi^\star(\theta) : \theta \sim \pi\}$, or at least its first moments. In UQ applications (see [19, 27]), the function $\phi^\star$ is typically given as the solution of an auxiliary problem, which requires some numerical computations: quite often, $\phi^\star$ solves a Partial Differential Equation (PDE); in our setting, the function $\phi^\star$ is defined for (π-almost) all θ, as the limit of an SA algorithm parameterized by θ.

In the last two decades, UQ has become a huge concern regarding both research and industrial applications; our goal is to study the UQ problem for the SA limits, which, to the best of our knowledge, has not been investigated so far.

For the UQ analysis, a first possible approach is based on crude Monte Carlo (MC) methods: they consist in sampling M values θ_m under π, and then compute, for each sample θ_m, an approximation $\widehat{\phi^\star(\theta_m)}$ of $\phi^\star(\theta_m)$. The distribution of the random variable $\{\phi^\star(\theta), \theta \sim \pi\}$ is then approximated by the empirical distribution of $\{\widehat{\phi^\star(\theta_m)} : 1 \leq m \leq M\}$. When $\phi^\star$ solves a PDE, a global error analysis is performed in [2], accounting for both the sampling error and the PDE discretization error, which are decoupled in some way. In our SA setting, a naive approach would be to compute $\widehat{\phi^\star(\theta_m)}$ as the output of a standard SA algorithm for fixed θ_m; the approach we propose below will be different since it will couple in an efficient way the outer Monte Carlo sampling of θ_m and the inner Monte Carlo sampling used to feed the SA algorithm.

A second method, developed in [20, 16], is a perturbative approach taking advantage of a stochastic expansion of $\{\phi^\star(\theta), \theta \sim \pi\}$ that is available when θ has small variations (a restriction that we do not need or want to impose in our case).

A third strategy, which dates back to Wiener [29] and has been developed in the fields of engineering and UQ in the 2000s (see [12, 19] and references therein), is based on chaos expansions. This technique, also known as the spectral method, consists in projecting the unknown function $\phi^\star : \Theta \mapsto \mathbb{R}^q$ on an orthonormal basis $\{\theta \mapsto B_i(\theta), i \in \mathbb{N}\}$ of the real-valued and squared-integrable (w.r.t. π) functions, computing the $\mathbb{R}^q$-valued coefficients $\{u_i, i \geq 0\}$ of $\phi = \phi^\star$ in its decomposition

$$\phi = \sum_{i \geq 0} u_i B_i. \tag{1.2}$$

In the most common case where $B_0 \equiv 1$, note that the expectation and the variance-covariance matrix are related to the $\mathbb{R}^q$-valued projection coefficients $\{u_i, i \geq \mathbb{N}\}$ through

$$\mathbb{E}_{\theta \sim \pi}[\phi(\theta)] = u_0 \text{ and } \mathbb{V}\mathrm{ar}_{\theta \sim \pi}(\phi(\theta)) = \sum_{i \geq 1} u_i u_i^\top.$$

Higher order moments are also explicitly related to the coefficients, in the case of polynomial basis (see [19, Appendix C]). This provides a tool to approximate the first and second moments of $\{\phi^\star(\theta), \theta \sim \pi\}$ from an approximation $\{\widehat{u}_i, i \geq 0\}$ of the coefficients $\{u_i, i \geq 0\}$. For an approximation of more general statistics, one can define an approximation $\widehat{\phi^\star}$ of the function $\phi^\star$ by using $\{\widehat{u}_i, i \geq 0\}$, sample independent and identically distributed (i.i.d.) variables θ_j's under the distribution π, compute $\widehat{\phi^\star}(\theta_j)$'s and obtain the empirical estimators.

When the function ϕ is explicitly known, estimating individual coefficients u_i in (1.2) is straightforward by MC simulations; nevertheless, the global convergence of a method where more and more coefficients are computed by Monte Carlo is subject to a nontrivial tuning of the speeds at which the number of coefficients and the number of simulations go to infinity (see [13]). In our case, the function $\phi = \phi^\star$ is not known explicitly. In [17], the authors provide a finite dimensional procedure to approximate the function $\phi^\star$ minimizing $\phi \mapsto \int_\Theta L(\phi(\theta), \theta)\pi(\mathrm{d}\theta)$, for some explicit function L, and they analyze the error due to finite dimensional truncation. The main drawback of this approach, is that it requires to fix the dimension in advance.

Our methodological contribution is the derivation of an algorithm for UQ analysis of SA limits: it is based on chaos expansion, it is able to overcome the finite dimension truncation by increasing the dimension, and it only uses Monte Carlo sampling for feeding the SA algorithm. Our method is among the family of infinite dimensional SA algorithms.

There exists a large number of works on such SA methods. Recently, numerous works have been devoted to statistical learning in Hilbert spaces, in particular, reproducing kernel Hilbert spaces (RKHS, see e.g. [9] and references therein). However, statistical learning in Hilbert spaces reduces to finite dimensional SA: based on N input/output examples $\{(x_i, y_i), 1 \leq i \leq N\}$, it consists in solving

$$\mathrm{argmin}_{\varphi \in \mathcal{H}} \frac{1}{N} \sum_{i=1}^{N} \left(L(\varphi(x_i), y_i) + \Omega(\varphi)\right) \tag{1.3}$$

where $\mathcal{H}$ is a RKHS associated to a positive-definite real-valued kernel K, L is a non-negative loss function and $\Omega(f)$ is a penalty term. By the Representer Theorem, the solution admits a representation of the form $\phi^\star = \sum_{i=1}^N \omega_i K(\cdot, x_i)$ so that, under regularity conditions on L and Ω, the solution of (1.3) can be solved by a SA algorithm in $\mathbb{R}^N$.

In [28], [6] and [31], the authors study SA in Hilbert spaces in the case $H(z, V) = \tilde{H}(z) + V$, where z here lives in a Hilbert space. The conditions of convergence are infinite dimensional formal analogues of those in the finite dimensional case (see Remark 3.1). Unfortunately, although interesting from a theoretical point of view, these SA algorithms are defined directly in the infinite dimensional Hilbert space, so that they are not feasible in practice.

An alternative is to iteratively increase the dimension, to ∞ in the limit, while remaining finite at each iteration. There have been several papers in this direction, generally known as the sieve approach. [14] proves almost-sure convergence in the norm topology for a modified Kiefer-Wolfowitz (see [15]) procedure in infinite dimensional Hilbert space using a sieve approach. [23] shows asymptotic normality for a modified sieve-type Robbins-Monro procedure. [30] proves almost-sure convergence in the weak topology for a sieve-type Robbins-Monro procedure. The latter three papers treat specific expressions of $H : \mathcal{H} \times \mathcal{V} \to \mathcal{H}$ while [8] combines

the unrealistic approach (as in [28, 6, 31]) with the sieve approach (as in [23, 14, 30]), deriving results on the convergence and asymptotic normality for SA with growing dimension in a quite general setting.

However, none of these literatures is adapted for dealing with UQ. Indeed, all of these previous papers in a infinite dimensional Hilbert space $\mathcal{H}$ solve problems of the form

$$\text{find } \phi^\star \in \mathcal{H} : \qquad \int H(\phi^\star, v)\mu(\mathrm{d}v) = 0,$$

so that, first, the distribution of V does not account for the uncertainty and, second, for any v, the computation of the quantity $H(\phi, v)$ may have a prohibitive computational cost depending on how a function ϕ appears in the definition of H.

Our SA algorithm in this paper combines *(i)* the sieve approach in the special case $\mathcal{H}$ is L^2_π for some probability distribution π, *(ii)* the UQ framework by allowing μ to depend upon $\theta \in \Theta$, $\theta \sim \pi$ (see Eq. (1.1)), and *(iii)* a tractable computational cost of the function H (see Eq. (1.1)). In the special case where μ does not depend on θ and $\mathcal{H} = L^2_\pi$, it can be compared to the method in [8]: see the TRMP algorithm in [8, Section II]. But our method is able to address the case when the underlying scalar product of L^2_π is not explicit. Moreover, our proof of convergence - while addressing the more general framework where the scalar product is approximated - relies on weaker assumptions (see Remark 3.1, where we show that, under our assumptions for convergence, the assumptions of [8] may fail to hold). As a result:

- A fully constructive, easy to implement, algorithm for UQ analysis of SA limits in a chaos expansion setup is obtained. It is dubbed USA (Uncertainty for Stochastic Approximation);
- A convergence proof is provided under easy-to-check hypotheses, in terms of underlying problems corresponding to fixed values of θ, avoiding conditions involving Hilbert space notions that are often hard to check in practice;
- Complexity issues, extensive reports and discussion on numerical tests are provided.

The paper is outlined as follows. USA is introduced in Section 2. Section 3 states the almost-sure convergence of USA and its L^p convergence with respect to the underlying Hilbert space norm. The proof is deferred to Section 4. Section 5 presents the results of numerical experiments, including a detailed discussion of the choice of the design parameters.

Note that beyond model uncertainty, applications of our approach include sensitivity analysis with respect to θ, or quasi-regression of an unknown function (see [1]), for instance in the context of outer Monte Carlo computations involving some unknown inner function $\phi^\star = \phi^\star(\theta)$. These applications are developed in the companion paper [4].

2. Problem Formulations and Algorithmic Solutions.

Let $\mathcal{V}$ be a metric space endowed with its Borel σ-field, Θ be a measurable subset of $\mathbb{R}^d$, and $H : \mathbb{R}^q \times \mathcal{V} \times \Theta \to \mathbb{R}^q$. Let π be a probability distribution on Θ and μ be a transition kernel from Θ to $\mathcal{V}$. For any measurable functions $f, g : \Theta \to \mathbb{R}$, we define the scalar product induced by π by

$$\langle f; g\rangle_\pi := \int_\Theta f(\theta)g(\theta)\pi(\mathrm{d}\theta). \tag{2.1}$$

By extension, for measurable functions $f = (f_1, \cdots, f_q) : \Theta \to \mathbb{R}^q$ and $g : \Theta \to \mathbb{R}$, we write in vector form

$$\langle f; g \rangle_\pi := \begin{bmatrix} \langle f_1; g \rangle_\pi \\ \cdots \\ \langle f_q; g \rangle_\pi \end{bmatrix}. \tag{2.2}$$

We denote by L^2_π the Hilbert space of functions $f : \Theta \to \mathbb{R}^q$ such that the norm $\|f\|_\pi := \sqrt{\sum_{i=1}^q \langle f_i; f_i \rangle_\pi}$ is finite. In the special case where $q = 1$, we write $L^{2,1}_\pi$.

UQ for SA limits consists in the study of the distribution of $\{\phi^\star(\theta), \theta \sim \pi\}$ where

$$\phi^\star \text{ in } L^2_\pi \text{ such that } \int_{\mathcal{V}} H(\phi^\star(\theta), v, \theta)\mu(\theta, \mathrm{d}v) = 0, \qquad \pi\text{-a.s.} \tag{2.3}$$

Hereafter, all the random variables are defined on the same probability space with probability and expectation resp. denoted by $\mathbb{P}$ and $\mathbb{E}$.

2.1. The naive SA algorithm. A naive approach for solving (2.3) would be to calculate $\phi^\star(\theta)$ for each value of θ separately, for example by the following standard SA scheme (see [5, 10, 18]): Given a deterministic sequence $\{\gamma_k, k \in \mathbb{N}\}$ of positive step sizes and a sequence of i.i.d. r.v. $\{V_k, k \in \mathbb{N}\}$ sampled from $\mu(\theta, \cdot)$, obtain $\phi^\star(\theta)$ as the limit of an iterative scheme

$$\phi^{k+1}(\theta) = \phi^k(\theta) - \gamma_{k+1} H(\phi^k(\theta), V_{k+1}, \theta). \tag{2.4}$$

Explicit conditions can be formulated to the effect that $\phi^\star(\theta) = \lim_k \phi^k(\theta)$ holds $\mathbb{P}$-a.s. (see e.g. [10, Chapter 1]). When $\lim_k k\gamma_k = O(1)$, the error $\mathbb{E}\left[|\phi^k(\theta) - \phi^\star(\theta)|^2\right]$ after k iterations (and thus k Monte Carlo samples) is $O(1/k)$ (see [10, Chapter 2]) - here the expectation is w.r.t. the random variable $\phi_k(\theta)$, for any fixed value θ. When $\lim_k k\gamma_k = +\infty$, the same rate of convergence can be reached by replacing $\phi^k(\theta)$ with an averaged estimator (see [24]).

However, except in the case where Θ is finite with few elements, the estimation of $\phi^\star(\theta)$, separately for each $\theta \in \Theta$, is unfeasible (or too demanding computationally).

2.2. Chaos Extension Setup. Let $\{\theta \mapsto B_i(\theta), i \in \mathbb{N}\}$ be an orthonormal basis of $L^{2,1}_\pi$ (for the scalar product (2.1)). Orthonormal polynomials are natural candidates, but there are other possibilities. For examples of the most commonly used bases, we refer the reader to [7, Chapter 2]. For $x, y \in \mathbb{R}^q$ we denote by $x \cdot y$ and $|x|$ the scalar product and the Euclidean norm in $\mathbb{R}^q$. We denote by l^2 the normed vector space of the $\mathbb{R}^q$-valued sequences $\{u_i, i \in \mathbb{N}\}$ with $\|u\|^2_{l^2} := \sum_{i \geq 0} |u_i|^2 < +\infty$. As is well known, given an orthonormal basis $\{B_i, i \in \mathbb{N}\}$ in $L^{2,1}_\pi$, any function $\phi \in L^2_\pi$ is characterized by a sequence $\{u_i, i \in \mathbb{N}\}$ in l^2 such that $\phi = \sum_{i \geq 0} u_i B_i$. Throughout the paper, we use the natural isomorphism $\mathbf{Is} : l^2 \to L^2_\pi$ given by

$$\phi = \mathbf{Is}(u) = \sum_{i \geq 0} u_i B_i, \text{ i.e. } u_i = \langle \phi; B_i \rangle_\pi \text{ for each } i \in \mathbb{N}, \tag{2.5}$$

and the corresponding isometry $\|\phi\|_\pi = \|u\|_{l^2}$ (see [21, Proposition 10.32]).

Assuming $\phi^\star \in L^2_\pi$, an alternative strategy for solving (2.3) would consist in the estimation of the $\mathbb{R}^q$-valued coefficients $\{u_i^\star, i \in \mathbb{N}\}$ of $\phi^\star$, combined with a truncation at a fixed level m of the expansion (2.5) and a Monte Carlo approximation of the coefficients $\{u_i^\star, i \leq m\}$, i.e., for $i \leq m$,

$$u_i = \langle \phi; B_i \rangle_\pi \approx \hat{u}_i := \frac{1}{M} \sum_{k=1}^{M} \widehat{\phi(\theta_{k,i})}\, B_i(\theta_{k,i}), \tag{2.6}$$

where $\{\theta_{k,i}, k \in \mathbb{N}, i \leq m\}$ are i.i.d. with distribution π and $\widehat{\phi(\theta_{k,i})}$ is an approximation of $\phi(\theta_{k,i})$. Let us discuss the computational cost of this approach, in the case $q = 1$ for ease of notation (and dimension d of θ). In the case of a Jacobi polynomial basis, the following control on the truncation error of ϕ holds (see [11, Theorem 6.4.2] or [7, Chapter 5]):

$$\left\| \sum_{i>m} u_i B_i \right\|_\pi^2 = O\left(m^{-\frac{2(\eta-1)}{d}} \right), \tag{2.7}$$

where η is the order of continuous differentiability of ϕ (in some cases the order may be strengthened to $O(m^{-\frac{2\eta}{d}})$). Neglecting the error associated with the approximation $\widehat{\phi(\theta_{k,i})} \approx \phi(\theta_{k,i})$, we have

$$\mathbb{E}\left[\left\| \sum_{i=0}^{m} (u_i - \hat{u}_i) B_i \right\|_\pi^2 \right] = O\left(\frac{m}{M} \right). \tag{2.8}$$

For balancing the error components (2.7) and (2.8), we must set $M = m^{1+\frac{2(\eta-1)}{d}}$. To reach a precision ϵ, m has to increase as $\epsilon^{-d/(2(\eta-1))}$ and M has to increase as $\epsilon^{-(1+d/(2(\eta-1)))}$. The computational cost in terms of number of Monte Carlo samples to estimate m coefficients is therefore $\epsilon^{-(1+d/(\eta-1))}$. This quantity suffers from the curse of dimensionality, which makes this approach fairly inefficient when combined with a nested procedure for the computation of $\widehat{\phi(\theta_{k,i})}$, e.g. through (2.4) if $\phi = \phi^\star$.

2.3. The USA Algorithm. Through the isomorphism (2.5), the problem (2.3) can be restated on l^2 as

$$u^\star \text{ in } l^2 : \quad \int_{\mathcal{V}} H\left(\sum_{i \geq 0} u_i^\star B_i(\theta), v, \theta \right) \mu(\theta, \mathrm{d}v) = 0, \ \pi\text{-a.s.}. \tag{2.9}$$

Note that, when $h(\phi^\star(\cdot), \cdot) \in L^2_\pi$ (see condition 3 below), the problem (2.3) is equivalent to finding $\phi^\star \in L^2_\pi$ such that

$$\phi^\star \in L^2_\pi : \quad \int_\Theta \left(\int_{\mathcal{V}} H\left(\phi^\star(\theta), v, \theta\right) \ \mu(\theta, \mathrm{d}v) \right) B_i(\theta) \pi(\mathrm{d}\theta) = 0_{\mathbb{R}^q}, \qquad \forall i \in \mathbb{N}. \tag{2.10}$$

This observation can be used for devising an original SA scheme for the $u_i^\star$ in (2.9). A first attempt in this direction is to restrict the problem to a set of functions ϕ of the form $\sum_{i=0}^m u_i B_i$,

for some fixed $m \in \mathbb{N}$. In this case an SA algorithm for the computation of $\{u_i^\star, i \leq m\}$ is given by

$$u_i^{k+1} = u_i^k - \gamma_{k+1}\, H\left(\sum_{j=0}^{m} u_j^k B_j(\theta_{k+1}), V_{k+1}, \theta_{k+1}\right) B_i(\theta_{k+1}), \quad i = 0, \cdots, m, \tag{2.11}$$

where the $\{(\theta_k, V_k), k \geq 0\}$ are i.i.d. with distribution $\pi(\mathrm{d}\theta)\mu(\theta, \mathrm{d}v)$. However, in practice, $\phi^\star$ is typically not of the form $\sum_{i=0}^m u_i B_i$ and, even when it is, we may not know for which m. We emphasize that, in the general case $\phi^\star \in L^2_\pi$, as the first argument of H in (2.11) is the current truncation $\sum_{i=0}^m u_i^k B_i(\theta_{k+1})$ and not $\phi^\star(\theta_{k+1})$, this algorithm does not converge to the projection of $\phi^\star$ onto the space spanned by $\{B_0, \cdots, B_m\}$. See the numerical evidence reported in Section 5.4.

Our algorithm is able to tackle the infinite dimensionality of the problem space L^2_π on which the problem is stated by increasing m along iterations, in order to recover in the limit the full sequence of the coefficients $\{u_i^\star, i \geq 0\}$ defining a solution $\phi^\star = \sum_i u_i^\star B_i$. Toward this aim, we introduce a sequence $\{m_k, k \in \mathbb{N}\}$ which specifies the number of coefficients u_i that are updated at the iteration k. The sequence $\{m_k, k \in \mathbb{N}\}$ is non-decreasing and converges to ∞.

The USA algorithm corresponds to the update of the sequence $\{u_i^k, i \geq 0\}$ through the following SA scheme, where $\Pi_{\mathcal{A}}$ denotes the projection on a suitable convex subset $\mathcal{A}$ of l^2. See Algorithm 2.1.

Algorithm 2.1 The USA algorithm for the coefficients of the basis decomposition of $\phi^\star$.

Input: Sequences $\{\gamma_k, k \geq 1\}$, $\{m_k, k \geq 1\}$, $\{M_k, k \geq 1\}$, $K \in \mathbb{N}$, $\{u_i^0, i = 0, \ldots, m_0\}$, a convex set $\mathcal{A} \subseteq l^2$.
for $k = 0, \ldots, K-1$ **do**
  for $s = 1 \ldots, M_{k+1}$ **do**
    Sample independently $(\theta_{k+1}^s, V_{k+1}^s)$ under the distribution $\pi(\mathrm{d}\theta)\mu(\theta, \mathrm{d}v)$
  end for
end for
For all $i > m_{k+1}$ define $u_i^k = 0$
for $i = 0, \ldots, m_{k+1}$ **do**
  $\hat{u}_i^{k+1} = u_i^k - \gamma_{k+1} M_{k+1}^{-1} \sum_{s=1}^{M_{k+1}} H\left(\sum_{j=0}^{m_k} u_j^k B_j(\theta_{k+1}^s), V_{k+1}^s, \theta_{k+1}^s\right) B_i(\theta_{k+1}^s)$
  $u^{k+1} = \Pi_{\mathcal{A}}(\hat{u}^{k+1})$
end for
return The vector $\{u_i^K, i = 0, \ldots, m_K\}$.

The motivations for the introduction of the projection set $\mathcal{A}$ and for the averaging over M_k draws at step k are discussed resp. in Sections 3.2 and 5.5.3.

The inputs of the algorithm are: a positive stepsize sequence $\{\gamma_k, k \geq 1\}$; two positive integer-valued sequences $\{m_k, k \geq 1\}$ and $\{M_k, k \geq 1\}$ corresponding to the number of non null coefficients in the approximation of $\phi^\star$ and to the number of Monte Carlo draws of the pair (θ, V) at each iteration k; an initial value $u_0 \in \mathbb{R}^{m_0}$; a total number of iterations K; a convex subset $\mathcal{A}$ of l^2 on which to project each newly updated sequence of coefficients.

The output of the algorithm is a sequence $u^K = \{u_i^K, i \le m_K\}$ approximating a solution $u^\star$ to the problem (2.9). The associated function

$$\phi^K := \sum_{i=0}^{m_K} u_i^K B_i. \tag{2.12}$$

is an approximation of a solution $\phi^\star$ to the problem (2.3).

3. The USA Algorithm Converges.

3.1. Assumptions.

Let

$$\mathcal{T}^\star := \left\{ u^\star \in l^2; \qquad \int_{\mathcal{V}} H\left(\sum_{i\ge 0} u_i^\star B_i(\theta), v, \theta \right) \mu(\theta, \mathrm{d}v) = 0, \ \pi\text{-a.s.} \right\}. \tag{3.1}$$

Note that (2.9) is equivalent to $u^\star \in \mathcal{T}^\star$. For simplicity of presentation above, we assumed uniqueness of $\phi^\star$, i.e. a singleton $\mathcal{T}^\star$. However, the USA algorithm is proved below to converge even in the case of multiple zeros. Allowing for multiple limits is quite standard in the SA literature. From the point of view of the application to UQ, it may seem meaningless to quantify the uncertainty of a non-uniquely defined quantity. However, enabling multiple limits appears to be the right setting when some components of the vector-valued function $\phi^\star(\cdot) \in \mathbf{Is}(\mathcal{T}^\star)$ are uniquely defined whereas some others are not. This encompasses the important case of computing quantiles and average quantiles (cf. [3]) of a (uncertain) distribution: the SA approximation for the quantile component may converge to several limits (especially when the distribution has atoms), while, for the average quantile component, the limit is unique (see e.g. [4, Lemma A.1.]).

C1. *The set $\mathcal{T}^\star$ is compact and non-empty.*

C2. *$\{M_k, k \ge 1\}$ and $\{m_k, k \ge 1\}$ are deterministic sequences of positive integers; $\{\gamma_k, k \ge 1\}$ is a deterministic sequence of positive real numbers such that, for some $\kappa > 0$,*

$$\sum_{k\ge 1} \gamma_k = +\infty, \ \sum_{k\ge 1} \gamma_k^{1+\kappa} < +\infty, \ \sum_{k\ge 1} \gamma_k^2 \frac{Q_{m_k}}{M_k} < +\infty, \ \sum_{k\ge 1} \gamma_k^{1-\kappa} q_{m_k} < +\infty, \tag{3.2}$$

where the sequences $\{q_m, m \in \mathbb{N}\}$ and $\{Q_m, m \in \mathbb{N}\}$ are defined by

$$q_m := \sup_{u^\star \in \mathcal{T}^\star} \sum_{i>m} |u_i^\star|^2, \qquad Q_m := \sup_{\theta\in\Theta} \sum_{i\le m} |B_i(\theta)|^2. \tag{3.3}$$

Since $\mathcal{T}^\star$ is compact, we have $\lim_m q_m = 0$ (cf. the proof of Lemma B.1). Assumption C2 requires, in particular, that $Q_m < +\infty$ for any m. If Θ is bounded, then this is verified for any continuous basis functions. In the case of polynomial basis, the coefficients Q_m are related to the Christoffel functions [22].

C3. *For any $z \in \mathbb{R}^q$,*

$$\int_{\Theta\times\mathcal{V}} |H(z, v, \theta)| \, \mu(\theta, \mathrm{d}v)\pi(\mathrm{d}\theta) < \infty;$$

For any $z \in \mathbb{R}^q$ and $\theta \in \Theta$,

$$h(z,\theta) = \int_{\mathcal{V}} H(z,v,\theta)\,\mu(\theta,\mathrm{d}v)$$

exists; For any $\phi \in L^2_\pi$, the mapping $h(\phi(\cdot),\cdot) : \theta \mapsto h(\phi(\theta),\theta)$ is in L^2_π; The mapping $\phi \mapsto h(\phi(\cdot),\cdot)$ from L^2_π into itself is continuous.

C4. *For π-almost every θ, for any $z_\theta, z^\star_\theta \in \mathbb{R}^q$ such that $h(z_\theta,\theta) \neq 0$ and $h(z^\star_\theta,\theta) = 0$,*

$$(z_\theta - z^\star_\theta)\cdot h(z_\theta,\theta) > 0.$$

Remark 3.1. *Previous works on SA in a Hilbert space $\mathcal{H}$ typically require an assumption of the type*

$$\int_\Theta (\phi(\theta) - \phi^\star(\theta))\cdot \hat{h}^m(\phi(\theta),\theta)\pi(\mathrm{d}\theta) > 0, \quad \forall \phi \in L^2_\pi \setminus \boldsymbol{Is}(\mathcal{T}^\star), \phi^\star \in \boldsymbol{Is}(\mathcal{T}^\star),$$

for m large enough, where $\hat{h}^m(\phi(\cdot),\cdot)$ is the approximation of $h(\phi(\cdot),\cdot)$ using the first m elements of a basis of $\mathcal{H}$: See e.g. [8, Assumption A3P(2)], which only requires the above condition for every $\phi \neq \phi^\star$ in the vector space spanned by the first m basis functions B_i. However, even this relaxed assumption does not hold in general in our setting. As a counter-example, one may take any $\phi^\star = \sum_{i\in\mathbb{N}} u^\star_i B_i$ with non null coefficients $u^\star_i$, $h(z,\theta) = z - \phi^\star(\theta)$, and $\phi = \phi^m$ given, for every m, as the truncation

$$\phi^m := Trunc_n(\phi^\star) = \sum_{i\leq m} u^\star_i B_i$$

of order m of $\phi^\star$. Then, as $Trunc_m(\phi^m - \phi^\star) = 0_{L^2_\pi}$ (by definition of ϕ^m), we have

$$\int_\Theta (\phi^m(\theta) - \phi^\star(\theta))\cdot \hat{h}^m(\phi(\theta),\theta)\pi(\mathrm{d}\theta) = \int (\phi^m(\theta) - \phi^\star(\theta))\cdot Trunc_m(\phi^m - \phi^\star)(\theta)\pi(\mathrm{d}\theta) = 0,$$

for every m.

By contrast, C4 is the standard assumption for SA with fixed θ.

C5. *a) There exists a constant C_H such that, for any $z \in \mathbb{R}^q$,*

$$\sup_{\theta\in\Theta} \int_{\mathcal{V}} |H(z,v,\theta)|^2 \mu(\theta,\mathrm{d}v) \leq C_H(1+|z|^2).$$

b) The map from L^2_π into $\mathbb{R}$ defined by $\phi \mapsto \int_{\mathcal{V}\times\Theta} |H(\phi(\theta),v,\theta)|^2 \pi(\mathrm{d}\theta)\mu(\theta,\mathrm{d}v)$ is bounded, i.e. it maps bounded sets into bounded sets.

Note that C5-b) implies that $\phi \mapsto h(\phi(\cdot),\cdot)$ is a bounded map from L^2_π into itself.

C6. *For any $B > 0$, there exists a constant $C_B > 0$ such that, for any $(\phi,\phi^\star) \in L^2_\pi \times \boldsymbol{Is}(\mathcal{T}^\star)$ with $\|\phi - \phi^\star\|_\pi \leq B$,*

$$\int (\phi - \phi^\star)(\theta)\cdot h(\phi(\theta),\theta)\,\pi(\mathrm{d}\theta) \geq C_B \min_{\bar{\phi}\in\boldsymbol{Is}(\mathcal{T}^\star)} \left\|\phi - \bar{\phi}\right\|^2_\pi .$$

Note that the above minimum exists since $\mathbf{Is}(\mathcal{T}^\star)$ is compact, by C1.

3.2. Projection Set. We address the convergence of the algorithm 2.1 for three possible choices regarding the projection set $\mathcal{A}$ (which always includes $\mathcal{T}^\star$).

Case 3.1. *$\mathcal{A} := l^2$.*

Case 3.2. *$\mathcal{A}$ is a closed ball of l^2 containing $\mathcal{T}^\star$.*

Case 3.3. *$\mathcal{A}$ is a closed convex set of l^2 containing $\mathcal{T}^\star$, with compact intersections to closed balls of l^2.*

Note that the projection set $\mathcal{A}$ is bounded in Case 3.2 and unbounded in the two other cases (for sure in Case 3.1 and potentially in 3.3).

Case 3.1 is the most convenient from the algorithmic viewpoint since no actual projection is required. However, it requires a stronger condition C5-a) to ensure the stability and an additional assumption C6 for the convergence.

The projection on a ball $\{u \in l^2 : \|u\|_{l^2} \leq B\}$ is given simply by

$$u \mapsto \min\left(1, \frac{B}{\|u\|_{l^2}}\right) u. \tag{3.4}$$

Hence, the projection required in Case 3.2 is quite straightforward. The milder assumption C5-b) is required for the stability but one still needs C6 for the convergence.

Case 3.3 requires a potentially nontrivial projection on a closed convex set: see e.g. Example 1 below. The stronger condition C5-a) is required for both the stability and the convergence, but C6 is not needed.

We now give an example of the set $\mathcal{A}$ in Case 3.3.

Example 1. *Given a positive sequence $\{a_n, n \in \mathbb{N}\}$ such that $\sum_{i\geq 0} a_i^2 < \infty$ and an increasing sequence of non-negative integers $\{d_n, n \in \mathbb{N}\}$, define the closed convex set $\mathcal{A}$:*

$$\mathcal{A} := \left\{ u \in l^2 : \sum_{d_n \leq i < d_{n+1}} |u_i|^2 \leq a_n^2 \quad \forall n \in \mathbb{N} \right\}. \tag{3.5}$$

When $d_0 = 0$, the set $\mathcal{A}$ is a compact convex subset of l^2 (see Lemma B.1). Otherwise, it is not necessarily compact. However, the set $\mathcal{A} \cap \{u \in l^2 : \sum_{i\geq 0} u_i^2 \leq B\}$ is a compact subset for any $B > 0$ (see Corollary 2). The orthogonal projection on $\mathcal{A}$ consists in projecting $(u_{d_n}, \ldots, u_{d_{n+1}-1})$ on the ball of radius a_n for all $n \in \mathbb{N}$.

3.3. Main Result.

Theorem 1. *Assume C1 to C4 and C5-a) if $\mathcal{A}$ is unbounded or C5-b) if $\mathcal{A}$ is bounded. Let there be given i.i.d. random variables $\{(\theta_k^s, V_k^s), 1 \leq s \leq M_k, k \geq 1\}$ with distribution $\pi(\mathrm{d}\theta)\mu(\theta, \mathrm{d}v)$. Let u^K and ϕ^K be the outputs of the USA Algorithm (cf. (2.12)).*

Stability. *For any $\phi^\star \in \boldsymbol{Is}(\mathcal{T}^\star)$, $\lim_{k\to+\infty} \left\|\phi^k - \phi^\star\right\|_\pi$ exists, is finite a.s., and we have*

$$\sup_{k\geq 0} \mathbb{E}\left[\left\|\phi^k - \phi^\star\right\|_\pi^2\right] < +\infty. \tag{3.6}$$

Convergence. *In addition, in case 3.3, and in cases 3.1 and 3.2 under the additional assumption C6, there exists a random variable ϕ^∞ taking values in $\boldsymbol{Is}(\mathcal{T}^\star)$ such that*

$$\lim_{k\to\infty} \left\|\phi^k - \phi^\infty\right\|_\pi = 0 \text{ a.s. and, for any } p \in (0,2), \lim_{k\to\infty} \mathbb{E}\left[\left\|\phi^k - \phi^\infty\right\|_\pi^p\right] = 0. \tag{3.7}$$

Remark 3.2. *The standard assumption ensuring a central limit theorem (CLT) for SA algorithms in a Hilbert space (cf. [8, Assumption B3(1)] or [23, Section 3, equation 3.3]) is not satisfied in our setup: as a counter-example, one can take a function $h(z,\theta)$ such that $\partial_z h(\phi^\star(\theta),\theta) = \theta$ and any polynomial basis, due to the recurrence relations of order two that are intrinsic to such bases. The study of convergence rates and CLT for the USA algorithm is therefore a problem per se, which we leave for future research.*

4. Proof of Theorem 1. For any $z = (z_1, \cdots, z_q) \in \mathbb{R}^q$ and any real-valued sequence $p := \{p_i, i \geq 0\}$ such that $\sum_{i\geq 0} p_i^2 < \infty$ we write

$$z \otimes p := ((z_1 p_0, \cdots, z_q p_0), (z_1 p_1, \cdots, z_q p_1), \cdots) \in l^2.$$

Set $\mathbf{B}^m(\theta) := (B_0(\theta), \ldots, B_m(\theta), 0, 0, \ldots)$. Define the filtration

$$\mathcal{F}_k := \sigma\left(\theta_\ell^s, V_\ell^s, 1 \leq s \leq M_\ell, 1 \leq \ell \leq k\right), \, k \in \mathbb{N}.$$

We fix $u^\star \in \mathcal{T}^\star$, which exists by C1, and we set $\phi^\star := \mathbf{Is}(u^\star)$.

4.1. Stability. The first step is to prove that the algorithm is stable in the sense that

$$\lim_k \left\|u^k - u^\star\right\|_{l^2} \text{ exists a.s. ,} \tag{4.1}$$

$$\sup_k \mathbb{E}\left[\left\|u^k - u^\star\right\|_{l^2}^2\right] < +\infty, \tag{4.2}$$

$$\liminf_{k\to\infty} \int_\Theta (\phi^k(\theta) - \phi^\star(\theta)) \cdot h(\phi^k(\theta),\theta)\pi(\mathrm{d}\theta) = 0, \quad \text{a.s.} \tag{4.3}$$

Using the definition of u^{k+1} in the USA algorithm and the property $\Pi_{\mathcal{A}}(u^\star) = u^\star$, we obtain (recalling that, in all cases 1 to 3, $\mathcal{T}^\star \subseteq \mathcal{A}$)

$$\begin{aligned}\left\|\phi^{k+1} - \phi^\star\right\|_\pi^2 = \left\|u^{k+1} - u^\star\right\|_{l^2}^2 &= \left\|\Pi_{\mathcal{A}}(\hat{u}^{k+1}) - \Pi_{\mathcal{A}}(u^\star)\right\|_{l^2}^2 \leq \left\|\hat{u}^{k+1} - u^\star\right\|_{l^2}^2 \\ &= \left\|u^k - u^\star - \gamma_{k+1}\mathcal{H}^k - \gamma_{k+1}\eta^{k+1}\right\|_{l^2}^2,\end{aligned}$$

where

$$\begin{aligned}\mathcal{H}^k :=& \mathbb{E}\left[\frac{1}{M_{k+1}}\sum_{s=1}^{M_{k+1}} H\left(\phi^k(\theta_{k+1}^s), V_{k+1}^s, \theta_{k+1}^s\right) \otimes \mathbf{B}^{m_{k+1}}(\theta_{k+1}^s)\middle|\mathcal{F}_k\right] \\ =& \int_{\Theta\times\mathcal{V}} H\left(\phi^k(\theta), v, \theta\right) \otimes \mathbf{B}^{m_{k+1}}(\theta)\, \pi(\mathrm{d}\theta)\mu(\theta,\mathrm{d}v) = \int_\Theta h(\phi^k(\theta),\theta) \otimes \mathbf{B}^{m_{k+1}}(\theta)\, \pi(\mathrm{d}\theta), \\ \eta^{k+1} :=& \frac{1}{M_{k+1}}\sum_{s=1}^{M_{k+1}} H\left(\phi^k(\theta_{k+1}^s), V_{k+1}^s, \theta_{k+1}^s\right) \otimes \mathbf{B}^{m_{k+1}}(\theta_{k+1}^s) - \mathcal{H}^k.\end{aligned}$$

For the equivalent definitions of $\mathcal{H}^k$, we used the Fubini theorem and C3. Observe that, by definition of $\mathbf{B}^{m_k}$, $\mathcal{H}^k$ and η^{k+1} are sequences in l^2 such that $\mathcal{H}_i^k = 0_{\mathbb{R}^q}$ and $\eta_i^{k+1} = 0_{\mathbb{R}^q}$ for all $i > m_{k+1}$. Define

$$\overline{\mathcal{H}}_i^k := \begin{cases} \mathcal{H}_i^k & i \leq m_{k+1}, \\ \int_\Theta h(\phi^k(\theta), \theta) B_i(\theta) \, \pi(\mathrm{d}\theta) & i > m_{k+1}. \end{cases}$$

Recalling that $u_i^k = 0_{\mathbb{R}^q}$ for $i > m_{k+1}$, we obtain

$$\begin{aligned} \left\| u^{k+1} - u^\star \right\|_{l^2}^2 &= \left\| u^k - u^\star \right\|_{l^2}^2 - 2\gamma_{k+1} \sum_{i=0}^{m_{k+1}} (u_i^k - u_i^\star) \cdot \mathcal{H}_i^k - 2\gamma_{k+1} \sum_{i=0}^{m_{k+1}} (u_i^k - u_i^\star) \cdot \eta_i^{k+1} \\ &\quad + 2\gamma_{k+1}^2 \sum_{i=0}^{m_{k+1}} \eta_i^{k+1} \cdot \mathcal{H}_i^k + \gamma_{k+1}^2 \left\| \eta^{k+1} \right\|_{l^2}^2 + \gamma_{k+1}^2 \left\| \mathcal{H}^k \right\|_{l^2}^2 \\ &= \left\| u^k - u^\star \right\|_{l^2}^2 - 2\gamma_{k+1} \sum_{i \geq 0} (u_i^k - u_i^\star) \cdot \overline{\mathcal{H}}_i^k - 2\gamma_{k+1} \sum_{i=0}^{m_{k+1}} (u_i^k - u_i^\star) \cdot \eta_i^{k+1} \\ &\quad + 2\gamma_{k+1}^2 \sum_{i=0}^{m_{k+1}} \eta_i^{k+1} \cdot \mathcal{H}_i^k + \gamma_{k+1}^2 \left\| \eta^{k+1} \right\|_{l^2}^2 + \gamma_{k+1}^2 \left\| \mathcal{H}^k \right\|_{l^2}^2 - 2\gamma_{k+1} \sum_{i > m_{k+1}} u_i^\star \cdot \overline{\mathcal{H}}_i^k. \end{aligned} \tag{4.4}$$

C4 implies that, for each θ,

$$\sum_{i \geq 0} \left((u_i^k - u_i^\star) \cdot h\left(\phi^k(\theta), \theta\right) \right) B_i(\theta) = (\phi^k(\theta) - \phi^\star(\theta)) \cdot h\left(\phi^k(\theta), \theta\right) \geq 0.$$

Taking expectation with respect to $\theta \sim \pi$ and applying the Fubini theorem (which follows from C3), we obtain for all $k \geq 0$

$$R^k := \int_\Theta (\phi^k(\theta) - \phi^\star(\theta)) \cdot h\left(\phi^k(\theta), \theta\right) \mathrm{d}\theta = \sum_{i \geq 0} (u_i^k - u_i^\star) \cdot \overline{\mathcal{H}}_i^k \geq 0. \tag{4.5}$$

Note also that $\sum_{i=0}^{+\infty} (u_i^k - u_i^\star) \cdot \mathcal{H}_i^k \in \mathcal{F}_k$. By definition, $\mathbb{E}\left[\eta_i^{k+1} | \mathcal{F}_k \right] = 0$, so that

$$\mathbb{E}\left[\sum_{i=0}^{m_{k+1}} \eta_i^{k+1} \cdot \mathcal{H}_i^k \middle| \mathcal{F}_k \right] = 0, \qquad \mathbb{E}\left[\sum_{i=0}^{m_{k+1}} (u_i^k - u_i^\star) \cdot \eta_i^{k+1} \middle| \mathcal{F}_k \right] = 0. \tag{4.6}$$

Let us consider the term $\left\|\eta^{k+1}\right\|_{l^2}^2$. We write

$$\begin{aligned}
\mathbb{E}\Big[\left\|\eta^{k+1}\right\|_{l^2}^2|\mathcal{F}_k\Big] &\leq \mathbb{E}\left[\left\|\frac{1}{M_{k+1}}\sum_{s=1}^{M_{k+1}} H\left(\phi^k(\theta_{k+1}^s), V_{k+1}^s, \theta_{k+1}^s\right)\otimes \mathbf{B}^{m_{k+1}}(\theta_{k+1}^s) - \mathcal{H}^k\right\|_{l^2}^2|\mathcal{F}_k\right] \\
&\leq \frac{1}{M_{k+1}}\int_{\Theta\times\mathcal{V}}\left\|H\left(\phi^k(\theta), v, \theta\right)\otimes \mathbf{B}^{m_{k+1}}(\theta)\right\|_{l^2}^2 \pi(\mathrm{d}\theta)\mu(\theta,\mathrm{d}v) \\
&= \frac{1}{M_{k+1}}\int_{\Theta\times\mathcal{V}}\left|H\left(\phi^k(\theta), v, \theta\right)\right|^2\left(\sum_{i=0}^{m_{k+1}} B_i(\theta)^2\right)\pi(\mathrm{d}\theta)\mu(\theta,\mathrm{d}v) \\
&\leq \frac{Q_{m_{k+1}}}{M_{k+1}}\int_{\Theta\times\mathcal{V}}\left|H\left(\phi^k(\theta), v, \theta\right)\right|^2\pi(\mathrm{d}\theta)\mu(\theta,\mathrm{d}v). \qquad (4.7)
\end{aligned}$$

Next we consider the term $2\gamma_{k+1}\sum_{i>m_{k+1}} u_i^\star\cdot\overline{\mathcal{H}}_i^k$. By using $2ab\leq a^2+b^2$ with $a\leftarrow(\gamma_{k+1}^{1-\kappa})^{1/2}|u_i^\star|$ and $b\leftarrow(\gamma_{k+1}^{1+\kappa})^{1/2}|\overline{\mathcal{H}}_i^k|$, we have

$$\begin{aligned}
\left|2\gamma_{k+1}\sum_{i>m_{k+1}} u_i^\star\cdot\overline{\mathcal{H}}_i^k\right| &\leq \gamma_{k+1}^{1-\kappa}\left(\sum_{i>m_{k+1}}^{+\infty}|u_i^\star|^2\right)+\gamma_{k+1}^{1+\kappa}\left(\sum_{i>m_{k+1}}^{+\infty}\left|\overline{\mathcal{H}}_i^k\right|^2\right) \\
&\leq \gamma_{k+1}^{1-\kappa}q_{m_{k+1}}+\gamma_{k+1}^{1+\kappa}\left\|\overline{\mathcal{H}}^k\right\|_{l^2}^2, \qquad (4.8)
\end{aligned}$$

where we used C2 in the last inequality. Note that

$$\begin{aligned}
\left\|\overline{\mathcal{H}}^k\right\|_{l^2}^2 &= \sum_{i=0}^{+\infty}\left|\int_\Theta h(\phi^k(\theta),\theta)B_i(\theta)\ \pi(\mathrm{d}\theta)\right|^2 = \int_\Theta\left|h(\phi^k(\theta),\theta)\right|^2\ \pi(\mathrm{d}\theta) \\
&\leq \int_{\Theta\times\mathcal{V}}\left|H\left(\phi^k(\theta),v,\theta\right)\right|^2\pi(\mathrm{d}\theta)\mu(\theta,\mathrm{d}v). \qquad (4.9)
\end{aligned}$$

Combining (4.4), (4.5), (4.6), (4.7), (4.8), and (4.9), we obtain

$$\begin{aligned}
\mathbb{E}\Big[\left\|u^{k+1}-u^\star\right\|_{l^2}^2|\mathcal{F}_k\Big] &\leq \left\|u^k-u^\star\right\|_{l^2}^2 - 2\gamma_{k+1}R^k+\gamma_{k+1}^{1-\kappa}q_{m_{k+1}} \\
&+\left(\gamma_{k+1}^2+\gamma_{k+1}^{1+\kappa}+\gamma_{k+1}^2\frac{Q_{m_{k+1}}}{M_{k+1}}\right)\int_{\Theta\times\mathcal{V}}\left|H\left(\phi^k(\theta),v,\theta\right)\right|^2\pi(\mathrm{d}\theta)\mu(\theta,\mathrm{d}v). \qquad (4.10)
\end{aligned}$$

To control the integral in (4.10), we distinguish two cases.

First case: $\mathcal{A}$ is unbounded. Using C5-a) we write

$$\begin{aligned}
\int_{\Theta\times\mathcal{V}}\left|H\left(\phi^k(\theta),v,\theta\right)\right|^2\pi(\mathrm{d}\theta)\mu(\theta,\mathrm{d}v) &\leq C_H\int_\Theta\left(1+\left|\phi^k(\theta)\right|^2\right)\pi(\mathrm{d}\theta) \\
&\leq C_1\left(1+\left\|u^k-u^\star\right\|_{l^2}^2\right),
\end{aligned}$$

where $C_1 := 2C_H(1+\sup_{u^\star\in\mathcal{T}^\star}\|u^\star\|_{l^2}^2)$. Note that C_1 is finite by C1.

Second case: $\mathcal{A}$ is bounded. Note that, by definition of u^k, there exists a constant B such that a.s. $\sup_{k\geq 0}\left\|u^k\right\|_{l^2}\leq B$. Assumption C5-b) implies that, for some finite $C_2>0$,

$$\sup_{k\geq 0}\int_{\mathcal{V}\times\Theta}|H(\phi^k(\theta),v,\theta)|^2\pi(\mathrm{d}\theta)\mu(\theta,\mathrm{d}v)\leq C_2.$$

In either case, we deduce from (4.10) that

$$\mathbb{E}\Big[\left\|u^{k+1}-u^\star\right\|_{l^2}^2|\mathcal{F}_k\Big]\leq\left\|u^k-u^\star\right\|_{l^2}^2-2\gamma_{k+1}R^k+\gamma_{k+1}^{1-\kappa}q_{m_{k+1}}+\left(\gamma_{k+1}^2+\gamma_{k+1}^{1+\kappa}+\gamma_{k+1}^2\frac{Q_{m_{k+1}}}{M_{k+1}}\right)(C_1\vee C_2)\left(1+\left\|u^k-u^\star\right\|_{l^2}^2\right). \tag{4.11}$$

Conclusion. In view of the above controls and C2, the assumptions of the Robbins-Siegmund lemma are verified (see [26]). An application of this lemma yields that $\lim_k\left\|u^k-u^\star\right\|_{l^2}^2$ exists and $\sum_{k\geq 0}\gamma_{k+1}R^k<+\infty$ a.s.. This concludes the proof of (4.1). Taking expectations in (4.11) and applying the Robbins-Siegmund lemma to the sequence $\mathbb{E}\left[\left\|u^k-u^\star\right\|_{l^2}^2\right]$ yields (4.2). Note also that

$$R:=\liminf_{k\to+\infty}R^k=0,\quad\text{a.s.}. \tag{4.12}$$

Indeed, on the event $\{R>0\}$, there exists a finite random index K such that $R^k>R/2$ holds for any $k\geq K$, which implies that $\sum_{k\geq 0}\gamma_{k+1}R^k<+\infty$ (as, by assumption, $\sum_{k\geq 1}\gamma_k=+\infty$). Therefore $\{R>0\}\subseteq\{\sum_{k\geq 0}\gamma_{k+1}R^k<+\infty\}$, where we saw above that $\{\sum_{k\geq 0}\gamma_{k+1}R^k<+\infty\}$ is a zero probability event. Hence so is $\{R>0\}$, which proves (4.3).

We know from (4.1) that $\lim_k\left\|\phi^k-\phi'\right\|_\pi$ exists a.s. for any $\phi'\in\mathbf{Is}(\mathcal{T}^\star)$. For later use we need the existence of this limit simultaneously for all $\phi'\in\mathbf{Is}(\mathcal{T}^\star)$ with probability one. Note that $\lim_k\left\|\phi^k-\phi'\right\|_\pi$ is continuous in ϕ' (by triangle inequality). Using that $\mathbf{Is}(\mathcal{T}^\star)$ is separable as a subset of a separable Hilbert space L^2_π, we deduce that

$$\lim_k\left\|\phi^k-\phi'\right\|_\pi\text{ exists for all }\phi'\in\mathbf{Is}(\mathcal{T}^\star),\quad\text{a.s.} \tag{4.13}$$

4.2. Proof of the Almost Sure Convergence in (3.7).

Proof for Case 3.1 or Case 3.2.. Under the assumption C1, $\mathbf{Is}(\mathcal{T}^\star)$ is bounded so that, by (4.1), the random variable $B:=\sup_{\phi^\star\in\mathcal{T}^\star}\sup_k\left\|\phi^k-\phi^\star\right\|_\pi$ is finite with probability one. Since by (4.12) $\liminf_k R^k=0$, with probability one, there exists a subsequence $\{\zeta(k),k\geq 1\}$ such that $\lim_k R^{\zeta(k)}=0$.

From (4.5) and by C6 applied with $\phi\leftarrow\phi^{\zeta(k)}$ and $\phi^\star\leftarrow\mathbf{Is}(u^\star)$, there exists a positive random variable C_B (finite a.s. and independent of k by definition of the r.v. B) such that

$$R^{\zeta(k)}\geq C_B\min_{\bar\phi\in\mathbf{Is}(\mathcal{T}^\star)}\left\|\phi^{\zeta(k)}-\bar\phi\right\|_\pi^2.$$

Let $\{\bar{\phi}^k, k \geq 0\}$ be an $\mathbf{Is}(\mathcal{T}^\star)$-valued sequence such that, for all k,

$$\min_{\bar{\phi} \in \mathbf{Is}(\mathcal{T}^\star)} \left\| \phi^{\zeta(k)} - \bar{\phi} \right\|_\pi^2 = \left\| \phi^{\zeta(k)} - \bar{\phi}^k \right\|_\pi^2 .$$

Such a sequence exists since $\mathcal{T}^\star$ is compact by C1. Using that $\lim_k R^{\zeta(k)} = 0$ we obtain $\lim_k \left\| \phi^{\zeta(k)} - \bar{\phi}^k \right\|_\pi = 0$ a.s.. Since the sequence $\{\bar{\phi}^k, k \geq 0\}$ is in a compact set $\mathbf{Is}(\mathcal{T}^\star)$ (see C1), up to extraction of a subsequence it converges to a random limit $\phi^\infty \in \mathbf{Is}(\mathcal{T}^\star)$. Hence

$$\lim_k \left\| \phi^{\zeta(k)} - \phi^\infty \right\|_\pi = 0 \text{ a.s..}$$

In view of (4.13), we deduce

$$\lim_k \left\| \phi^k - \phi^\infty \right\|_\pi = \lim_k \left\| \phi^{\zeta(k)} - \phi^\infty \right\|_\pi = 0 \quad \text{a.s..}$$

This concludes the proof of (3.7).

Proof for Case 3.3.. Since by (4.12) $\liminf_k R^k = 0$ with probability one, there exists a (random) subsequence $\{\zeta(k), k \geq 1\}$ such that $\lim_k R^{\zeta(k)} = 0$ a.s. Since the sequence $\{u^{\zeta(k)}, k \geq 0\}$ is bounded in l^2 a.s.(as $\lim_k \left\| u^k - u^\star \right\|_{l^2}$ exists a.s.) and belongs to the convex set $\mathcal{A}$ by construction, hence it belongs to a compact set (see Corollary 2). Therefore we can assume (up to extraction of another subsequence) the existence of $u^\infty \in L^2_\pi$ such that $\lim_k \left\| u^{\zeta(k)} - u^\infty \right\|_{l^2} = 0$ a.s. We now prove that u^∞ is a $\mathcal{T}^\star$-valued random variable (possibly depending on the choice of $u^\star \in \mathcal{T}^\star$). Set $\phi^\infty := \mathbf{Is}(u^\infty)$ and define

$$R^\infty := \int_\Theta \left(\phi^\infty - \phi^\star\right)(\theta) \cdot h(\phi^\infty(\theta), \theta)\, \pi(\mathrm{d}\theta).$$

Then for any $j \geq 1$,

$$\begin{aligned} R^j - R^\infty = & \int_\Theta \left(\phi^j - \phi^\infty\right)(\theta) \cdot h\left(\phi^j(\theta), \theta\right) \pi(\mathrm{d}\theta) \\ & + \int_\Theta \left(\phi^\infty - \phi^\star\right)(\theta) \cdot \left(h\left(\phi^j(\theta), \theta\right) - h\left(\phi^\infty(\theta), \theta\right)\right) \pi(\mathrm{d}\theta). \end{aligned}$$

By either C5-b) or C5-a) (depending on whether $\mathcal{A}$ is bounded or not) and since $\sup_k \left\| u^k \right\|_{l^2} < \infty$ a.s., we have $\sup_k \left\| h\left(\phi^{\zeta(k)}(\cdot), \cdot\right) \right\|_\pi^2 < \infty$ a.s.. Since

$$\lim_k \left\| \phi^{\zeta(k)} - \phi^\infty \right\|_\pi = \lim_k \left\| u^{\zeta(k)} - u^\infty \right\|_{l^2} = 0, \quad \text{a.s.,}$$

hence

$$\lim_k \int_\Theta \left(\phi^{\zeta(k)} - \phi^\infty\right)(\theta) \cdot h\left(\phi^{\zeta(k)}(\theta), \theta\right) \pi(\mathrm{d}\theta) = 0, \quad \text{a.s.}$$

Furthermore, since, by C3, $\phi \mapsto h(\phi(\cdot), \cdot)$ is continuous in L^2_π, we have

$$\lim_k \int_\Theta \left(\phi^\infty - \phi^\star\right)(\theta) \cdot \left(h\left(\phi^{\zeta(k)}(\theta), \theta\right) - h\left(\phi^\infty(\theta), \theta\right)\right) \pi(\mathrm{d}\theta) = 0 \text{ a.s.}$$

Hence $0 = \lim_k R^{\zeta(k)} = R^\infty$ a.s. In view of the definition of R^∞ and of C4, we deduce that $u^\infty \in \mathcal{T}^\star$ a.s.. In view of (4.13), this implies that $\lim_k \left\| \phi^k - \phi^\infty \right\|_\pi = \lim_k \left\| \phi^{\zeta(k)} - \phi^\infty \right\|_\pi = 0$.

4.3. Proof of the L^2-Control (3.6) and of the L^p-Convergence in (3.7). The L^2-control

$$\sup_{k\geq 0}\mathbb{E}\left[\left\|\phi^k-\phi^\infty\right\|_\pi^2\right]<+\infty$$

follows directly from (4.2) and the boundedness of $\mathcal{T}^\star$ (see C1). This proves (3.6). Let $C>0$ and $p\in(0,2)$. We write

$$\mathbb{E}\left[\left\|\phi^k-\phi^\infty\right\|_\pi^p\right]=\mathbb{E}\left[\left\|\phi^k-\phi^\infty\right\|_\pi^p\ 1_{\{\|\phi^k-\phi^\infty\|_\pi>C\}}\right]+\mathbb{E}\left[\left\|\phi^k-\phi^\infty\right\|_\pi^p\ 1_{\{\|\phi^k-\phi^\infty\|_\pi\leq C\}}\right].$$

The first term on the right hand side converges to 0 as $C\to+\infty$, uniformly in k: indeed, we have

$$\mathbb{E}\left[\left\|\phi^k-\phi^\infty\right\|_\pi^p 1_{\{\|\phi^k-\phi^\infty\|_\pi>C\}}\right]\leq\frac{\sup_{l\geq 0}\mathbb{E}\left[\left\|\phi^l-\phi^\infty\right\|_\pi^2\right]}{C^{2-p}}.$$

For any fixed $C>0$, the second term converges to zero by the dominated convergence theorem. This concludes the proof of Theorem 1.

5. Numerical Investigations. This section is devoted to the numerical analysis of the convergence of the USA algorithm. Its parameterization is discussed and the sensitivity of its performance with respect to the design parameters is tested empirically.

Notably, the possibility of letting the number m_k of estimated coefficients tend to infinity appears not only as a necessary ingredient for proving the theoretical convergence (see Theorem 1), but also as an important feature for its numerical performance, regarding, in particular, the estimation of the lower order coefficients $u_i^\star$ and the mitigation of the burn-in phase. We illustrate this assertion numerically, by testing both the genuine USA algorithm with increasing m_k and the fixed dimension version with $m_k=m$ (for different values of m), respectively referred to as the "increasing m_k" and the "fixed m" algorithms henceforth. The speed of the dimension growth turns out to be a determining factor of the practical convergence rate of the algorithm. A correct tuning of this speed allows achieving the right balance between the truncation error, i.e. the error due to the non-estimation of the coefficients beyond the m_k^{th} one, and the estimation error on the "active" coefficients up to m_k. Balancing these two contributions of the error seems to be the way to reach an optimal performance of the algorithm.

5.1. Design Parameterization of the USA Algorithm. When running the USA algorithm, the user has to choose some design parameters: given a problem of the form (2.9) and the corresponding sequence $\{q_m, m\in\mathbb{N}\}$ via (3.2), the user has to choose the orthogonal basis $\{B_i(\theta), i\in\mathbb{N}\}$, which fixes in turn the sequence $\{Q_m, m\in\mathbb{N}\}$. It remains to choose $\{\gamma_k, k\in\mathbb{N}\}$, $\{m_k, k\in\mathbb{N}\}$ and $\{M_k, k\in\mathbb{N}\}$. In this section, we consider sequences of the form

$$\gamma_k=k^{-a},\quad m_k=\lfloor k^b\rfloor+1,\quad M_k=\lfloor k^p\rfloor+1, \tag{5.1}$$

for $a,p\geq 0$ and $b>0$, and we discuss how to choose these constants assuming that

$$q_m=O\left(m^{-\delta}\right),\qquad Q_m=O\left(m^{\Delta}\right), \tag{5.2}$$

for some $\delta > 0$ and $\Delta \geq 0$.

An easy calculation shows that C2 is satisfied ($\kappa > 0$ ensuring C2 exists) if

$$0 < a \leq 1, \qquad 2 - \delta b < 2a, \qquad b\Delta + 1 < 2a + p. \tag{5.3}$$

Given $\delta > 0$ and $\Delta \geq 0$, there always exist a, b, p satisfying these conditions. Figure 1 displays the lines $x \mapsto 1$, $x \mapsto 2(1-x)/\delta$ and $x \mapsto (2x-1)/\Delta$ for different values of the pair (δ, Δ) with $\Delta > 0$. The colored area corresponds to the points (a, b) satisfying the conditions (5.3) in the case $p = 0$, i.e. in the case where the numbers of Monte Carlo draws is constant over iterations. Note that this set becomes all the more restrictive that $\delta \to 0$ and $\Delta \to \infty$. Choosing $p > 0$ gives more flexibility, but it also leads to higher computational cost (since the number of Monte Carlo simulations increases along iterations, see the discussion in Section 5.5.3).

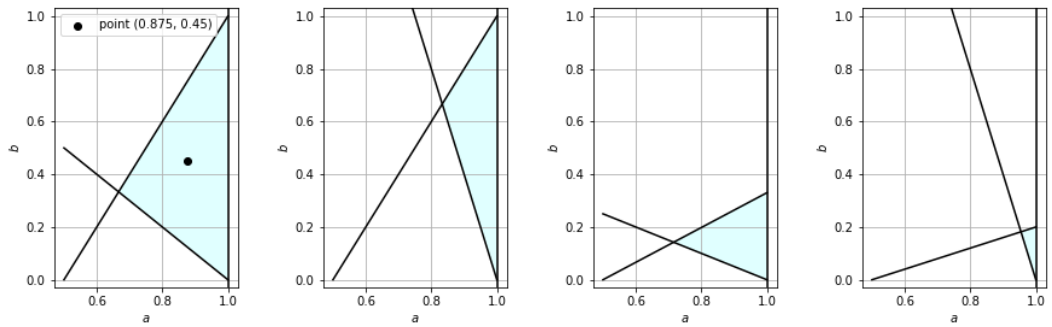

Figure 1. *For different values of (δ, Δ), in the case $p = 0$, the colored area is the admissible set of points (a, b) satisfying (5.3). From left to right: $(\delta, \Delta) = (2, 1), (0.5, 1), (4, 3)$, and $(0.5, 5)$.*

5.2. Benchmark Problem.

We consider the problem (2.9) in the case where

$$\Theta = [-\pi, \pi], \qquad \pi(\mathrm{d}\theta) = \frac{1}{2\pi}\mathbf{1}_{[-\pi,\pi]}\,\mathrm{d}\theta. \tag{5.4}$$

We perform tests for two different models of function H:

1. $H_1(z, v, \theta) := (z - \phi^\star(\theta))\left(1 + \dfrac{\cos(v)}{2}\cos(z - \phi^\star(\theta))\right) + v,$
2. $H_2(z, v, \theta) := (z - \phi^\star(\theta))\left(1 + \dfrac{\cos(v)}{2}\sin(z - \phi^\star(\theta))\right),$

for a common function $\phi^\star : \Theta \to \mathbb{R}$ given by

$$\phi^\star(\theta) := \left|\frac{4}{5} + \frac{1}{4}\exp(\sin(\theta)) - \cosh(\sin(\theta)^2)\right|(1 + \sin(2\theta)), \tag{5.5}$$

and where, for any $\theta \in \Theta$, the conditional distribution $\mu(\theta, \mathrm{d}v)$ is a centered Gaussian law with variance θ^2. In both cases we have $q = 1$ and $\mathbf{Is}(\mathcal{T}^\star) = \{\phi^\star\}$. It is easily checked that for any $z \in \mathbb{R}$, $\theta \in \Theta$ and $i = 1, 2$ we have

$$\int_{\mathcal{V}} |H_i(z, v, \theta)|^2 \mu(\theta, \mathrm{d}v) \leq 8|z - \phi^\star(\theta)|^2 + 2\theta^2, \qquad (z - \phi^\star(\theta)) \cdot h_i(z, \theta) \geq \frac{1}{2}(z - \phi^\star(\theta))^2,$$

where h_i is derived from H_i following (1.1). Hence, the assumptions C3, C4, C5, and C6 are satisfied for both models.

The two models for H above correspond to two possible behaviors of the martingale increment sequence $\{\eta^k, k \in \mathbb{N}\}$ (cf. Section 4.1): $\mathbb{E}\left[\left\|\eta^k\right\|_{l^2}\right]$ bounded away from 0 (case of Model 1) or $\mathbb{E}\left[\left\|\eta^k\right\|_{l^2}\right] \to 0$ (case of Model 2). While the first case is more general, the second one may also appear in practice and leads to quite different behavior of the USA algorithm, requiring a different tuning of the parameters.

In real-life applications, the target function $\phi^\star$ is bound to be less challenging than the present one, e.g. monotone and/or convex/concave with respect to θ or some of its components (for instance in the context of financial applications, see e.g. [4]). Moreover, the user may be interested with a few coefficients $u_i^\star$ only, whereas we show numerical results up to $m_K = 250$ below.

The choice of $\mathcal{N}(0, \theta^2)$ for the kernel $\mu(\theta, \mathrm{d}v)$ is purely illustrative. This distribution could be replaced by any other one (simulatable i.i.d.) without expectable impact regarding the qualitative conclusions drawn from the numerical experiments below.

Finally, for the orthonormal basis $\{B_i, i \in \mathbb{N}\}$, we choose the normalized trigonometric basis on $\Theta = [-\pi, \pi]$ (cf. [7, Section 2.1]). Therefore, we have $\sup_{i\in\mathbb{N}} \sup_\Theta |B_i(\theta)| < +\infty$, so that

$$Q_m = O(m),$$

i.e. $\Delta = 1$ in (5.1). Since $\phi^\star$ extended by periodicity outside $[-\pi, \pi]$ is piecewise continuously differentiable, its truncation error satisfies (5.2) with $\delta = 2$ (see Lemma A.1). Numerically, one can check that the practical rate of convergence lies somewhere between 2 and 3, i.e. the theoretical value $\delta = 2$ above is reasonably sharp (meaning that our example $\phi^\star$ is close to a "real" $\delta = 2$ example and not much "easier", which also motivated our choice of this particular function $\phi^\star$).

5.3. Performance Criteria. In the numerical experiments that follow, we compare the performances of the algorithms with increasing m_k and fixed m, for different choices of (a, b, p). The comparison relies on the root-mean-square errors, where the exact expectation is approximated by the mean value over 50 independent runs of the algorithms. After K iterations, the square of the total error $\mathcal{E}^2$ is decomposed into the mean squared SA error $\mathcal{E}_{sa}^2$, which is the error restricted to the $(m_K + 1)$ estimated coefficients, and the squared truncation error $\mathcal{E}_{tr}^2$, i.e. $\mathcal{E}^2 = \mathcal{E}_{sa}^2 + \mathcal{E}_{tr}^2$ where

$$\mathcal{E}^2 = \mathbb{E}\left[\left\|u^K - u^\star\right\|_{l^2}^2\right], \quad \mathcal{E}_{sa}^2 = \mathbb{E}\left[\sum_{i=0}^{m_K}(u_i^K - u_i^\star)^2\right], \quad \text{and} \quad \mathcal{E}_{tr}^2 = \sum_{i=m_K+1}^{+\infty}(u_i^\star)^2$$

(recalling $u_i^K = 0$ for $i > m_K$). The benchmark values for the coefficients $u_i^\star$ are pre-calculated by high-precision numerical integration. With the exception of Figures 3 and 5[left], all our graphs are error plots in log-log scale.

5.4. Impact of the Increasing Dimension. In this section, we discuss the role of the sequence $\{m_k, k \in \mathbb{N}\}$. Since $(\delta, \Delta) = (2, 1)$, the set of admissible pairs (a, b) for our benchmark problem is given by the leftmost graph of Figure 1.

We take $a = 0.875$, which is in the middle of the corresponding admissible interval. For Model 1 (i.e. $H = H_1$), a heuristic may be applied for a clever choice of b. In finite dimensional SA schemes, the squared L^2-error after k iterations is typically of the order of $\gamma_k = k^{-a}$ (see [10, Chapter 2]). For the USA algorithm, we may expect (in the case $p = 0$) a growth of the variance at least proportional to the dimension $m_k \approx k^b$. This suggests a heuristic guess for the SA-error order given by $1/k^{a-b}$. Now, by (5.2) (with in our case $\delta = 2$), the truncation error is of order $k^{-b\delta}$. Hence, to optimize the convergence rate, we take b such that $b\delta = a - b$. This approximately corresponds to $b = 0.3$, which is the value that we use in our tests for Model 1. In the case of Model 2 (i.e. $H = H_2$), this rule of thumb does not apply, because the variance of the martingale increments goes to 0. In this case we simply take $b = 0.45$, so that $(a, b) = (0.875, 0.45)$ lies in the middle of the admissible set (see Figure 1[left]). Also note that the range $[0.25, 0.5]$ for b is reasonable in view of the total number K of iterations that we commonly use in the algorithm and of the number m_K of the coefficients of interest.

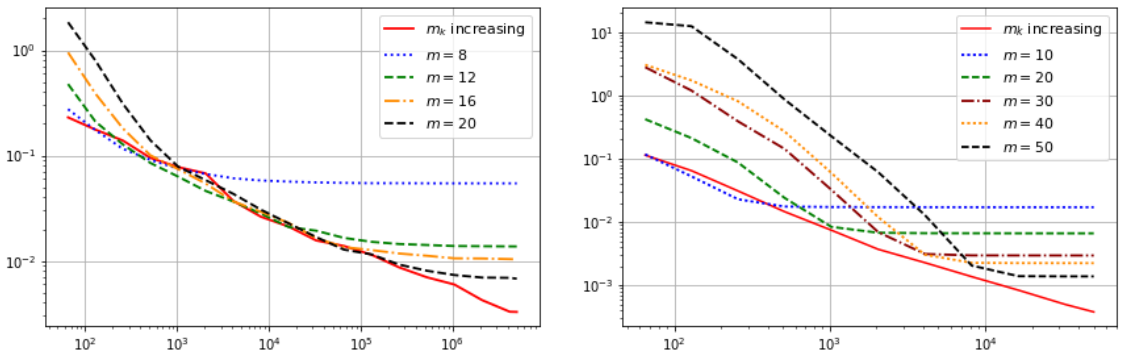

Figure 2. *The total error $\mathcal{E}$ as a function of the number of iterations, for different choices of the sequence $\{m_k, k \in \mathbb{N}\}$: m_k increasing (solid line) and: [left] for Model 1 with $m_k = m = 8, 12, 16, 20$ (other lines); [right] for Model 2 with $m_k = m = 10, 20, 30, 40, 50$ (other lines).*

Figure 2 displays the total error $\mathcal{E}$ for different strategies on the sequence $\{m_k, k \in \mathbb{N}\}$: the solid line is the case $m_k = \lfloor k^b \rfloor + 1$ (with $b = 0.3$ and 0.45 for Models 1 and 2 respectively), while the other lines correspond to the cases $m_k = m = 8, 12, 16, 20$ for Model 1 and $m_k = m = 10, 20, 30, 40, 50$ for Model 2 (larger values of m are used here since the convergence is expected to be faster for Model 2). The performance of the algorithm with increasing m_k is similar or better throughout the whole iteration path. This holds true in the burn-in phase, which is typically related to disproportion of the first values of γ_k and the magnitude of the solution (estimated coefficients). In fact, with increasing m_k, the dimension gradually grows with k, with larger values of γ_k naturally associated with the estimation of the first (and larger, for classical basis) coefficients, whereas, when $m_k = m$ is constant, the higher order "small" coefficients are involved from the very beginning along with the larger values of γ_k, leading to a longer burn-in phase. It is also true on the convergence part, where the algorithms with fixed m only converge up to a certain accuracy depending on the value of m. The better performance of the increasing m_k version is more explicit for Model 2; while for Model 1, fixed m versions have similar performance within certain ranges of values of K. However, in

practice we do not know in advance the length of the burn-in phase or the magnitude of the truncation error for various m. Hence, the genuine USA algorithm with increasing m_k, which provides optimal performance without the need for additional knowledge, is always preferable.

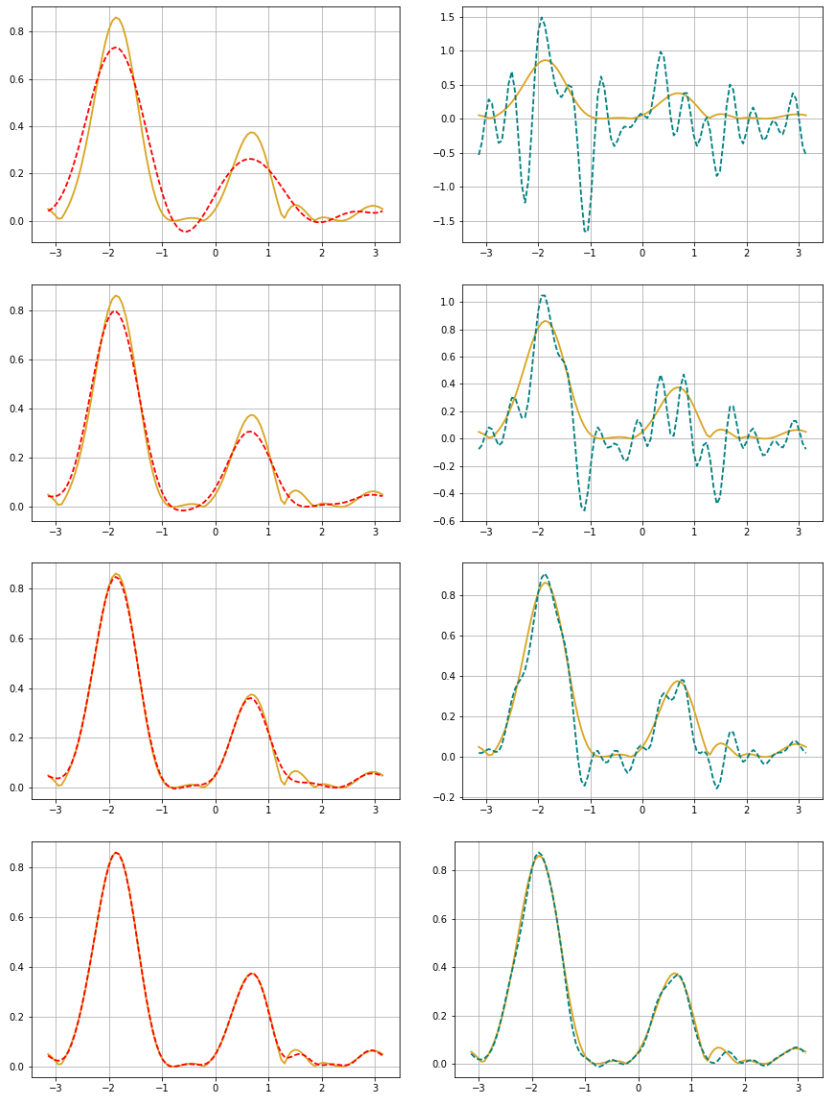

Figure 3. *The functions $\phi^\star$ and ϕ^K are displayed in respective solid line and dashed lines, as a function of $\theta \in [-\pi, \pi]$. On the left, $\{m_k, k \in \mathbb{N}\}$ is increasing and on the right, it is constant and equal to $m = 30$. From top to bottom, $K \in \{128, 256, 512, 1024\}$.*

Let us now analyze the weak burn-in phase performance of the fixed m version for Model 2 (cf. Figure 2[right]). Figure 3 displays the result of a single run of the USA algorithm in this case. In dashed line, the function $\theta \mapsto \phi^K(\theta)$ is displayed for $\theta \in [-\pi, \pi]$. For comparison, the function $\theta \mapsto \phi^\star(\theta)$ is displayed in solid line. To illustrate the advantage of the USA algorithm on the burn-in phase due to gradual dimension growth, we show the estimated function ϕ^K for

different values of K (from top to bottom, $K \in \{128, 256, 512, 1024\}$) and for m_k increasing (left panels) versus $m_k = m = 30$ for any k (right panels). The increasing dimension m_k leads to a smoother convergence, with intermediate iterations looking closer to a projection of $\phi^\star$ on the subspace spanned by a smaller number of basis functions. We conclude that in the case where the variance of the martingale increment is small or goes to 0, the progressive dimension growth plays a key role in the USA algorithm performance.

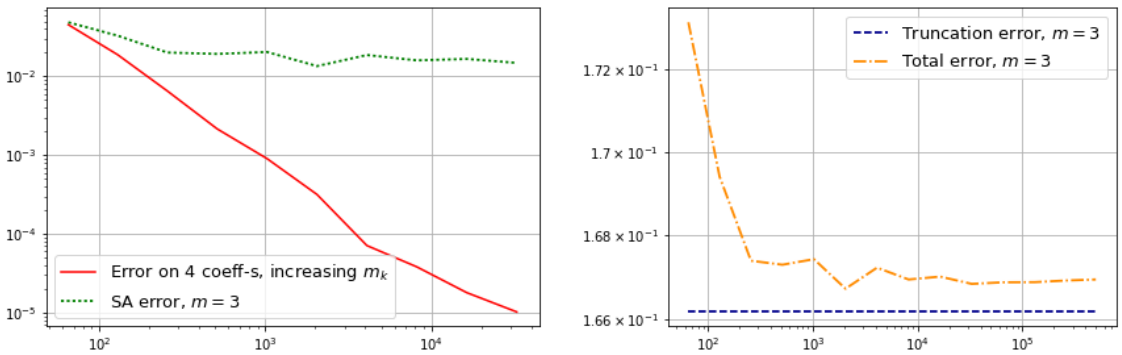

Figure 4. *Model 2: [left] in the case $m_k \to \infty$ (solid line) and $m_k = m = 3$ (dotted line), the SA error $\left(\mathbb{E}\sum_{i=0}^{3}(u_i^K - u_i^\star)^2\right)^{1/2}$ as a function of the number of iterations K; [right] in the case $m_k = m = 3$, the truncation error $\mathcal{E}_{tr}$ (dashed line) and the total error $\mathcal{E}$ (dash-dot line) displayed as a function of K.*

In Figure 4, we show that increasing m_k is also key for an accurate determination of the lower order coefficients (e.g. in the case where only the first few coefficients of the expansion of $\phi^\star$ are of interest to the user). In fact, as already mentioned in Section 2.3, the algorithm with fixed m does typically not converge to the first $(m+1)$ coefficients of the decomposition of $\phi^\star$. In Figure 4[left], the L^2-error on the first 4 coefficients is displayed as a function of the number of iterations K, for two strategies on m_k (case of Model 2): the solid line is the case $m_k = O(k^b)$ with $b = 0.45$ and the dotted line is the case $m = 3$. In Figure 4[right], the total error $\mathcal{E}$ and the truncation error $\mathcal{E}_{tr}$ are displayed, resp. in dash-dot line and dashed line in the case m_k is the constant sequence equal to $m = 3$. These figures show that, when $m_k \to +\infty$, USA converges (which is the claim of Theorem 1), whereas, when $m_k = m$ for any k, it does not: the total error does not reach the truncation error since there is a non vanishing bias on the estimation of the first $(m+1)$ coefficients (the SA-error $\mathcal{E}_{sa}$ does not vanish when $K \to +\infty$).

For Model 1, similar effects (not reported here) are visible, even if a bit less obvious due to the slower convergence of both versions (fixed and increasing dimension) of the algorithm in this case.

5.5. Impact of the Design Parameters for the Increasing m_k USA Algorithm. In this section, we discuss the impact of the choice of a, b, and p on the performance on the USA algorithm.

5.5.1. Role of b. In this paragraph, we set $a = 0.875$, $p = 0$, and we test different values of b. The range of admissible values of b is $(0.125, 0.75)$. We take $b \in \{0.2, 0.25, 0.3, 0.35, 0.4, 0.45\}$

for Model 1 and $b \in \{0.3, 0.4, 0.5, 0.6, 0.7\}$ for Model 2 (as we will see below, these values of b are enough to explain the behavior of the algorithm for both models).

Figure 5 displays the evolution of the total error $\mathcal{E}$ as a function of the number of iterations K for different values of b, and for both models of H. For Model 1, the variance increases with the dimension, which makes the SA error larger as b increases. At the same time, the truncation error decreases at the rate $b\delta$, so that for too small values of b, the truncation error dominates the SA error. Hence there is a trade-off between the two errors, with optimal values of b somewhere in the range. This phenomenon is observed on Figure 5[left, center]. For $b = 0.2, 0.25, 0.3$, the total error is dominated by the truncation error, while from $b = 0.35$ the error is dominated by the SA error so that, as b increases further, the convergence becomes slower due to additional variance which increases the SA error. For Model 2, since the variance of the martingale increments goes to 0, the effect of additional variance due to a larger dimension is not visible. We observe that larger values of b implies a better convergence, up to $b = 0.70$. However, as we may see, the gain in the rate of convergence from taking larger b decreases as it tends to the boundary of the admissible interval. In addition, this analysis in terms of the number of iterations K does not take into account the higher computational cost due to a dimension growing faster and each iteration becoming longer when b is larger. For example, for $b = 0.70$, we made only $K = 2500$ iterations, because the computational effort becomes too large beyond this. To conclude, we suggest that in this case optimal values of b (for given a) in terms of both convergence and computational cost lie near the middle of the admissible interval.

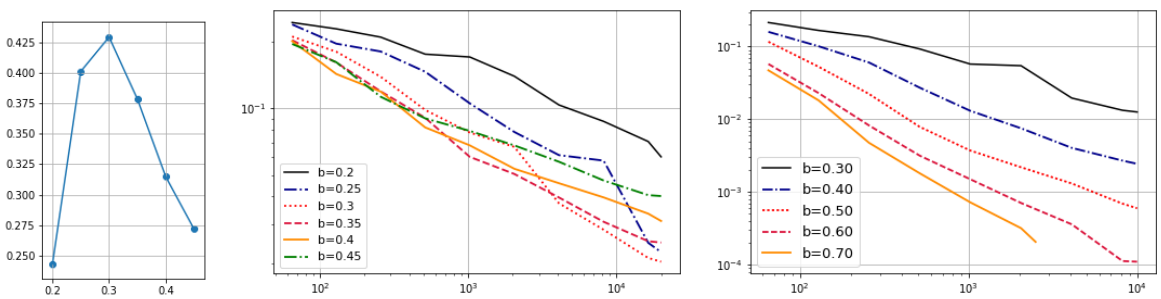

Figure 5. *[left] Empirical L^2 convergence rate for $b \in \{0.2, 0.25, 0.3, 0.35, 0.4, 0.45\}$, Model 1. Total error $\mathcal{E}$ as a function of the number of iterations K, for different values of b; [middle] Model 1, $b \in \{0.2, 0.25, 0.3, 0.35, 0.4, 0.45\}$; [right] Model 2, $b \in \{0.3, 0.4, 0.5, 0.6, 0.7\}$.*

5.5.2. Role of a. In this paragraph, again for $p = 0$, taking $b = 0.3$ for Model 1 and $b = 0.45$ for Model 2 as in Section 5.4, we compare different values of a.

Figure 6 displays the total error $\mathcal{E}$ as a function of the number of iterations K for different values of a for both models. For Model 1, we see that the convergence rate is better for larger values of a. This is in line with classical results for finite dimensional stochastic approximation, whereby the L^2 error is of order $\gamma_k^{1/2} \sim k^{-a/2}$ (see [10, Chapter 2]). For Model 2, varying a does not produce much effect (except for a slight decline as a approaches 1), because the step-size controls the variance of the corresponding martingale noise, but in the case of Model 2, the variance of the martingale increments goes to 0 anyway.

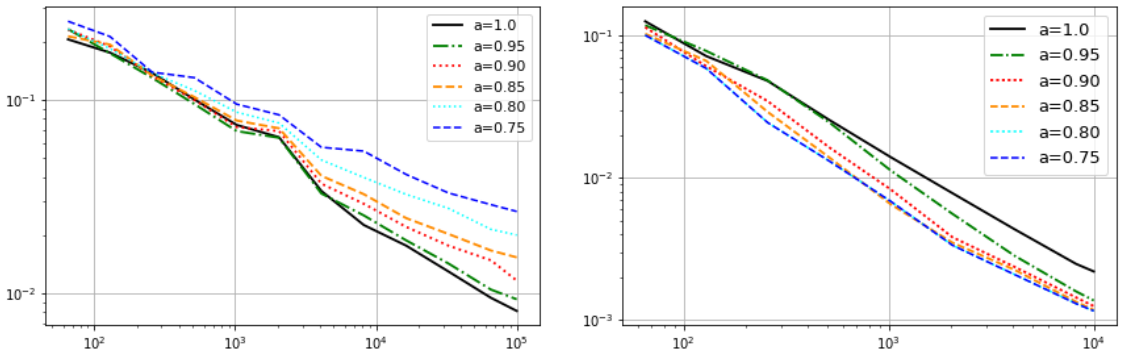

Figure 6. *The total error $\mathcal{E}$ as a function of the number of iterations, for different values of a in $\{0.75, 0.80, 0.85, 0.9, 0.95, 1.0\}$: [left] Model 1 and [right] Model 2.*

5.5.3. Role of p. In this section we consider the case $p > 0$, i.e. the number of Monte Carlo samples at each iteration increases along the USA iterations. One may check that all the triples of the parameters (a, b, p) used below lie in the admissible set (cf. (5.3)).

In the analysis that follows, we want to keep track of the dependence of the error with respect to a computational cost proxied by the total number of Monte Carlo draws of the pair (θ, v), i.e., after K iterations, $\sum_{k=0}^{K-1} M_k \approx O(K^{p+1})$. As we want to have the same dimension growth speed with respect to the computational cost for different tests, we take $b = \bar{b}(p+1)$, with $\bar{b} = 0.3$ for Model 1 and $\bar{b} = 0.45$ for Model 2.

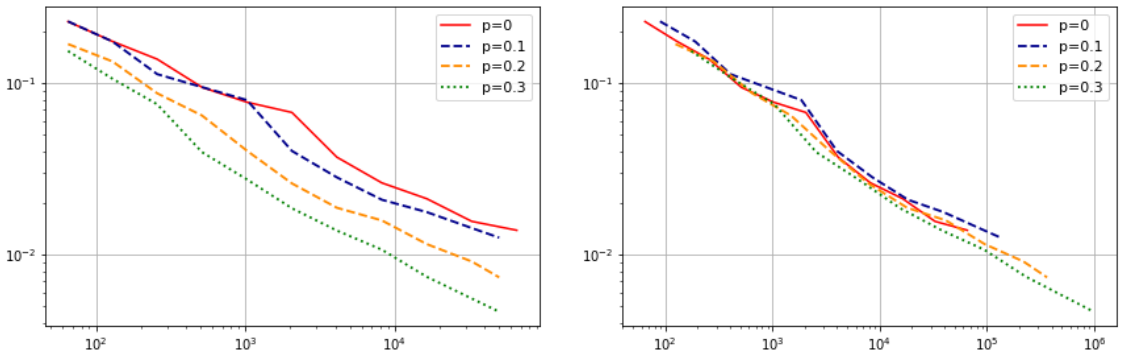

Figure 7. *Model 1, total error $\mathcal{E}$ of the USA algorithm for different values of $p \in \{0, 0.1, 0.2, 0.3\}$ as a function of the number of iterations [left] and of the total number of Monte Carlo draws [right]. Here $a = 0.875$ and $b = 0.3(p+1)$.*

We set $a = 0.875$ (as in Sections 5.4 and 5.5.1). Figure 7 displays for Model 1 the total error $\mathcal{E}$ as a function of the number of iterations (left) and as a function of the total number of Monte Carlo draws (right) for triples of the form $(a, \bar{b}(p+1), p)$ with various p. The results show that, even though larger p yield a better convergence in terms of the number of iterations K, there is no much difference when the computational cost is taken into account (i.e. in terms of the number of Monte Carlo draws). The results are very similar for Model 2 and therefore

not reported here.

5.6. Conclusion. To summarize, we observed from our experiments, that the increasing dimension feature of the USA algorithm is essential for the asymptotic convergence, as well as for the lower order coefficients estimation; it also yields a milder burn-in phase. As expected, the convergence is generally faster for Model 2 due to reduced variance effect. In the more general setting (Model 1), the parameter b plays a crucial role in the performance of the algorithm, a good rule of thumb (in the case $p = 0$) being to take b satisfying $b\delta = a - b$. However, in the special case where the variance of the martingale increments goes to 0, this rule does not work and one should take larger values of b. Empirical convergence rates are naturally bounded by $a/2$ for Model 1 (which follows from corresponding results for finite dimensional SA), while for Model 2 the convergence is much faster. Finally, taking $p > 0$ may potentially be useful for verifying the assumptions of Theorem 1 in some cases, but it makes no real difference in terms of convergence speed when the latter is assessed with respect to the total number of simulations.

Appendix A. Truncation error for trigonometric basis.

Lemma A.1. *Let $\phi : \mathbb{R} \to \mathbb{R}$ be 2π-periodic and piecewise continuously differentiable. Let $\{u_i, i \in \mathbb{N}\}$ be the coefficients of its decomposition with respect to the normalized trigonometric basis (cf. [7, Section 2.1]). Then for some $C > 0$*

$$\sum_{i=m+1}^{+\infty} |u_i|^2 \leq Cm^{-2}.$$

Proof. Consider the Fourier decomposition of the function ϕ' on $[-\pi, \pi]$:

$$\phi'(x) = v_0 + \sum_{m\geq 1} (v_{2m-1} \sin(mx) + v_{2m} \cos(mx)).$$

Using that ϕ is 2π-periodic (i.e. $\phi(-\pi) = \phi(\pi)$), it is not hard to deduce via integration by parts that, for any $m \geq 1$, $u_{2m-1} = v_{2m}/m$, and $u_{2m} = -v_{2m-1}/m$. Hence,

$$\sum_{i=2m-1}^{+\infty} |u_i|^2 \leq \frac{1}{m^2} \sum_{i=2m-1}^{+\infty} |v_i|^2 \leq \frac{\|\phi'\|_\pi^2}{m^2},$$

which implies the result. ∎

Appendix B. Compact sets in l^2.

Lemma B.1. *For a positive sequence $\{a_n, n \in \mathbb{N}\}$ such that $\sum_{i\geq 0} a_i^2 < \infty$ and an increasing sequence of non-negative integers $\{d_n, n \in \mathbb{N}\}$ such that $d_0 = 0$, the closed convex set $\mathcal{A}$:*

$$\mathcal{A} := \left\{ u \in l^2 : \sum_{d_n \leq i < d_{n+1}} |u_i|^2 \leq a_n^2 \quad \forall n \in \mathbb{N} \right\} \tag{B.1}$$

is compact.

Proof. By [17, Theorem 3] a subset $\mathcal{A}$ of l^2 is relatively compact if and only if

$$\sup_{u\in\mathcal{A}} \sum_{i\geq n} |u_i|^2 \text{ is finite for every } n \text{ and converges to 0 as } n \to +\infty.$$

For $\mathcal{A}$ given by (B.1) it is clear that for l such that $d_l \leq n$ we have

$$\sup_{u\in\mathcal{A}} \sum_{i\geq n} |u_i|^2 \leq \sup_{u\in\mathcal{A}} \sum_{i\geq d_l} |u_i|^2 \leq \sum_{j\geq l} a_j^2 \to 0$$

as $n, l \to +\infty$. Since $\mathcal{A}$ is also closed we deduce that it is compact. ∎

Corollary 2. *Let $\mathcal{A}$ be defined by (B.1) (with d_0 not necessarily 0). For any constant $B > 0$ the set $\{u \in \mathcal{A} : \|u\|_{l^2} < B\}$ is convex compact.*

Proof. The result follows directly from Lemma B.1 using that $\sum_{i<d_0} |u_i|^2 \leq B^2$ for any $u \in \mathcal{A}$. ∎

REFERENCES

[1] J. An and A. Owen, *Quasi-regression*, Journal of Complexity, 17 (2001), pp. 588–607.

[2] I. Babuska, R. Tempone, and G. Zouraris, *Galerkin finite element approximations of stochastic elliptic partial differential equations*, SIAM J. Numer. Anal., 42 (2004), pp. 800–825.

[3] O. Bardou, N. Frikha, and G. Pages, *Computing VaR and CVaR using stochastic approximation and adaptive unconstrained importance sampling*, Monte Carlo Methods and Applications, 15 (2009), pp. 173–210.

[4] D. Barrera, S. Crépey, B. Diallo, G. Fort, E. Gobet, and U. Stazhynski, *Stochastic approximation schemes for economic capital and risk margin computations*. Submitted preprint available at https://hal-univ-tlse2.archives-ouvertes.fr/IMT/hal-01710394v1, 2018.

[5] A. Benveniste, M. Metivier, and P. Priouret, *Adaptive algorithms and stochastic approximations*, vol. 22 of Applications of Mathematics, Springer-Verlag, 1990.

[6] N. Berman and A. Shwartz, *Abstract stochastic approximations and applications*, Stochastic Processes and their Applications, 31 (1989), pp. 133–149.

[7] C. Canuto, M. Hussaini, A. Quarteroni, and T. Zang, *Spectral Methods: Fundamentals in Single Domains*, Springer Verlag, 2006.

[8] X. Chen and H. White, *Asymptotic properties of some projection-based Robbins-Monro procedures in a Hilbert space*, Studies in Nonlinear Dynamics and Econometrics, 6 (2002), pp. 1–55.

[9] A. Dieuleveut and F. Bach, *Nonparametric stochastic approximation with large step-sizes*, The Annals of Statistics, 44 (2016), pp. 1363–1399.

[10] M. Duflo, *Random Iterative Models*, vol. 34 of Applications of Mathematics, Springer, 1997.

[11] D. Funaro, *Polynomial Approximation of Differential Equations*, vol. 8 of Lecture Notes in Physics Monographs, Springer-Verlag Heidelberg, 1992.

[12] R. Ghanem and P. Spanos, *Stochastic finite elements: a spectral approach*, Courier Corporation, 2003.

[13] E. Gobet and K. Surana, *A new sequential algorithm for L_2-approximation and application to Monte-Carlo integration*, Preprint available on https://hal.archives-ouvertes.fr/hal-00972016, (2014).

[14] L. Goldstein, *Minimizing noisy functionals in Hilbert space: An extension of the Kiefer-Wolfowitz procedure*, Journal of Theoretical Probability, 1 (1988), pp. 189–204.

[15] J. Kiefer and J. Wolfowitz, *Stochastic estimation of the maximum of a regression function*, Annals of Mathematical Statistics, 23 (1952), pp. 462–466.

[16] M. Kleiber and T. Hien, *The Stochastic Finite Element Method: basic perturbation technique and computer implementation*, John Wiley & Sons Ltd, 1992.

[17] A. Kulkarni and V. Borkar, *Finite dimensional approximation and Newton-based algorithm for stochastic approximation in Hilbert space*, Automatica, 45 (2009), pp. 2815–2822.

[18] H. Kushner and G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*, vol. 35 of Application of Mathematics, Springer, 1997.
[19] O. Le Maitre and O. Knio, *Spectral Methods for Uncertainty Quantification. With Applications to Computational Fluid Dynamics, Scientific Computation*, Springer Science & Business Media, 2010.
[20] W. Liu, T. Belytschko, and A. Mani, *Random field finite elements*, International journal for numerical methods in engineering, 23 (1986), pp. 1831–1845.
[21] J. Muscat, *Functional Analysis. An Introduction to Metric Spaces, Hilbert Spaces, and Banach Algebras*, Springer, 2014.
[22] P. Nevai, *Géza Freud, orthogonal polynomials and Christoffel functions. a case study*, Journal of Approximation Theory, 48 (1986), pp. 3–167.
[23] R. Nixdorf, *An invariance principle for a finite dimensional stochastic approximation method in a Hilbert space*, Journal of Multivariate Analysis, 15 (1984), pp. 252–260.
[24] B. Polyak and A. Juditsky, *Acceleration of stochastic approximation by averaging*, SIAM J. Control Optim., 30 (1992), pp. 838–855.
[25] H. Robbins and S. Monro, *A stochastic approximation method*, Annals of Mathematical Statistics, 22 (1951), pp. 400–407.
[26] H. Robbins and D. Siegmund, *A convergence theorem for nonnegative almost supermartingales and some applications*, in Optimizing Methods in Statistics, J. Rustagi, ed., Academic Press, New-York, 1971, pp. 233–257.
[27] R. Smith, *Uncertainty quantification. Theory, implementation, and applications*, vol. 12 of Computational Science & Engineering, Springer, Philadelphia, PA, 2014.
[28] H. Walk, *An invariance principle for the Robbins-Monro process in a Hilbert space*, Zeitschrift fur Warscheinlichkeitstheorie und Verwandte Gebiete, 39 (1977), pp. 135–150.
[29] N. Wiener, *The homogeneous chaos*, American Journal of Mathematics, 60 (1938), pp. 897–936.
[30] G. Yin, *On H-valued stochastic approximation: Finite dimensional projections*, Stochastic Analysis and Applications, 10 (1992), pp. 363–377.
[31] G. Yin and Y. Zhu, *On H-valued Robbins-Monro processes*, Journal of Multivariate Analysis, 34 (1990), pp. 116–140.