

A non-adapted sparse approximation of PDEs with stochastic inputs [☆]

Alireza Doostan ^{a,*}, Houman Owhadi ^b

^a Aerospace Engineering Sciences Department, University of Colorado, Boulder, CO 80309, USA

^b Applied & Computational Mathematics Department, California Institute of Technology, Pasadena, CA 91125, USA

ARTICLE INFO

Article history:

Received 10 June 2010

Received in revised form 24 October 2010

Accepted 4 January 2011

Available online 9 January 2011

Keywords:

Polynomial chaos

Uncertainty quantification

Stochastic PDE

Compressive sampling

Sparse approximation

ABSTRACT

We propose a method for the approximation of solutions of PDEs with stochastic coefficients based on the direct, i.e., non-adapted, sampling of solutions. This sampling can be done by using any legacy code for the deterministic problem as a black box. The method converges in probability (with probabilistic error bounds) as a consequence of sparsity and a concentration of measure phenomenon on the empirical correlation between samples. We show that the method is well suited for truly high-dimensional problems.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

Realistic analysis and design of complex engineering systems require not only a fine understanding and modeling of the underlying physics and their interactions, but also a significant recognition of intrinsic uncertainties and their influences on the quantities of interest. Uncertainty Quantification (UQ) is an emerging discipline that aims at addressing the latter issue; it aims at meaningful characterization of uncertainties in the physical models from the available measurements and efficient propagation of these uncertainties for a quantitative validation of model predictions.

Despite recent growing interests in UQ of complex systems, it remains a grand challenge to efficiently propagate uncertainties through systems characterized by a large number of uncertain sources where the so-called curse-of-dimensionality is yet an unsolved problem. Additionally, development of *non-intrusive* uncertainty propagation techniques is of essence as the analysis of complex multi-disciplinary systems often requires the use of sophisticated coupled deterministic solvers in which one cannot readily intrude to set up the necessary propagation infrastructure.

Sampling methods such as the Monte Carlo simulation and its several variants had been utilized for a long time as the primary scheme for uncertainty propagation. However, it is well understood that these methods are generally inefficient for large-scale systems due to their slow rate of convergence. There has been an increasing recent interest in developing alternative numerical methods that are more efficient than the Monte Carlo techniques. Most notably, the stochastic Galerkin schemes using Polynomial Chaos (PC) bases [38,28,67,2,63] have been successfully applied to a variety of engineering problems and are extremely useful when the number of uncertain parameters is not large. In their original form, the stochastic Galerkin schemes are *intrusive*, as one has to modify the deterministic solvers for their implementation. *Stochastic collocation* schemes [58,45,66,1,51] belong to a different class of methods that rely upon (isotropic) sparse grid integration/interpolation in the stochastic space of the problem to reduce the curse-of-dimensionality associated with the conventional

[☆] A preliminary version of this work has appeared in a 2009 CTR Annual Research Brief report [35].

* Corresponding author. Tel.: +1 303 492 7572; fax: +1 303 492 4990.

E-mail addresses: doostan@colorado.edu (A. Doostan), owhadi@caltech.edu (H. Owhadi).

tensor-product integration/interpolation rules. As their construction is primarily based on the input parameter space, the computational cost of both stochastic Galerkin and collocation techniques increases rapidly for large number of independent input uncertainties.

More recently, efforts have been made to construct *solution-adaptive* uncertainty propagation techniques that exploit any structures in the solution to decrease the computational cost. Among them are the multi-scale model reduction of [33] and the sparse decomposition of [60,8,6,7] for the stochastic Galerkin technique, adaptive sparse regression approaches of [10,11] for PC expansions, anisotropic and adaptive sparse grids of [50,44] for the stochastic collocation scheme, and the low-rank solution approximations of [52,53,34].

In the present study, we are interested in cases where the quantity of interest is *sparse* at the stochastic level, i.e., it can be accurately represented with only *few* terms when linearly expanded into a stochastic, e.g., polynomial chaos, basis. Interestingly, sparsity is salient in the analysis of high-dimensional problems where the number of energetic basis functions (those with large coefficients) is small relative to the cardinality of the full basis. For instance, it has been shown in [60,8] that, under some mild conditions, solutions to linear elliptic stochastic PDEs with high-dimensional random coefficients admit sparse representations with respect to the PC basis. Consequently, an approach based on a zero-dimensional algebraic stochastic problem has been proposed in [8] to detect the sparsity pattern, which then guides the stochastic Galerkin analysis of the original problem. Moreover, a “quasi”-best N -term approximation for a class of elliptic stochastic PDEs has been proposed in [7].

In this work, using *concentration of measure* inequalities and *compressive sampling* techniques, we derive a method for PC expansion of sparse solutions to stochastic PDEs. The proposed method is

- Non-intrusive: it is based on the direct random sampling of the PDE solutions. This sampling can be done by using any legacy code for the deterministic problem as a black box.
- Non-adapted: it does not tailor the sampling process to identify the important dimensions at the stochastic level.
- Provably convergent: we obtain probabilistic bounds on the approximation error proving the stability and convergence of the method.
- Well-suited to problems with high-dimensional random inputs.

Compressive sampling is an emerging direction in signal processing that aims at recovering *sparse* signals accurately (or even exactly) from a small number of their random projections [20,21,17,30,15,18,16,22,14]. A sparse signal is simply a signal that has only few significant coefficients when linearly expanded into a basis, e.g., $\{\psi_\alpha\}$.

For sufficiently sparse signals, the number of samples needed for a successful recovery is typically less than what is required by the Shannon–Nyquist sampling principle. Generally speaking, a successful signal reconstruction by compressive sampling is conditioned upon:

- *Sufficient* sparsity of the signal; and
- *Incoherent* random projections of the signal.

A square-measurable stochastic function $u(\omega)$, defined on a suitable probability space $(\Omega, \mathcal{F}, \mathcal{P})$ can be expanded into a mean-squared convergent series of the chaos polynomial bases, i.e., $u(\omega) \approx \sum_\alpha c_\alpha \psi_\alpha(\omega)$, with some cardinality P . The stochastic function $u(\omega)$ is then sparse in PC basis $\{\psi_\alpha\}$, if only a small fraction of coefficients c_α are significant. In this case, under certain conditions, the sparse PC coefficients \mathbf{c} may be computed accurately and robustly using only $N \ll P$ random realizations of $u(\omega)$ via compressive sampling. Given N random samples of $u(\omega)$, compressive sampling aims at finding the sparsest (or nearly sparsest) coefficients \mathbf{c} from an optimization problem of the form

$$(P_{s,\delta}) : \min_{\mathbf{c}} \|\mathbf{W}\mathbf{c}\|_s \quad \text{subject to} \quad \|\Psi\mathbf{c} - \mathbf{u}\|_2 \leq \delta, \quad (1)$$

where $\|\mathbf{W}\mathbf{c}\|_s$, with $s = \{0, 1\}$ and some positive diagonal weight matrix \mathbf{W} , is a measure of the sparsity of \mathbf{c} and $\|\Psi\mathbf{c} - \mathbf{u}\|_2$ is a measure of the accuracy of the truncated PC expansion in estimating the $u(\omega)$ samples. The N -vector \mathbf{u} contains the independent random samples of $u(\omega)$ and the rows of the $N \times P$ matrix Ψ consist of the corresponding samples of the PC basis $\{\psi_\alpha\}$.

Throughout the rest of this manuscript, we will elaborate on the formulation of the compressive sampling problem (1) and the required conditions under which it leads to an accurate and stable approximation of an arbitrary sparse stochastic function as well as sparse solutions to linear elliptic stochastic PDEs. Although we choose to study this particular class of stochastic PDEs, we stress that the proposed algorithms and theoretical developments are far more general and may be readily applied to recover sparse solution of other stochastic systems.

In Section 2, we describe the setup of the problem of interest for which numerical experiments are performed. We then, in Section 3, lay out the main contributions of the present work. In particular, the approximation of sparse stochastic functions as well as the elliptic stochastic PDEs using the compressive sampling technique are introduced in Sections 3.2–3.4. Sections 3.5 and 3.6 discuss some of the implementation details of the present technique. To demonstrate the accuracy and efficiency of the proposed procedures, in Section 4, we perform two numerical experiments on 1- D (in space) linear elliptic stochastic differential equations with high-dimensional random diffusion coefficients.

2. Problem setup

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a complete probability space where \mathcal{P} is a probability measure on the σ -field \mathcal{F} . We consider the following elliptic stochastic PDE defined on a bounded Lipschitz continuous domain $\mathcal{D} \subset \mathbb{R}^D, D = 1, 2, 3$, with boundary $\partial\mathcal{D}$,

$$\begin{aligned} -\nabla \cdot (a(\mathbf{x}, \omega) \nabla u(\mathbf{x}, \omega)) &= f(\mathbf{x}) \quad \mathbf{x} \in \mathcal{D}, \\ u(\mathbf{x}, \omega) &= 0 \quad \mathbf{x} \in \partial\mathcal{D}, \end{aligned} \tag{2}$$

\mathcal{P} -a.s. $\omega \in \Omega$. The diffusion coefficient $a(\mathbf{x}, \omega)$ is a stochastic function defined on $(\Omega, \mathcal{F}, \mathcal{P})$ and is the source of uncertainty in (2). We assume that $a(\mathbf{x}, \omega)$ is specified by a truncated Karhunen–Loève–“like” expansion

$$a(\mathbf{x}, \omega) = \bar{a}(\mathbf{x}) + \sum_{i=1}^d \sqrt{\lambda_i} \phi_i(\mathbf{x}) y_i(\omega), \tag{3}$$

where $(\lambda_i, \phi_i), i = 1, \dots, d$, are the eigenpairs of the covariance function $C_{aa}(\mathbf{x}_1, \mathbf{x}_2) \in L_2(\mathcal{D} \times \mathcal{D})$ of $a(\mathbf{x}, \omega)$ and $\bar{a}(\mathbf{x})$ is the mean of $a(\mathbf{x}, \omega)$. We further assume that $a(\mathbf{x}, \omega)$ satisfies the following conditions:

A-I. For all $\mathbf{x} \in \mathcal{D}$, there exists constants a_{\min} and a_{\max} such that

$$0 < a_{\min} \leq a(\mathbf{x}, \omega) \leq a_{\max} < \infty \quad \mathcal{P} - a.s. \quad \omega \in \Omega. \tag{4}$$

A-II. The covariance function $C_{aa}(\mathbf{x}_1, \mathbf{x}_2)$ is *piecewise analytic* on $\mathcal{D} \times \mathcal{D}$ [57,8], implying that there exist real constants c_1 and c_2 such that for $i = 1, \dots, d$,

$$0 \leq \lambda_i \leq c_1 e^{-c_2 i^\kappa} \tag{5}$$

and

$$\forall \alpha \in \mathbb{N}^d : \sqrt{\lambda_i} \|\partial^\alpha \phi_i\|_{L^\infty(\mathcal{D})} \leq c_1 e^{-c_2 i^\kappa}, \tag{6}$$

where $\kappa := 1/D$ and $\alpha \in \mathbb{N}^d$ is a fixed multi-index. Notice that the decay rates in Eqs. (5) and (6) will be *algebraic* if $C_{aa}(\mathbf{x}_1, \mathbf{x}_2)$ has $C^s(\mathcal{D} \times \mathcal{D})$ regularity for some $s > 0$ [57].

A-III. The random variables $\{y_k(\omega)\}_{k=1}^d$ are independent and uniformly distributed on $\Gamma_k := [-1, 1], k = 1, \dots, d$, with probability density function $\rho_k(y_k) = 1/2$ defined over Γ_k . The joint probability density function of the random vector $\mathbf{y} := (y_1, \dots, y_d)$ is then given by $\rho(\mathbf{y}) := \prod_{k=1}^d \rho_k(y_k)$.

Remark. The algorithm proposed in the paper requires the existence of a sparse solution. The only role of assumption A-II is to guarantee the existence of such a sparse solution for problem (2). It is not necessary for the application and the validity of the proposed algorithm. In particular, if the coefficient $a(\mathbf{x}, \omega)$ is only essentially bounded, the proposed algorithm will be accurate as long as a sparse approximation exists (because the recovery does not depend on the regularity of a).

Given the finite-dimensional uncertainty representation in (3), the solution $u(\mathbf{x}, \omega)$ of (2) also admits a finite-dimensional representation, i.e.,

$$u(\mathbf{x}, \mathbf{y}) := u(\mathbf{x}, y_1(\omega), \dots, y_d(\omega)) : \mathcal{D} \times \Gamma \rightarrow \mathbb{R}, \tag{7}$$

where $\Gamma := \prod_{k=1}^d \Gamma_k$.

In what follows, we first briefly outline the Legendre spectral stochastic discretization of $u(\mathbf{x}, \mathbf{y})$ and subsequently introduce our approach based on compressive sampling to obtain such a discretization.

3. Numerical approach

3.1. Spectral stochastic discretization

In the context of the spectral stochastic methods [38,28,67,2,64], the solution $u(\mathbf{x}, \mathbf{y})$ of (2) is represented by an infinite series of the form

$$u(\mathbf{x}, \mathbf{y}) = \sum_{\alpha \in \mathbb{N}_0^d} c_\alpha(\mathbf{x}) \psi_\alpha(\mathbf{y}), \tag{8}$$

where $\mathbb{N}_0^d := \{(\alpha_1, \dots, \alpha_d) : \alpha_j \in \mathbb{N} \cup \{0\}\}$ is the set of multi-indices of size d defined on non-negative integers. The basis functions $\{\psi_\alpha(\mathbf{y})\}$ are multi-dimensional Legendre polynomials, referred to as the Legendre polynomial chaos, and are orthogonal with respect to the joint probability measure $\rho(\mathbf{y})$ of the random vector \mathbf{y} . Each basis function $\psi_\alpha(\mathbf{y})$ is a tensor product of univariate Legendre polynomials $\psi_{\alpha_i}(y_i)$ of degree $\alpha_i \in \mathbb{N}_0^1$, i.e.,

$$\psi_\alpha(\mathbf{y}) = \psi_{\alpha_1}(y_1) \psi_{\alpha_2}(y_2) \cdots \psi_{\alpha_d}(y_d), \quad \alpha \in \mathbb{N}_0^d. \tag{9}$$

We here assume that the univariate Legendre polynomials $\psi_{x_i}(y_i)$ are also normalized such that

$$\int_{\Gamma_i} \psi_{x_i}^2(y_i) \rho_i(y_i) dy_i = 1, \quad i = 1, \dots, d. \tag{10}$$

The exact generalized Fourier coefficients $c_\alpha(\mathbf{x})$ in (8), referred to as the PC coefficients, are computed by the projection of $u(\mathbf{x}, \mathbf{y})$ onto each basis function $\psi_\alpha(\mathbf{y})$,

$$c_\alpha(\mathbf{x}) = \mathbb{E}[u(\mathbf{x}, \mathbf{y})\psi_\alpha(\mathbf{y})] = \int_{\Gamma} u(\mathbf{x}, \mathbf{y})\psi_\alpha(\mathbf{y})\rho(\mathbf{y}) d\mathbf{y}. \tag{11}$$

Here, \mathbb{E} denotes the expectation operator. In practice, the expansion (8) is finite; that is, only a finite number of spectral modes is needed to approximate $u(\mathbf{x}, \mathbf{y})$ within a given target accuracy. Traditionally, a finite-order truncation of the basis $\{\psi_\alpha(\mathbf{y})\}$ is adopted for the approximation, i.e.,

$$u_p(\mathbf{x}, \mathbf{y}) := \sum_{\alpha \in A_{p,d}} c_\alpha(\mathbf{x})\psi_\alpha(\mathbf{y}), \tag{12}$$

where the set of multi-indices $A_{p,d}$ is

$$A_{p,d} := \{\alpha \in \mathbb{N}_0^d : \|\alpha\|_1 \leq p, \|\alpha\|_0 \leq d\} \tag{13}$$

and has the cardinality

$$P := |A_{p,d}| = \frac{(p+d)!}{p!d!}. \tag{14}$$

Here, $\|\alpha\|_1 = \sum_{i=1}^d \alpha_i$ and $\|\alpha\|_0 = \#\{i : \alpha_i > 0\}$ are the total order (degree) and dimensionality of the basis function $\psi_\alpha(\mathbf{y})$, respectively. The approximation is then refined by increasing p to achieve a given target accuracy. Under assumptions A-I, A-II, and A-III stated in Section 2, the solution $u(\mathbf{x}, \mathbf{y})$ is analytic with respect to the random variables $\{y_i\}_{i=1}^d$ (see [1]), and as p increases, the approximation (12) converges exponentially fast in the mean-squares sense [2,1,8].

Definition (Sparsity). The solution $u(\mathbf{x}, \mathbf{y})$ is said to be (nearly) sparse if only a small fraction of coefficients $c_\alpha(\mathbf{x})$ in (12) are dominant and contribute to the solution statistics.

As will be described next, a sparse solution $u(\mathbf{x}, \mathbf{y})$ may be accurately recovered using $N \ll P$ random realizations $\{u(\mathbf{x}, \mathbf{y}_i)\}_{i=1}^N$ using compressive sampling techniques. This has to be compared, for instance, with the least-squares regression-type techniques, [40], that normally require $N \gg P$ samples for an accurate recovery.

3.2. Sparse recovery using compressive sampling

Compressive sampling is an emerging theory in the field of signal and image processing [20,21,17,30,15,18,16,22,14]. It hinges around the idea that a set of incomplete random observations of a sparse signal can be used to accurately, or even exactly, recover the signal (provided that the basis in which the signal is sparse is known). In particular, the number of such observations may be much smaller than the cardinality of the signal. In the context of the Legendre PC expansion of a random function $u(\mathbf{y})$, compressive sampling may be interpreted as follows. If $u(\mathbf{y})$ is sparse with respect to the Legendre PC basis, that is, many of the PC coefficients c_α are negligible, then $u(\mathbf{y})$ can be accurately recovered using $N \ll P$ random realizations $\{u(\mathbf{y}_i)\}_{i=1}^N$, where P is the cardinality of the Legendre PC basis $\{\psi_\alpha\}$. We next elaborate on the above statement and address how and under what conditions such a sparse reconstruction is successfully achieved for a general stochastic function $u(\mathbf{y})$ as well as the solution to the problem (2).

Let $\{\mathbf{y}_i\}_{i=1}^N$ be i.i.d. samples of \mathbf{y} and $\{u(\mathbf{y}_i)\}_{i=1}^N$ the corresponding realizations of the stochastic function $u(\mathbf{y})$. For problem (2), $\{u(\mathbf{y}_i)\}_{i=1}^N$ are realizations of the solution $u(\mathbf{x}, \mathbf{y})$ at a fixed point \mathbf{x} in \mathcal{D} . For the time being, let us assume that the p th-order PC basis $\{\psi_\alpha\}$ is a complete basis to expand $u(\mathbf{y})$; we will relax this assumption as we proceed. Given pairs of $\{\mathbf{y}_i\}_{i=1}^N$ and $\{u(\mathbf{y}_i)\}_{i=1}^N$, we write

$$u(\mathbf{y}_i) = \sum_{\alpha \in A_{p,d}} c_\alpha \psi_\alpha(\mathbf{y}_i), \quad i = 1, \dots, N, \tag{15}$$

or equivalently,

$$\Psi \mathbf{c} = \mathbf{u}. \tag{16}$$

Here $\mathbf{c} \in \mathbb{R}^P$ is the vector of PC coefficients c_α to be determined, $\mathbf{u} \in \mathbb{R}^N$ is the vector of samples of $u(\mathbf{y})$, and each column of the measurement matrix $\Psi \in \mathbb{R}^{N \times P}$ contains samples of the j th element of the PC basis, i.e.,

$$\Psi[i,j] = \psi_{x_j}(\mathbf{y}_i), \quad i = 1, \dots, N, \quad j = 1, \dots, P. \tag{17}$$

We are interested in the case that the number N of solution samples is much smaller than the unknown PC coefficients P , i.e., $N \ll P$. Without any additional constraints on \mathbf{c} , the underdetermined linear system (16) is ill-posed and, in general, has infinitely many solutions. When \mathbf{c} is sparse; that is, only a small fraction of the coefficients c_α are significant, the problem (16) may be regularized to ensure a well-posed solution. Such a regularization may be achieved by seeking a solution \mathbf{c} with the minimum number of non-zeros. This can be formulated in the optimization problem

$$(P_0) : \min_{\mathbf{c}} \|\mathbf{c}\|_0 \quad \text{subject to} \quad \Psi \mathbf{c} = \mathbf{u}, \tag{18}$$

where the semi-norm $\|\mathbf{c}\|_0 := \#\{\alpha: c_\alpha \neq 0\}$ is the number of non-zero components of \mathbf{c} . In general, the global minimum solution of (P_0) is not unique and is NP-hard to compute: the cost of a global search is exponential in P . Further developments in compressive sampling resulted in a convex relaxation of problem (P_0) by minimization of the ℓ_1 -norm of the solution \mathbf{c} instead, i.e.,

$$(P_1) : \min_{\mathbf{c}} \|\mathbf{W}\mathbf{c}\|_1 \quad \text{subject to} \quad \Psi \mathbf{c} = \mathbf{u}, \tag{19}$$

where \mathbf{W} is a diagonal matrix whose $[j,j]$ entry is the ℓ_2 -norm of the j th column of Ψ and $\|\cdot\|_1$ denotes the ℓ_1 -norm. Notice that the ℓ_1 -norm is the closest convex function to the ℓ_0 -norm that compels the small coefficients c_α to be zero, thus promoting the sparsity in the solution. The purpose of weighting the ℓ_1 cost function with \mathbf{W} is to prevent the optimization from biasing toward the non-zero entries in \mathbf{c} whose corresponding columns in Ψ have large norms. The problem (P_1) is entitled Basis Pursuit (BP) [20] and its solution can be obtained by linear programming. Since the ℓ_1 -norm functional $\|\mathbf{c}\|_1$ is convex, the optimization problem (P_1) admits a unique solution that coincides with the unique solution to problem (P_0) for sufficiently sparse \mathbf{c} and with some constraints on the measurement matrix Ψ ; e.g., see [14].

In general, the p th-order PC basis is not complete for the exact representation of $u(\mathbf{y})$; therefore, we have to account for the truncation error. This can be accommodated in (P_0) and (P_1) by allowing a non-zero residual in the constraint $\Psi \mathbf{c} = \mathbf{u}$. Therefore, as in Sections 3.2.1 and 3.2.3 of [14], the proposed algorithms in this paper are error-tolerant versions of (P_0) and (P_1) , with error tolerance δ , i.e.,

$$(P_{0,\delta}) : \min_{\mathbf{c}} \|\mathbf{c}\|_0 \quad \text{subject to} \quad \|\Psi \mathbf{c} - \mathbf{u}\|_2 \leq \delta \tag{20}$$

and

$$(P_{1,\delta}) : \min_{\mathbf{c}} \|\mathbf{W}\mathbf{c}\|_1 \quad \text{subject to} \quad \|\Psi \mathbf{c} - \mathbf{u}\|_2 \leq \delta, \tag{21}$$

respectively. The latter problem is named Basis Pursuit Denoising (BPDN) in [20] and may be solved using techniques from quadratic programming. We leave the discussion on the available algorithms for solving problems $(P_{1,\delta})$ and $(P_{0,\delta})$ to Section 3.6. Instead, we henceforth delineate on sufficient conditions under which the BPDN problem $(P_{1,\delta})$ leads to a successful Legendre PC expansion of a general sparse stochastic function $u(\mathbf{y})$ and, subsequently, the sparse solution $u(\mathbf{x}, \mathbf{y})$ to the problem (2). Our results are extensions of those in [32, 14], adapted to the case where the measurement matrix Ψ consists of random evaluations of the Legendre PC basis $\{\psi_\alpha\}$. With slight differences that will be remarked accordingly, similar results hold for the case of the $(P_{0,\delta})$ problem.

Notation. We write $u_p^{1,\delta}(\mathbf{y}) := \sum_{\alpha \in A_{p,d}} c_\alpha^{1,\delta} \psi_\alpha(\mathbf{y})$ the Legendre PC approximation of $u(\mathbf{y})$ with coefficients $\mathbf{c}^{1,\delta}$ computed from the ℓ_1 -minimization problem $(P_{1,\delta})$ in (21). Additionally, we write $u_p^0(\mathbf{y})$ the Legendre PC expansion $u_p^0(\mathbf{y}) := \sum_{\alpha \in A_{p,d}^\epsilon} c_\alpha^0 \psi_\alpha(\mathbf{y})$, for some index set $A_{p,d}^\epsilon \subseteq A_{p,d}$ (or $A_{p_e, v_e} \subseteq A_{p,d}$) to be defined accordingly. Similar notations are used for PC expansion of the solution $u(\mathbf{x}, \mathbf{y})$ to problem (2).

Theorem 3.1 (General stability of $(P_{1,\delta})$). *Let $u(\mathbf{y})$ be an essentially bounded function of i.i.d. random variables $\mathbf{y} := (y_1, \dots, y_d)$ uniformly distributed on $\Gamma := [-1, 1]^d$. Define*

$$S_{\max} := \frac{N}{64P^{4c_{p,d}}(\ln P)} \tag{22}$$

with

$$c_{p,d} := \frac{\ln 3}{2} \frac{p}{\ln \left(\frac{(p+d)!}{p!d!} \right)}. \tag{23}$$

If there exists a Legendre PC expansion $u_p^0(\mathbf{y})$ for some index set $A_{p,d}^\epsilon \subseteq A_{p,d}$ such that $\|u - u_p^0\|_{L^\infty(\Gamma)} \leq \epsilon$ and

$$S < S_{\max}, \tag{24}$$

with $S := |A_{p,d}^\epsilon|$, then with probability

$$\text{Prob}_1 \geq 1 - 4P^{2-2S_{\max}} - P^{-8S_{\max}} - P^{-8S_{\max}P^{4c_{p,d}}}, \tag{25}$$

(on the N samples $\{u(\mathbf{y}_i)\}_{i=1}^N$) and for some constants c_1 and c_2 , the solution $u_p^{1,\delta}(\mathbf{y})$ must obey

$$\|u - u_p^{1,\delta}\|_{L^2(I)} \leq c_1 \epsilon + c_2 \frac{\delta}{\sqrt{N}}. \tag{26}$$

In simple words, **Theorem 3.1** states that if an essentially bounded stochastic function $u(\mathbf{y})$ admits a sparse Legendre PC expansion, then the ℓ_1 -minimization problem $(P_{1,\delta})$ can accurately recover it from a sufficiently large number of random solution realizations. The recovery is stable under the truncation error $\|\Psi \mathbf{c} - \mathbf{u}\|_2$ and is within a distance of the exact solution that is proportional to the error tolerance δ . It is worth highlighting that no prior knowledge of the sparsity pattern of the PC coefficients \mathbf{c} is needed for an accurate recovery.

Remark. Based on the conditions (22) and (24), the number N of random samples has to grow like $P^{4c_{p,d}} \ln P$ and also be proportional to the number of dominant coefficients $S = |A_{p,d}^\epsilon|$. Given any order p of the PC expansion, for sufficiently high-dimensional problems, the constant $c_{p,d} < 1/4$ (see **Lemma 3.3** and **Fig. 1**), thus justifying $N \ll P$. In fact, the conditions (22) and (24) are too pessimistic; in practice, the number of random samples required for an accurate recovery is much smaller than the theoretical value in (22). We will elaborate on this statement in Section 3.3.

Remark. Although the BPDN reconstruction is achieved by minimizing the ℓ_1 -norm of the solution, based on (26), the approximation also converges to the exact solution in the mean-squares sense.

Remark. A similar theorem holds for the case of the sparse approximation using the ℓ_0 -minimization problem $(P_{0,\delta})$ in (20). In this case, the condition (22) has to be replaced with $S_{\max} := \frac{N}{16p^{4c_{p,d}} (\ln P)}$ which is, in theory, milder than that of the $(P_{1,\delta})$ problem. The error estimate (26) also holds with a larger probability, but with different constants c_1 and c_2 .

Remark. In this paper (see also [35]) we have chosen the samples $\{\mathbf{y}_i\}_{i=1}^N$ from the uniform density $\rho(\mathbf{y}) = \prod_{k=1}^d \rho_k(y_k)$ where $\rho_k(y_k) = 1/2$ is defined over $[-1, 1]$. It is important to observe that this choice is arbitrary and not necessary for sparse recovery. In particular, in $d = 1$, selecting the samples $\{y_i\}_{i=1}^N$ from the Chebyshev probability measure $\rho(y) = \pi^{-1}(1 - y^2)^{-1/2}$ defined over $[-1, 1]$ allows for fewer samples for sparse recovery [56]. In higher dimensions (for the multivariate case of tensor product Legendre systems), the generalization of the estimates of [56] leads to constants depending on the number of tensorized polynomials. This could explain why mixed performance were observed between the two approaches in the high-dimensional cases (e.g., $d = 40$). The optimal selection of the samples $\{\mathbf{y}_i\}_{i=1}^N$ requires further study.

3.3. General stability of $(P_{1,\delta})$ and $(P_{0,\delta})$: proofs and more

The ability of problems $(P_{1,\delta})$ and $(P_{0,\delta})$ in accurately approximating the sparse PC coefficients \mathbf{c} in (12), hence $u(\mathbf{y})$, depends on two main factors: (i) the sparsity of the PC coefficients \mathbf{c} and (ii) the *mutual coherence* of the measurement matrix Ψ . In fact, the number N of random solution realizations required for a successful sparse approximation is dictated by these two factors. While sparsity is a characteristic of the function of interest $u(\mathbf{y})$ and the selected basis, the mutual coherence of the measurement matrix Ψ is universal as it only depends on the choice of the expansion basis, e.g., the Legendre PC basis $\{\psi_\alpha\}$, and the sampling process from which Ψ is assembled. While, in the present study, we are interested in approximating stochastic functions $u(\mathbf{y})$ that are sparse with respect the Legendre PC basis, we note that such an assumption may not be valid for an arbitrary function $u(\mathbf{y})$. In this case, our sparse approximation still converges to the actual solution but, perhaps, not using as few as $N \ll P$ random solution realizations. The problem (2) is a special instance of stochastic PDEs for which the sparsity of solution with respect to the Legendre PC basis is derived analytically [8]. We will briefly review the sparsity

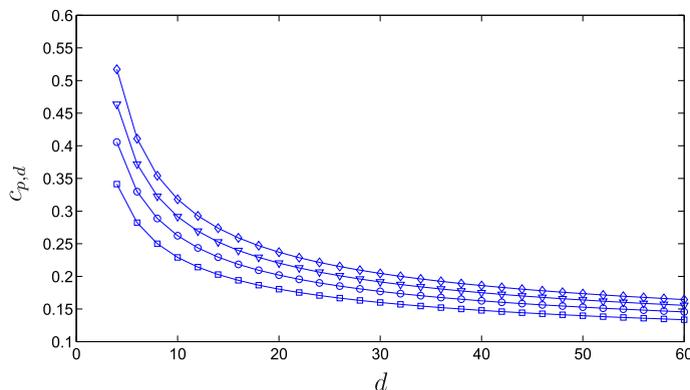


Fig. 1. Decay of $c_{p,d}$ as a function of d . $p = 1(\square)$; $p = 2(\circ)$; $p = 3(\nabla)$; $p = 4(\diamond)$.

analysis of [8] in Section 3.4. Subsequently, we give the definition of the mutual coherence of Ψ and discuss its role in our sparse approximation using $(P_{1,\delta})$ and $(P_{0,\delta})$.

3.3.1. Mutual coherence of Ψ

Definition (Mutual coherence [32]). The mutual coherence $\mu(\Psi)$ of a matrix $\Psi \in \mathbb{R}^{N \times P}$ is the maximum of absolute normalized inner-products of its columns. Let ψ_j and ψ_k be two distinct columns of Ψ . Then,

$$\mu(\Psi) := \max_{1 \leq j, k \leq P, j \neq k} \frac{|\psi_j^T \psi_k|}{\|\psi_j\|_2 \|\psi_k\|_2}. \tag{27}$$

In plain words, the mutual coherence is a measure of how close to orthogonal a matrix is. Clearly, for any general matrix Ψ ,

$$0 \leq \mu(\Psi) \leq 1, \tag{28}$$

where the lower bound is achieved, for instance, by unitary matrices. However, for the case of $N < P$, the mutual coherence $\mu(\Psi)$ is strictly positive. It is well understood that measurement matrices with smaller mutual coherence have a better ability to recover a sparse solution using compressive sampling techniques, e.g., see Lemma 3.4. Therefore, we shall proceed to examine the mutual coherence of the random measurement matrix Ψ in (16). We first observe that, by the orthogonality of the Legendre PC basis, i.e., $\mathbb{E}[\psi_{\alpha_j} \psi_{\alpha_k}] = \delta_{jk}$, and the strong law of large numbers, the mutual coherence $\mu(\Psi)$ converges to zero almost surely for asymptotically large random sample sizes N . However, it is essential for our purpose to (i) investigate if a desirably small $\mu(\Psi)$ can be achieved by a sample size $N \ll P$ and (ii) quantify how large $\mu(\Psi)$ can get for a finite N . These are addressed in the following theorem.

Theorem 3.2 (Bound on the mutual coherence $\mu(\Psi)$). Let $\Psi \in \mathbb{R}^{N \times P}$, as defined in (16), be the measurement matrix corresponding to N independent random realizations of the Legendre polynomial chaos basis of order p in d i.i.d. uniform random variables \mathbf{y} . There exists a positive constant $c_{p,d} := \frac{\ln 3}{2} \frac{p}{\ln \left(\frac{(p+d)!}{p!d!} \right)}$ depending on p and d , such that if

$$0 \leq r = 2\sqrt{\zeta P^{4c_{p,d}} (\ln P) / N} \leq 1/2 \tag{29}$$

for some $\zeta > 1$, then

$$\text{Prob} \left[\mu(\Psi) \geq \frac{r}{1-r} \right] \leq 4P^{2-2\zeta}. \tag{30}$$

Fig. 1 illustrates the decay of $c_{p,d}$, for several values of p , as a function of d . Based on Theorem 3.2, for cases where the number d of random variables \mathbf{y} is large enough such that $c_{p,d} < 1/4$, it is sufficient to have $N \sim \mathcal{O}(16P^{4c_{p,d}} \ln P) \ll P$ to keep $\mu(\Psi)$ bounded from above with a large probability. Notice that such a requirement on $c_{p,d}$ is particularly suited to high-dimensional problems.

Remark. We observe that, given the choice of r in (29), the upper bound on $\mu(\Psi)$ in (30) decays like $1/\sqrt{N}$ for asymptotically large N , which is consistent with the Central Limit Theorem.

In order to prove Theorem 3.2, we first need to compute the maximum of the Legendre PC basis functions $\psi_{\alpha}(\mathbf{y})$. This is given in the following lemma.

Lemma 3.3 (Maximum of $\psi_{\alpha}(\mathbf{y})$). Let $\{\psi_{\alpha}(\mathbf{y})\}$ be the Legendre polynomial chaos basis of total order p in d i.i.d uniform random variables \mathbf{y} (as defined in (13)) and with cardinality P . Then,

$$\|\psi_{\alpha}\|_{L^{\infty}(I)} \leq P^{c_{p,d}} \tag{31}$$

with a constant

$$c_{p,d} := \frac{\ln 3}{2} \frac{p}{\ln \left(\frac{(p+d)!}{p!d!} \right)}. \tag{32}$$

Proof. Given the equality

$$\|\psi_{\alpha_i}\|_{L^{\infty}(I_i)} = \sqrt{2\alpha_i + 1},$$

we have

$$\|\psi_{\alpha}\|_{L^{\infty}(I)} = \prod_{i=1}^d \sqrt{2\alpha_i + 1}.$$

Using the constraint $\alpha_i \in \mathbb{N}_0^1$ and $\sum_{i=1}^d \alpha_i \leq p$, the right-hand-side is maximized when p of α_i s are equal to one and $d - p$ equal to zero (we assume $d \geq p$). We deduce that

$$\|\psi_{\alpha}\|_{L^{\infty}(\Gamma)} \leq 3^{\frac{p}{2}} = P^{\frac{\ln 3}{2} \frac{p}{\ln P}} = P^{c_{p,d}}.$$

We are now ready to prove **Theorem 3.2**. \square

Proof of Theorem 3.2. The mutual coherence $\mu(\Psi)$ is

$$\mu(\Psi) = \max_{1 \leq j, k \leq P, j \neq k} \frac{\left| \frac{1}{N} \sum_{i=1}^N \psi_{\alpha_j}(\mathbf{y}_i) \psi_{\alpha_k}(\mathbf{y}_i) \right|}{\left(\frac{1}{N} \sum_{i=1}^N \psi_{\alpha_j}^2(\mathbf{y}_i) \right)^{1/2} \left(\frac{1}{N} \sum_{i=1}^N \psi_{\alpha_k}^2(\mathbf{y}_i) \right)^{1/2}}. \tag{33}$$

Given the independence of samples $\{\mathbf{y}_i\}_{i=1}^N$ and using the McDiarmid’s inequality (Lemma 1.2 in [46]), we obtain

$$\text{Prob} \left[\left| \frac{1}{N} \sum_{i=1}^N \psi_{\alpha_j}(\mathbf{y}_i) \psi_{\alpha_k}(\mathbf{y}_i) \right| \geq r \right] \leq 2 \exp \left(\frac{-2Nr^2}{(2\|\psi_{\alpha_j}(\mathbf{y}_i) \psi_{\alpha_k}(\mathbf{y}_i)\|_{L^{\infty}(\Gamma_i)})^2} \right). \tag{34}$$

Using **Lemma 3.3**, we may probabilistically bound the numerator in (33) as

$$\text{Prob} \left[\left| \frac{1}{N} \sum_{i=1}^N \psi_{\alpha_j}(\mathbf{y}_i) \psi_{\alpha_k}(\mathbf{y}_i) \right| \geq r \right] \leq 2 \exp \left(\frac{-Nr^2}{2P^{4c_{p,d}}} \right),$$

in which we exploit the orthonormality of $\psi_{\alpha_j}(\mathbf{y})$ and $\psi_{\alpha_k}(\mathbf{y})$, i.e., $\mathbb{E}[\psi_{\alpha_j} \psi_{\alpha_k}] = \delta_{jk}$. Similarly, for $j = 1, \dots, P$, we have

$$\text{Prob} \left[\frac{1}{N} \sum_{i=1}^N \psi_{\alpha_j}^2(\mathbf{y}_i) \leq 1 - r \right] \leq \exp \left(\frac{-Nr^2}{P^{4c_{p,d}}} \right) \leq \exp \left(\frac{-Nr^2}{2P^{4c_{p,d}}} \right).$$

Therefore,

$$\text{Prob} \left[\frac{\left| \frac{1}{N} \sum_{i=1}^N \psi_{\alpha_j}(\mathbf{y}_i) \psi_{\alpha_k}(\mathbf{y}_i) \right|}{\left(\frac{1}{N} \sum_{i=1}^N \psi_{\alpha_j}^2(\mathbf{y}_i) \right)^{1/2} \left(\frac{1}{N} \sum_{i=1}^N \psi_{\alpha_k}^2(\mathbf{y}_i) \right)^{1/2}} \geq \frac{r}{1-r} \right] \leq 4 \exp \left(\frac{-Nr^2}{2P^{4c_{p,d}}} \right) \tag{35}$$

and

$$\text{Prob} \left[\mu(\Psi) \geq \frac{r}{1-r} \right] \leq 4P^2 \exp \left(\frac{-Nr^2}{2P^{4c_{p,d}}} \right). \tag{36}$$

Taking

$$r = 2\sqrt{\zeta P^{4c_{p,d}} (\ln P) / N}, \tag{37}$$

for some $\zeta > 1$, we arrive at the statement of the **Theorem 3.2**. \square

To summarize, we observe that with large probability, the mutual coherence $\mu(\Psi)$ of the measurement matrix Ψ in (16) can be arbitrarily bounded from above by increasing the number N of independent random solution realizations.

Following [32,14], we next state a *sufficient* condition on the sparsity of $u(\mathbf{y})$ (or, equivalently, the mutual coherence of Ψ) such that the problem $(P_{1,\delta})$ leads to a stable and accurate sparse approximation. By stability, we simply mean that the PC coefficients $\mathbf{c}^{1,\delta}$ recovered from the ℓ_1 -minimization problem $(P_{1,\delta})$ do not blow up in the presence of the truncation error δ .

Lemma 3.4 (A condition on sparsity for stability of $(P_{1,\delta})$). *Let $u_p^0(\mathbf{y})$ be a sparse Legendre PC approximation of an essentially bounded function $u(\mathbf{y})$ associated with an index set $\Lambda_{p,d}^{\epsilon} \subseteq \Lambda_{p,d}$. Assume that the vector of PC coefficients \mathbf{c}^0 satisfies the sparsity condition*

$$\|\mathbf{c}^0\|_0 = |\Lambda_{p,d}^{\epsilon}| < (1 + 1/\mu(\Psi))/4. \tag{38}$$

Then, with probability at least $1 - \exp(-\frac{N}{8P^{4c_{p,d}}})$ (on the N samples $\{\mathbf{y}_i\}_{i=1}^N$) and for all $\delta \geq 0$, the solution $u_p^{1,\delta}$ must obey

$$\mathbb{E} \left[\left(u_p^0 - u_p^{1,\delta} \right)^2 \right] \leq \frac{4}{3N} \frac{(\delta + \|\Psi \mathbf{c}^0 - \mathbf{u}\|_2)^2}{1 - \mu(\Psi)(4\|\mathbf{c}^0\|_0 - 1)}. \tag{39}$$

Proof. Using **Theorem 3.1** of [32], we obtain that if \mathbf{c}^0 satisfies the sparsity condition $\|\mathbf{c}^0\|_0 < (1 + 1/\mu(\Psi))/4$, then

$$\sum_{\alpha} (c_{\alpha}^{1,\delta} - c_{\alpha}^0)^2 \|\psi_{\alpha}\|_2^2 \leq \frac{(\delta + \|\Psi \mathbf{c}^0 - \mathbf{u}\|_2)^2}{1 - \mu(\Psi)(4\|\mathbf{c}^0\|_0 - 1)},$$

where $\|\psi_{\alpha}\|_2$ is the ℓ_2 -norm of the column of Ψ corresponding to the index α . The presence of $\|\psi_{\alpha}\|_2$ is due to the fact that the columns of Ψ are not normalized. Next, using McDiarmid’s inequality and the independence of the entries $\Psi[i, j]$ for distinct rows i , we obtain that

$$\text{Prob} \left[\sum_{\alpha} (c_{\alpha}^{1,\delta} - c_{\alpha}^0)^2 \|\psi_{\alpha}\|_2^2 \leq \frac{3N}{4} \sum_{\alpha} (c_{\alpha}^{1,\delta} - c_{\alpha}^0)^2 \right] \leq \exp \left(-\frac{N}{8\|\psi_{\alpha}\|_{L^{\infty}(T)}^4} \right). \tag{40}$$

We conclude using Lemma 3.3 and the fact that, due to the orthonormality of $\{\psi_{\alpha}(\mathbf{y})\}$, we have $\mathbb{E} \left[(u_p^0 - u_p^{1,\delta})^2 \right] = \|\mathbf{c}^0 - \mathbf{c}^{1,\delta}\|_2^2$. \square

Remark. The error bound in (39) is not tight; in fact, the actual error is significantly smaller than the upper bound given in (39). More importantly, according to [32], the sparsity condition (38) is unnecessarily too restrictive. In practice, both far milder sparsity conditions are needed and much smaller errors are achieved.

Remark. We will later use the sparsity condition (38) to derive the sufficient condition (22) (together with (24)) on the number N of random samples needed for a successful recovery. As the condition (38) is too restrictive, the theoretical lower bound on N given in (22) and (24) is too pessimistic.

Remark. According to Lemma 3.4, we do not need to know *a priori* the sparse index set $A_{p,d}^{\epsilon}$; only the sparsity condition (38) is required.

Remark. With slight modifications, a similar argument as in Lemma 3.4 may be asserted for the solution of ℓ_0 -minimization problem $(P_{0,\delta})$. Specifically, in that case, we only require a sparsity limit $\|\mathbf{c}^0\|_0 = |A_{p,d}^{\epsilon}| < (1 + 1/\mu(\Psi))/2$ to achieve the error estimate $\mathbb{E} \left[(u_p^0 - u_p^{0,\delta})^2 \right] \leq \frac{4}{3N} \frac{(\delta + \|\Psi \mathbf{c}^0 - \mathbf{u}\|_2)^2}{1 - \mu(\Psi)(2\|\mathbf{c}^0\|_0 - 1)}$.

Notice that the normalized truncation error

$$\epsilon_N^2 := \frac{\|\Psi \mathbf{c}^0 - \mathbf{u}\|_2^2}{N} \tag{41}$$

is the sample average estimate of the mean-squares sparse approximation error $\mathbb{E} \left[(u - u_p^0)^2 \right]$ and is a function of samples $\{\mathbf{y}_i\}_{i=1}^N$ in addition to the order p and the dimensionality d of the sparse PC expansion. For a given set of N independent random samples $\{\mathbf{y}_i\}_{i=1}^N$ and realizations $\{u(\mathbf{y}_i)\}_{i=1}^N$, we may probabilistically bound ϵ_N^2 using McDiarmid’s inequality, i.e.,

$$\text{Prob} \left[\epsilon_N^2 \geq \mathbb{E} \left[(u - u_p^0)^2 \right] + r \right] \leq \exp \left(-2N \frac{r^2}{\|u - u_p^0\|_{L^{\infty}(T)}^4} \right). \tag{42}$$

Proof of Theorem 3.1. Given the statement of the Lemma 3.4, we have

$$\|u - u_p^{1,\delta}\|_{L^2(T)} \leq \|u - u_p^0\|_{L^2(T)} + \|u_p^0 - u_p^{1,\delta}\|_{L^2(T)} \leq \epsilon + \frac{2}{\sqrt{3N}} \frac{\delta + \|\Psi \mathbf{c}^0 - \mathbf{u}\|_2}{\sqrt{1 - \mu(\Psi)(4S - 1)}}$$

for all $\delta > 0$. Furthermore, selecting $r = \frac{1}{4} \|u - u_p^0\|_{L^{\infty}(T)}^2$ in (42) leads to

$$\epsilon_N^2 =: \frac{\|\Psi \mathbf{c}^0 - \mathbf{u}\|_2^2}{N} \leq \|u - u_p^0\|_{L^2(T)}^2 + \frac{1}{4} \|u - u_p^0\|_{L^{\infty}(T)}^2 \leq \frac{5}{4} \epsilon^2$$

with probability at least $1 - P^{-8S_{\max} P^{4\epsilon p, d}}$. Finally, by taking

$$c_1 := 1 + \frac{\sqrt{5}}{\sqrt{3}\sqrt{1 - \mu(\Psi)(4S - 1)}} \quad \text{and} \quad c_2 := \frac{2}{\sqrt{3}\sqrt{1 - \mu(\Psi)(4S - 1)}},$$

we arrive at the statement of the Theorem 3.1. \square

The stability analyses presented above were tied to the case of a generic stochastic function $u(\mathbf{y})$ which was assumed to be sparse with respect to the Legendre PC basis. We will next extend our analysis to the specific case of problem (2) whose solution $u(\mathbf{x}, \mathbf{y})$ is provably sparse.

3.4. Stability of $(P_{1,\delta})$ for stochastic PDE (2)

We here briefly summarize the results of [60,8] on the sparsity of the Legendre PC expansion of the solution $u(\mathbf{x}, \mathbf{y})$ to the problem (2). Alternative to the p th-order truncated PC expansion of (12), one may ideally seek a proper index set $A_{p,d}^\epsilon \subseteq A_{p,d}$, with sufficiently large p , such that for a given accuracy ϵ

$$A_{p,d}^\epsilon := \arg \min \left\{ |\tilde{A}_{p,d}| : \tilde{A}_{p,d} \subseteq A_{p,d}, \|u - \tilde{u}_p\|_{H_0^1(\mathcal{D}, L^\infty(\Gamma))} \leq \epsilon \right\} \tag{43}$$

in which $\tilde{u}_p(\mathbf{x}, \mathbf{y}) := \sum_{\alpha \in \tilde{A}_{p,d}} \tilde{c}_\alpha(\mathbf{x}) \psi_\alpha(\mathbf{y})$. This will lead to the so-called *sparse approximation* of $u(\mathbf{x}, \mathbf{y})$ if

$$|A_{p,d}^\epsilon| \ll |A_{p,d}| = P, \tag{44}$$

where $A_{p,d}$ is defined in (13). Such a reduction in the number of basis functions in (44) is possible as, given the accuracy ϵ , the *effective dimensionality* v of $u(\mathbf{x}, \mathbf{y})$ in Γ is potentially smaller than the apparent dimensionality d . More precisely, under assumptions A-I, A-II, and A-III stated in Section 2, the analyses of [60,8] imply that the discretization of $u(\mathbf{x}, \mathbf{y})$ using a sparse index set $A_{p,v}$,

$$A_{p,v} := \{ \alpha \in \mathbb{N}_0^d : \|\alpha\|_1 \leq p, \|\alpha\|_0 \leq v \leq d \}, \tag{45}$$

preserves the exponential decay of the approximation error in the $H_0^1(\mathcal{D}, L^\infty(\Gamma))$ sense. For the sake of completeness, we cite this from [8] in the following lemma.

Lemma 3.5 (Proposition 3.10 of [8]). *Given assumptions A-I, A-II, and A-III in Section 2, there exist constants $c_1, c_2, c_3, c_4 > 0$, depending only on $a(\mathbf{x}, \omega)$ and $f(\mathbf{x})$ but independent of d, p, v , such that*

$$\|u - u_{p,v}\|_{H_0^1(\mathcal{D}, L^\infty(\Gamma))} \leq c_1 \left(e^{-c_2 v^{1+\kappa}} + e^{c_3 v(\ln d + \ln p) - c_4 p} \right), \tag{46}$$

for any $d, p, v \in \mathbb{N}$ with $v \leq d$ and $\kappa = 1/D$.

In particular, for $d \geq c_d |\ln \epsilon|^{1/\kappa}$, choosing

$$p_\epsilon = \lceil c_p d^{\kappa} \rceil \leq p \quad \text{and} \quad v_\epsilon = \lceil c_v d^{\kappa/(\kappa+1)} \rceil \leq d, \tag{47}$$

leads to

$$\|u - u_{p_\epsilon, v_\epsilon}\|_{H_0^1(\mathcal{D}, L^\infty(\Gamma))} \leq \epsilon, \tag{48}$$

where $u_{p_\epsilon, v_\epsilon}$ is now defined on a sparse index set

$$A_{p_\epsilon, v_\epsilon} := \{ \alpha \in \mathbb{N}_0^d : \|\alpha\|_1 \leq p_\epsilon, \|\alpha\|_0 \leq v_\epsilon \} \tag{49}$$

with cardinality

$$|A_{p_\epsilon, v_\epsilon}| \lesssim \epsilon^{-1/\gamma}, \tag{50}$$

for some arbitrary large $\gamma > 0$ and constants c_d, c_p , and c_v independent of d, p_ϵ , and v_ϵ [8].

In practice, the sparse set $A_{p,d}^\epsilon$ in (43) (or equivalently $A_{p_\epsilon, v_\epsilon}$ in (49)) is not known *a priori*. In [8], an approach based on an algebraic purely-stochastic problem is proposed to adaptively identify $A_{p_\epsilon, v_\epsilon}$. Having done this, the coefficients of the spectral modes are computed via the (intrusive) stochastic Galerkin scheme [38,67]. Alternatively, in this work, we apply our sparse approximation using $(P_{1,\delta})$ and $(P_{0,\delta})$ to compute $u(\mathbf{x}, \mathbf{y})$. The implementation of $(P_{1,\delta})$ and $(P_{0,\delta})$ is non-intrusive; only random realizations of the solution are needed. Moreover, we do not adapt the sampling process to identify the important dimensions at the stochastic level; therefore, our constructions are non-adapted.

We next state our main result on the convergence and stability of problem $(P_{1,\delta})$ (or $(P_{0,\delta})$) for the approximation of the sparse solution to the stochastic PDE (2). Our result is based upon the combination of Lemma 3.5 with Theorem 3.1.

Theorem 3.6 (Stability of $(P_{1,\delta})$ for stochastic PDE (2)). *Let $u_p^{1,\delta}(\mathbf{x}, \mathbf{y})$ be the p th-order Legendre PC approximation of $u(\mathbf{x}, \mathbf{y})$ in (2) where the coefficients $c_\alpha^{1,\delta}(\mathbf{x})$ are obtained from the ℓ_1 -minimization problem $(P_{1,\delta})$ with N independent realizations of $u(\mathbf{x}, \mathbf{y})$ and arbitrary $\delta \geq 0$. Write $\kappa = 1/D$. Let $c_{p,d}$ be defined by (23) and let S_{max} be defined by (22). Let $\gamma > 0$ be arbitrary.*

There exists constants c_1, c_2, c_3, c_4, c_5 independent from p, d, κ, N such that if $\lceil c_2 d^\kappa \rceil \leq p$ and $\lceil c_3 d^{\kappa/(\kappa+1)} \rceil \leq d$, then with probability at least

$$\text{Prob}_1 \geq 1 - 4p^{2-2S_{max}} - p^{-8S_{max}} - p^{-8p^{\kappa} p_d S_{max}}, \tag{51}$$

the solution $u_p^{1,\delta}$ must obey

$$\|u - u_p^{1,\delta}\|_{L^2(\mathcal{D}, L^2(\Gamma))} \leq c_4 \epsilon + c_5 \frac{\delta}{\sqrt{N}} \tag{52}$$

with

$$\epsilon := \max \left(\frac{1}{S_{\max}^\gamma}, \exp \left(- \left(\frac{d}{c_1} \right)^\kappa \right) \right). \tag{53}$$

Proof. We first note that, given the conditions of Lemma 3.5, the solution to problem (2) admits a sparse Legendre PC expansion u_p with sparsity $S = |A_{p_e, v_e}| \lesssim \epsilon^{-1/\gamma}$ when $\epsilon \geq \exp \left(- \left(\frac{d}{c_1} \right)^\kappa \right)$ for some constants c_1 and (arbitrary) $\gamma > 0$. Notice that the sparse approximation u_p has an accuracy better than ϵ in the $H_0^1(\mathcal{D}, L^\infty(\Gamma))$ sense. As a consequence of Lemma 3.4,

$$\|u_p^0 - u_p^{1,\delta}\|_{L^2(\mathcal{D}, L^2(\Gamma))}^2 \leq \frac{4}{3N} \frac{\left(\delta + \|\Psi \mathbf{c}^0 - \mathbf{u}\|_{L^2(\mathcal{D}, \ell_2(\mathbb{R}^N))} \right)^2}{1 - \mu(\Psi)(4\|\mathbf{c}^0\|_0 - 1)} \tag{54}$$

with probability at least $1 - \exp\left(-\frac{N}{8P^{4c_{p,d}}}\right)$ (on the N realizations $\{u(\mathbf{x}, \mathbf{y}_i)\}_{i=1}^N$). Additionally, the normalized error $\epsilon_N^2 := \frac{\|\Psi \mathbf{c}^0 - \mathbf{u}\|_{L^2(\mathcal{D}, \ell_2(\mathbb{R}^N))}^2}{N}$ is bounded from above in probability through

$$\text{Prob} \left[\epsilon_N^2 \geq \|u - u_p^0\|_{L^2(\mathcal{D}, L^2(\Gamma))}^2 + r \right] \leq \exp \left(-2N \frac{r^2}{\|u - u_p^0\|_{L^2(\mathcal{D}, L^\infty(\Gamma))}^4} \right). \tag{55}$$

Following Theorem 3.2, it is sufficient to have random solution realizations of size $N \geq 64P^{4c_{p,d}}(\ln P)S$, to meet the sparsity requirement $S = |A_{p_e, v_e}| < (1 + 1/\mu(\Psi))/4$ with probability at least $1 - 4P^{2-2S_{\max}}$ where $S_{\max} := \frac{N}{64P^{4c_{p,d}}(\ln P)}$. On the other hand, given N random realizations of solution, we require $\epsilon \geq \frac{1}{S_{\max}^\gamma}$ to satisfy the sparsity condition. Using the triangular and Poincaré inequalities, with probability at least $1 - P^{-8S_{\max}}$, we have

$$\|u - u_p^{1,\delta}\|_{L^2(\mathcal{D}, L^2(\Gamma))} \leq c_D \|u - u_p^0\|_{H_0^1(\mathcal{D}, L^2(\Gamma))} + \|u_p^0 - u_p^{1,\delta}\|_{L^2(\mathcal{D}, L^2(\Gamma))} \leq c_D \epsilon + \frac{2}{\sqrt{3N}} \frac{\delta + \|\Psi \mathbf{c}^0 - \mathbf{u}\|_{L^2(\mathcal{D}, \ell_2(\mathbb{R}^N))}}{\sqrt{1 - \mu(\Psi)(4S - 1)}}$$

for all $\delta > 0$. Moreover, by choosing $r = \frac{1}{4} \|u - u_p^0\|_{L^2(\mathcal{D}, L^\infty(\Gamma))}^2$ in (55),

$$\epsilon_N^2 := \frac{\|\Psi \mathbf{c}^0 - \mathbf{u}\|_{L^2(\mathcal{D}, \ell_2(\mathbb{R}^N))}^2}{N} \leq \|u - u_p^0\|_{L^2(\mathcal{D}, L^2(\Gamma))}^2 + \frac{1}{4} \|u - u_p^0\|_{L^2(\mathcal{D}, L^\infty(\Gamma))}^2 \leq \frac{5}{4} c_D^2 \epsilon^2$$

with probability at least $1 - P^{-8S_{\max}P^{4c_{p,d}}}$. Finally, by taking

$$c_4 := c_D \left(1 + \frac{\sqrt{5}}{\sqrt{3}\sqrt{1 - \mu(\Psi)(4S - 1)}} \right) \quad \text{and} \quad c_5 := \frac{2}{\sqrt{3}\sqrt{1 - \mu(\Psi)(4S - 1)}},$$

we arrive at the statement of Theorem 3.6. \square

Remark. We note that the sparsity of the solution $u(\mathbf{x}, \mathbf{y})$ to problem (2) is due to the fact that the effective dimensionality v_e is potentially smaller than d . However, as given in Theorem 3.1, as far as the sparsity condition (38) is satisfied, similar stability results are valid for situations where dominant basis are defined over all the dimensions.

3.5. Choosing truncation error tolerance δ

An important component of the sparse approximation using $(P_{1,\delta})$ and $(P_{0,\delta})$ is the selection of the truncation error tolerance δ . Although the stability bounds given in Lemma 3.4 are valid for any $\delta \geq 0$, the actual error and the sparsity level of the solution to $(P_{1,\delta})$ and $(P_{0,\delta})$ depend on the choice of δ . Ideally, we desire to choose $\delta \approx \|\Psi \mathbf{c}^0 - \mathbf{u}\|_2$; while larger values of δ deteriorate the accuracy of the approximation, as in Lemma 3.4, smaller choices of δ may result in over-fitting the solution samples and, thus, less sparse solutions. In practice, as the exact values of the PC coefficients \mathbf{c}^0 are not known, the exact values of the truncation error $\|\Psi \mathbf{c}^0 - \mathbf{u}\|$ and, consequently, δ are not known *a priori*. Therefore, δ has to be estimated, for instance, using statistical techniques such as the cross-validation [12,65].

In this work, we propose a heuristic cross-validation algorithm to estimate δ . We first divide the N available solution samples to N_r reconstruction and N_v validation samples such that $N = N_r + N_v$. The idea is to repeat the solution of $(P_{1,\delta})$ (or $(P_{0,\delta})$) on the reconstruction samples and with multiple values of truncation error tolerance δ_r . We then set $\delta = \sqrt{\frac{N}{N_r}} \hat{\delta}_r$ in which $\hat{\delta}_r$ is such that the corresponding truncation error on the N_v validation samples is minimum. This is simply motivated by the fact that the truncation error on the validation samples is large for values of δ_r considerably larger and smaller than $\|\Psi \mathbf{c}^0 - \mathbf{u}\|_2$ evaluated using the reconstruction samples. While the former is expected from the upper bound on the approximation error in Lemma 3.4, the latter is due to the overfitting the reconstruction samples. The following exhibit outlines the estimation of δ using the above cross-validation approach:

Algorithm for cross-validation estimation of δ :

- Divide the N solution samples to N_r reconstruction and N_v validation samples.
- Choose multiple values for δ_r such that the exact truncation error $\|\Psi \mathbf{c}^0 - \mathbf{u}\|_2$ of the reconstruction samples is within the range of δ_r values.
- For each value of δ_r , solve $(P_{1,\delta})$ (or $(P_{0,\delta})$) using the N_r reconstruction samples.
- For each value of δ_r , compute the truncation error $\delta_v := \|\Psi \mathbf{c}^{1,\delta_r} - \mathbf{u}\|_2$ (or $\delta_v := \|\Psi \mathbf{c}^{0,\delta_r} - \mathbf{u}\|_2$) of the N_v validation samples.
- Find the minimum value of δ_v and its corresponding $\hat{\delta}_r := \delta_r$.
- Set $\delta = \sqrt{\frac{N}{N_r}} \hat{\delta}_r$.

In the numerical experiments of Section 4, we repeat the above cross-validation algorithm for multiple replications of the reconstruction and validation samples. The estimate of $\delta = \sqrt{\frac{N}{N_r}} \hat{\delta}_r$ is then based on the value of $\hat{\delta}_r$ for which the average of the corresponding truncation errors δ_v , over all replications of the validation samples, is minimum. This resulted in more accurate solutions in our numerical experiments.

3.6. Algorithms

There are several numerical algorithms for solving problems $(P_{0,\delta})$ and $(P_{1,\delta})$ each with different optimization kernel, computational complexity, and degree of accuracy. An in-depth discussion on the performance of these algorithms is outside the scope of the present work; however, below we name some of the available options for each problem and briefly describe the algorithms that have been utilized in our numerical experiments. For comprehensive discussions on this subject, the interested reader is referred to [14,62,37,4,69,61].

Problem $(P_{0,\delta})$: A brute force search through all possible support sets in order to identify the correct sparsity for the solution $\mathbf{c}^{0,\delta}$ of $(P_{0,\delta})$ is NP-hard and not practical. Greedy pursuit algorithms form a major class of schemes to tackle the solution of $(P_{0,\delta})$ with a tractable computational cost. Instead of performing an exhaustive search for the support of the sparse solution, these solvers successively find one or more components of the solution that result in the largest improvement in the approximation. Some of the standard greedy pursuit algorithms are Orthogonal Marching Pursuit (OMP) [55,27], Regularized OMP (ROMP)[49], Stagewise OMP (StOMP) [31], Compressive Sampling MP (CoSaMP) [48], Subspace Pursuit [24], and Iterative Hard Thresholding (IHT) [26]. Under well-defined conditions, all of the above schemes provide stable and accurate solutions to $(P_{0,\delta})$ in a reasonable time.

In the present study, we employ the OMP algorithm to approximate the solution of $(P_{0,\delta})$. Starting from $\mathbf{c}^{0,\delta,(0)} = \mathbf{0}$ and an empty active column set of Ψ , at any iteration k , OMP identifies only one column to be added to the active column set. The column is chosen such that the ℓ_2 -norm of the residual, $\|\Psi \mathbf{c}^{0,\delta,(k)} - \mathbf{u}\|_2$, is maximally reduced. Having specified the active column set, a least-squares problem is solved to compute the solution $\mathbf{c}^{0,\delta,(k)}$. The iterations are continued until the error truncation tolerance δ is achieved. In general, the complexity of the OMP algorithm is $\mathcal{O}(S \cdot N \cdot P)$ where $S := \|\mathbf{c}^{0,\delta}\|_0$ is number of non-zero (dominant) entries of $\mathbf{c}^{0,\delta}$. The following exhibit depicts a step-by-step implementation of the OMP algorithm.

Orthogonal Matching Pursuit (OMP) Algorithm:

- Set $k = 0$.
 - Set the initial solution $\mathbf{c}^{0,\delta,(0)} = \mathbf{0}$ and residual $\mathbf{r}^{(0)} = \mathbf{u} - \Psi \mathbf{c}^{0,\delta,(0)} = \mathbf{u}$.
 - Set the solution support index set $\mathcal{I}^{(0)} = \emptyset$.
- While $\|\mathbf{u} - \Psi \mathbf{c}^{0,\delta,(k)}\|_2 > \delta$ perform:
 - For all $j \notin \mathcal{I}^{(k)}$ evaluate $\epsilon(j) = \|\psi_j \alpha_j - \mathbf{r}^{(k)}\|_2$ with $\alpha_j = \psi_j^T \mathbf{r}^{(k)} / \|\psi_j\|_2^2$.
 - Set $k = k + 1$.
 - Update the support index set $\mathcal{I}^{(k)} = \mathcal{I}^{(k-1)} \cup \{\arg \min_j \epsilon(j)\}$.
 - Solve for $\mathbf{c}^{0,\delta,(k)} = \arg \min_{\mathbf{c}^{0,\delta}} \|\mathbf{u} - \Psi \mathbf{c}^{0,\delta}\|_2$ subject to $\text{Support}\{\mathbf{c}^{0,\delta}\} = \mathcal{I}^{(k)}$.
 - Update the residual $\mathbf{r}^{(k)} = \mathbf{u} - \Psi \mathbf{c}^{0,\delta,(k)}$.
- Output the solution $\mathbf{c}^{0,\delta} = \mathbf{c}^{0,\delta,(k)}$.

Although we chose OMP in our analysis, we note that further studies are needed to identify the most appropriate greedy algorithm for the purpose of this study.

Problem ($P_{1,\delta}$): The majority of available solvers for ℓ_1 -minimization are based on alternative formulations of ($P_{1,\delta}$), such as the ℓ_1 -norm regularized least-squares problem

$$(QP_\lambda) : \min_c \frac{1}{2} \|\Psi \mathbf{c} - \mathbf{u}\|_2^2 + \lambda \|\mathbf{W}\mathbf{c}\|_1, \tag{56}$$

or the LASSO problem [59],

$$(LS_\tau) : \min_c \frac{1}{2} \|\Psi \mathbf{c} - \mathbf{u}\|_2^2 \quad \text{subject to} \quad \|\mathbf{W}\mathbf{c}\|_1 \leq \tau. \tag{57}$$

It can be shown that for an appropriate choice of scalars δ , λ , and τ , the problems ($P_{1,\delta}$), (QP_λ), and (LS_τ) share the same solution [62,14,61]. Among others, the problem (QP_λ) is of particular interest as it is an unconstrained optimization problem. Numerous solvers based on the *active set* [54,36], *interior-point continuation* [21,42] and *projected gradient* [25,23,39,13,9,62,3,4] methods have been developed for solving the above formulations of the ℓ_1 -minimization problem.

In our numerical experiments, we adopt the Spectral Projected Gradient algorithm (SPGL1) proposed in [62] and implemented in the MATLAB package SPGL1 [5] to solve the ℓ_1 -minimization problem ($P_{1,\delta}$) in (21). SPGL1 is based on exploring the so-called Pareto curve, describing the tradeoff between the ℓ_2 -norm of the truncation error $\|\Psi \mathbf{c} - \mathbf{u}\|_2$ and the (weighted) ℓ_1 -norm of the solution $\|\mathbf{W}\mathbf{c}\|_1$, for successive solution iterations. At each iteration, the LASSO problem (57) is solved using the spectral projected gradient technique with a worst-case complexity of $\mathcal{O}(P \ln P)$ where P is the number of columns in Ψ . Given the error tolerance δ , a scalar equation is solved to identify a value for τ such that the (LS_τ) solution of (57) is identical to that of ($P_{1,\delta}$) in (21). Besides being efficient for large-scale systems where Ψ may not be available explicitly, the SPGL1 algorithm is specifically effective for our application of interest as the truncation error $\|\Psi \mathbf{c} - \mathbf{u}\|_2$ is known only approximately.

In the next section, we explore some aspects of the proposed scheme through its application to a 1- D (in space) linear elliptic stochastic differential equations with high-dimensional random diffusion coefficients.

4. Numerical examples

We consider the solution of a one-dimensional, i.e., $D = 1$, version of problem (2),

$$-\frac{d}{dx} \left(a(x, \omega) \frac{du(x, \omega)}{dx} \right) = 1, \quad x \in \mathcal{D} = (0, 1), \tag{58}$$

$$u(0, \omega) = u(1, \omega) = 0,$$

where the stochastic diffusion coefficient $a(x, \omega)$ is given by the expansion

$$a(x, \omega) = \bar{a} + \sigma_a \sum_{i=1}^d \sqrt{\lambda_i} \phi_i(x) y_i(\omega). \tag{59}$$

Here, $\{\lambda_i\}_{i=1}^d$ and $\{\phi_i(x)\}_{i=1}^d$ are, respectively, d largest eigenvalues and the corresponding eigenfunctions of the Gaussian covariance kernel

$$C_{aa}(x_1, x_2) = \exp \left[-\frac{(x_1 - x_2)^2}{l_c^2} \right], \tag{60}$$

in which l_c is the correlation length of $a(x, \omega)$ that prescribes the decay of the spectrum of C_{aa} in (60). Random variables $\{y_i(\omega)\}_{i=1}^d$ are assumed to be independent and uniformly distributed on $[-1, 1]$. The coefficient σ_a controls the variability of $a(x, \omega)$.

We verify the accuracy and efficiency of the present sparse approximation schemes for both moderate and high-dimensional diffusion coefficient $a(x, \omega)$. These two cases are obtained, respectively, by assuming $(l_c, d) = (1/5, 14)$ and $(l_c, d) = (1/14, 40)$ in (60) and (59). We further assume that $\bar{a} = 0.1$, $\sigma_a = 0.03$ when $d = 14$, and $\sigma_a = 0.021$ when $d = 40$. These choices ensure that all realizations of $a(x, \omega)$ are strictly positive on $\mathcal{D} = (0, 1)$. Table 1 summarizes the assumed parameters for the two test cases.

For both cases, the spatial discretization is done by the Finite Element Method using quadratic elements. A mesh convergence analysis is performed to ensure that spatial discretization errors are inconsequential.

The solution statistics are computed using the conventional Monte Carlo simulation, the isotropic sparse grid stochastic collocation with the Clenshaw–Curtis abscissas [66,1], and the proposed sparse approximation techniques. We use the Basis Pursuit Denoising (BPDN) solver implemented in SPGL1 [5,62] to solve the ℓ_1 -minimization problem ($P_{1,\delta}$) and the Orthogonal Matching Pursuit (OMP) solver in SparseLab [29] to approximate the ℓ_0 -minimization problem ($P_{0,\delta}$). We compare the errors in the mean, standard deviation, and root mean-squares of the solution error at $x = 0.5$ using the above methods. The details of the analysis are reported below.

Table 1Choices of parameters defining the stochastic description of diffusion coefficient $a(x, \omega)$ in Eq. (59).

Case	\bar{a}	σ_a	l_c	d
I	0.1	0.030	1/5	14
II	0.1	0.021	1/14	40

4.1. Case I: $d = 14$

We consider an increasing number $N = \{29, 120, 200, 280, 360, 421, 600\}$ of random solution samples to evaluate the solution u at $x = 0.5$ and, subsequently, to compute the PC coefficients of the solution using ℓ_1 - and ℓ_0 -minimization. These samples are nested in the sense that we recycle the previous samples when we perform calculations with larger sample sizes. The nested sampling property of our scheme is of paramount importance in large scale calculations where the computational cost of each solution evaluation is enormous. We note that sample sizes $N = 29$ and $N = 421$, respectively, correspond to the number of nested abscissas in the level $l = 1$ and level $l = 2$ of the isotropic stochastic collocation with the Clenshaw–Curtis rule.

As elucidated in Section 3.3, the accuracy of our sparse reconstruction depends on the mutual coherence $\mu(\Psi)$, the sample size N , and the truncation error $\|\Psi\mathbf{c} - \mathbf{u}\|_2$ (hence δ). In order to reduce the approximation error, we need to reduce $\|\Psi\mathbf{c} - \mathbf{u}\|_2$, which may be done by increasing p and, therefore, P . However, with a fixed number N of samples, an increase in P may result in a larger mutual coherence and, thus, the degradation of the reconstruction accuracy. Therefore, in practice, we start by approximating the lower order PC expansions when N is small and increase p when larger number of samples become available. Notice that such an adaptivity with respect to the order p is a natural way of refining the accuracy of PC expansions, for instance, when the intrusive stochastic Galerkin scheme is adopted [38]. In particular, in this example, for sample sizes $N = \{29, 120\}$, we attempt to estimate the coefficients of the 3rd-order Legendre PC expansion, i.e. $p = 3$ and $P = 680$. For larger sample sizes N , we also include the first 320 basis function from the 4th-order chaos, thus resulting in $P = 1000$. Since all of the 4th-order basis functions are not employed, we need to describe the ordering of our basis construction. We sort the elements of $\{\psi_\alpha(\mathbf{y})\}$ such that, for any given order p , the random variables y_i with smaller indices i contribute first in the basis.

For each analysis, we estimate the truncation error tolerance δ based on the cross-validation algorithm described in Section 3.5. For each N , we use $N_r \approx 3N/4$ of the samples (reconstruction set) to compute the PC coefficients \mathbf{c}^{1, δ_r} and the rest of the samples (validation set) are used to evaluate the truncation error δ_v . The cross-validation is performed for four replications of reconstruction and validation sample sets. We then find the value $\hat{\delta}_r$ that minimizes the average of δ_v over the four replications of the cross-validation samples. Given an estimate of the truncation error tolerance $\delta \approx \sqrt{4/3} \hat{\delta}_r$, we then use all N samples to compute the coefficients $\mathbf{c}^{1, \delta}$.

Fig. 2 compares the ‘exact’ PC coefficients with those obtained using BPDN and OMP solvers. We only demonstrate the results corresponding to sample sizes $N = \{120, 600\}$. An ‘exact’ solution is computed using the level $l = 8$ stochastic collocation for which the approximation errors are negligible in our comparisons. We observe that BPDN tends to give less sparse solutions compared to OMP. This is due to the facts that (i) the solution is not exactly sparse, i.e., there are many non-zero (although negligible) coefficients c_α , (ii) the ℓ_1 cost function does not impose a sufficiently large penalty on the small coefficients as does the ℓ_0 cost function, and (iii) the truncation error tolerance δ may be under-estimated. To reduce this issue, a number of modifications, including the reweighted ℓ_1 -minimization [19,68,47,41], have been introduced in the literature that are the subjects of our future work. In contrary, OMP results in more sparse solutions as it adds basis function one-at-a-time until the residual falls below the truncation error. However, as is seen in Fig. 2(b) and (d), a number of small coefficients are still over-estimated. This is primarily due to under-estimation of the truncation error tolerance δ in the cross-validation algorithm.

The convergence of the mean, standard deviation, and root mean-squares of the approximation error for $u(0.5, \mathbf{y})$ is illustrated in Fig. 3(a)–(c), respectively. For the case of stochastic collocation, we apply sparse grid quadrature (cubature) integration rule to directly compute the mean and the standard deviation. The root mean-squares error of the Monte Carlo and the stochastic collocation solution are evaluated by estimating the corresponding PC coefficients using sampling and sparse grid quadrature integration, respectively, and then comparing them with the exact coefficients.

To make a meaningful comparison, for each N , the samples used to compute the solution statistics by the conventional Monte Carlo, BPDN, and OMP are identical. In this sense, the sparse approximation using ℓ_1 - and ℓ_0 -minimizations may be viewed as only post-processing steps in the Monte Carlo simulation. As the sample sizes are finite, the estimates of the PC coefficients \mathbf{c} are sample dependent and are in fact random. To demonstrate the convergence of the algorithm with respect to different sets of samples, we repeat the analysis for two independent sets of N samples and report the corresponding statistics errors with solid and dashed lines. Although for different solution samples of size N the estimates of \mathbf{c} and the solution statistics are not identical, the approximation converges, with large probability, for any set of samples with sufficiently large size N (see Theorem 3.6).

Fig. 3(d) illustrates the statistical estimation of δ using the cross-validation approach described in Section 3.5. The estimation of δ is slightly different in BPDN and OPM, this is a consequence of different reconstruction accuracy of these

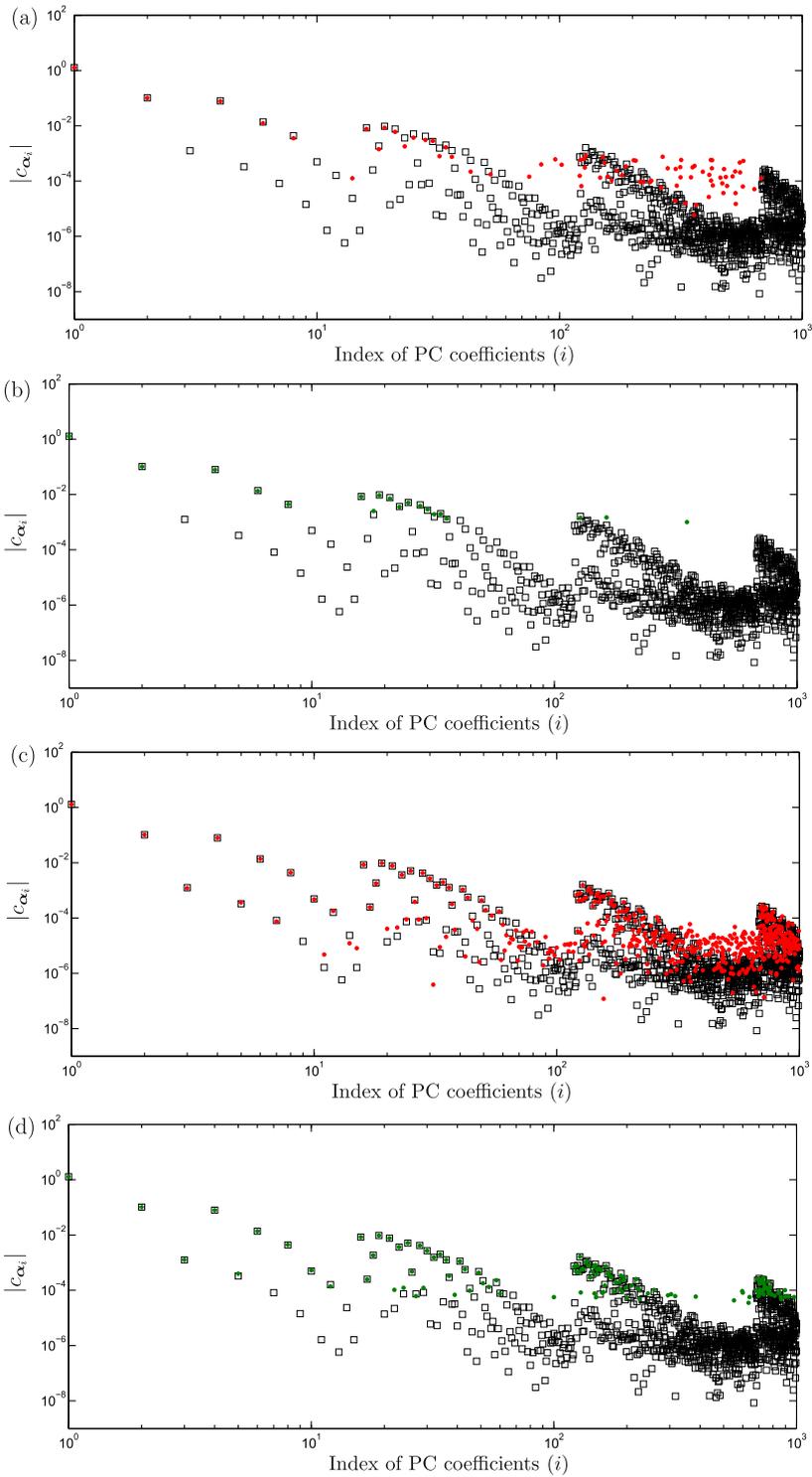


Fig. 2. Approximation of polynomial chaos (PC) coefficients \mathbf{c} of $u(0.5, \mathbf{y})$ using BPDN and OMP for $d = 14$. (a) BPDN with $N = 120$ samples, (b) OMP with $N = 120$ samples, (c) BPDN with $N = 600$ samples, and (d) OMP with $N = 600$ samples. ‘Exact’ coefficients computed from level 8 stochastic collocation with the Clenshaw–Curtis abscissas (\square); BPDN and OMP (\bullet).

techniques. Moreover, the solution of the BPDN algorithm is less sensitive to small perturbations in the truncation error δ compared to that of the OMP algorithm. This is justified by the fact that the ℓ_0 -norm is highly discontinuous.

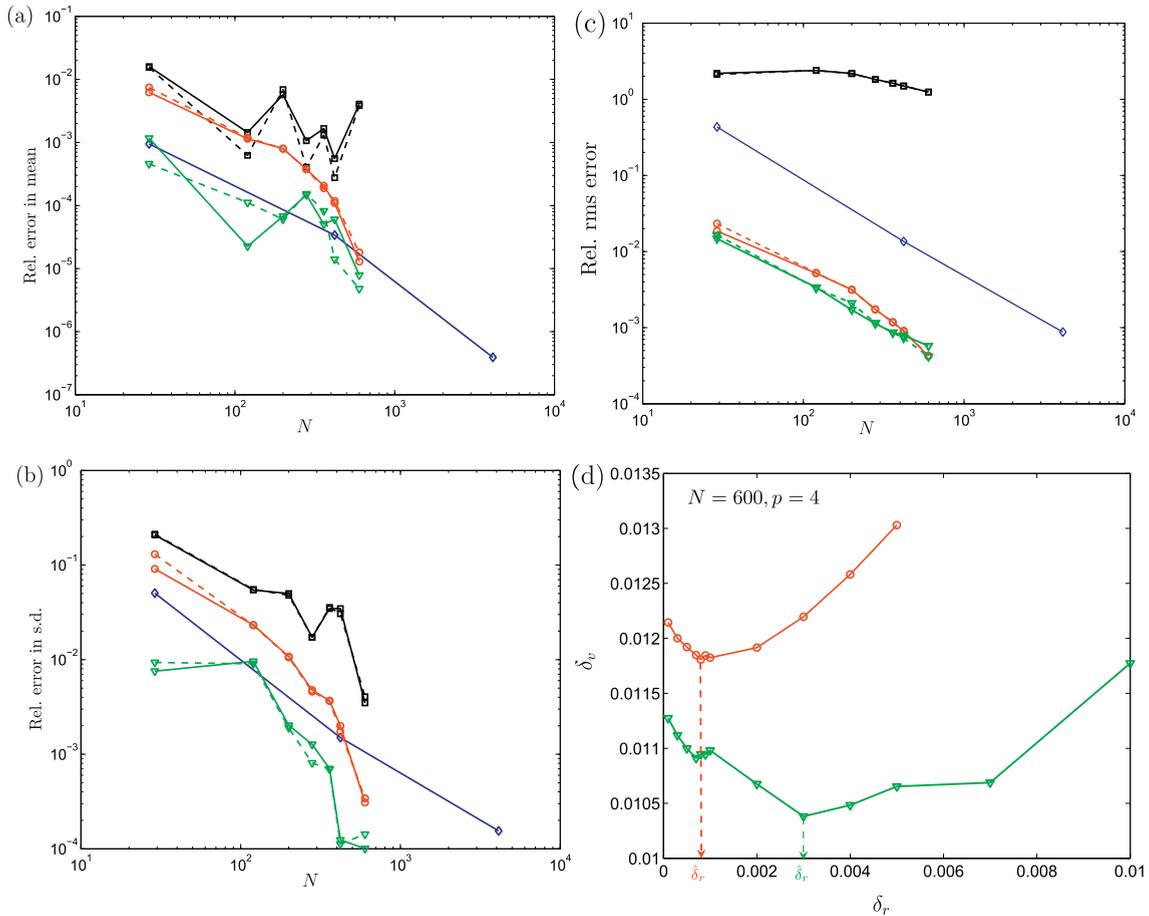


Fig. 3. Comparison of relative error in solution statistics at $x = 0.5$ for the Monte Carlo simulation, isotropic sparse grid stochastic collocation with the Clenshaw–Curtis abscissas, and the proposed sparse approximations (BPDN and OMP) for $d = 14$. Two sets of independent random realizations of $u(0.5, \mathbf{y})$ are generated first and are used for the Monte Carlo simulation, BPDN, and OMP. The solid and dashed lines correspond to the first and second sets of samples, respectively. (a) Relative error in the mean; (b) relative error in the standard deviation; (c) relative root-mean-squares (rms) error; and (d) estimation of δ using cross-validation: δ is computed from δ_r for which δ_v is minimum. (Monte Carlo simulation (\square); stochastic collocation (\diamond); BPDN (\circ); OMP (∇)).

Remark. Despite the conventional implementation of the stochastic collocation approach where the approximation refinement requires a certain number of extra samples, the ℓ_1 - and ℓ_0 -minimizations may be implemented using arbitrary numbers of additional samples, which is an advantage, particularly, when only a limited number of samples can be afforded.

4.2. Case II: $d = 40$

The objective of this example is to highlight that a sparse reconstruction may lead to significant computational savings for problems with high-dimensional random inputs. Similar to the analysis of Case I described in Section 4.1, we compute the solution statistics using multiple numbers of independent samples. More specifically, we evaluate the solution at $x = 0.5$ for independent samples of size $N = \{81, 200, 400, 600, 800, 1000\}$. The number of grid points in the level $l = 1$ and $l = 2$ of the Clenshaw–Curtis rule in dimension $d = 40$ is $N = 81$ and $N = 3281$, respectively. To obtain a reference solution, the 3rd order PC coefficients \mathbf{c} of the solution at $x = 0.5$ are computed using level $l = 5$ stochastic collocation with the Clenshaw–Curtis rule.

For $N = \{81, 200\}$ we only estimate the coefficients associated with the 2nd-order PC expansion, i.e., $p = 2$ and $P = 861$. For larger sample sizes, we also include the first 639 basis functions from the 3rd-order chaos, thus leading to $P = 1500$. For each combination of N and p , we estimate the truncation error δ using an identical cross-validation procedure described in Section 4.1. Fig. 4 illustrates the estimation of PC coefficients of $u(0.5, \mathbf{y})$ with BPDN and OMP algorithms with $N = 200$ and $N = 1000$. We again note that the recovered solution from the BPDN algorithm is less sparse as compared to that of the OMP approach, although the over-estimated coefficients (mostly from the second order terms) are indeed small. As the sample size N is increased, we are naturally able to recover more dominant coefficients on the expansion. Fig. 5 depicts the

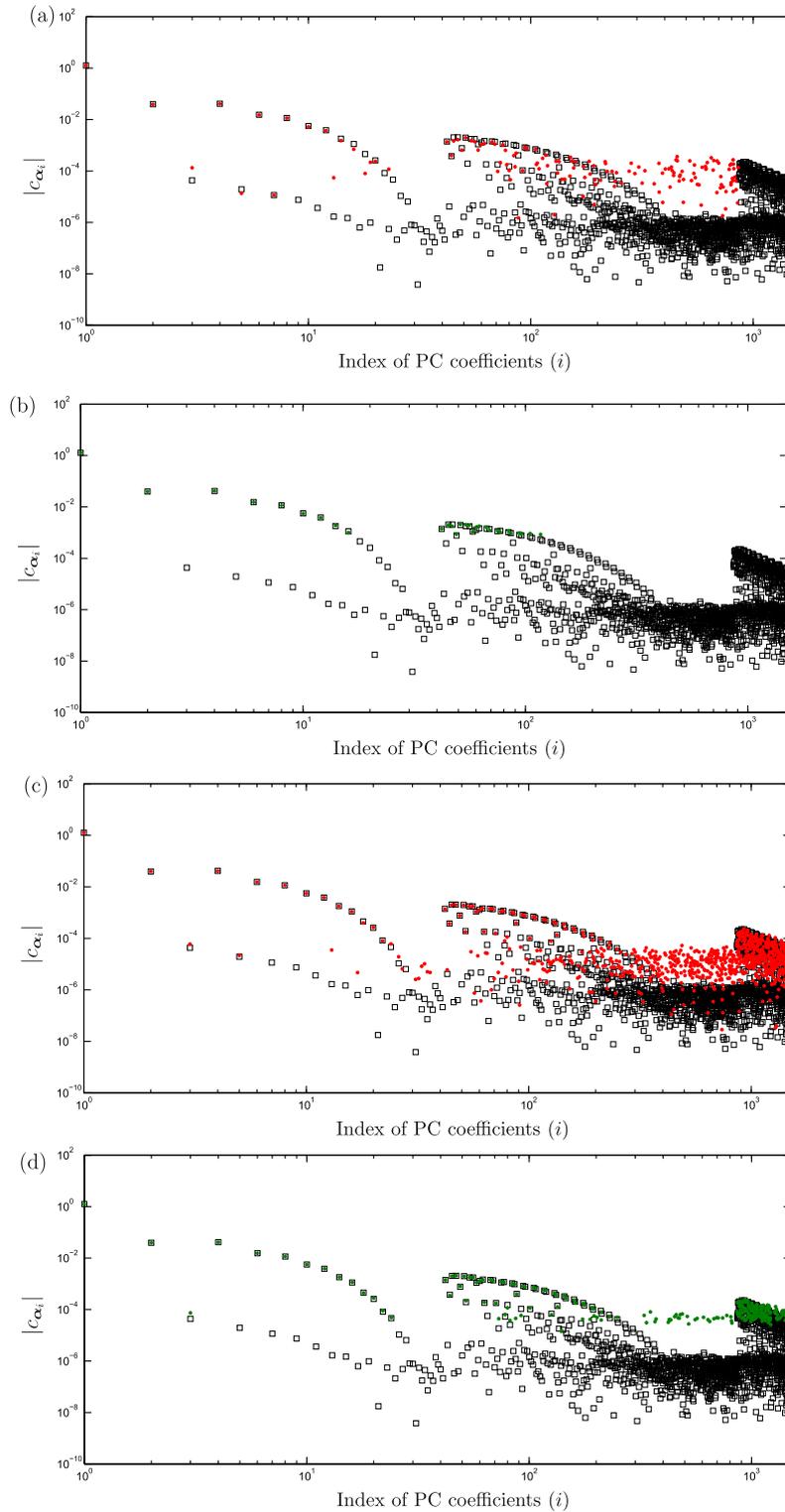


Fig. 4. Approximation of polynomial chaos (PC) coefficients \mathbf{c} of $u(0.5, \mathbf{y})$ using BPDN and OMP for $d = 40$. (a) BPDN with $N = 200$ samples, (b) OMP with $N = 200$ samples, (c) BPDN with $N = 1000$ samples, and (d) OMP with $N = 1000$ samples. ‘Exact’ coefficients computed from level 5 stochastic collocation with the Clenshaw–Curtis abscissas (\square); BPDN and OMP (\bullet).

convergence of the statistics of the solution as functions of the sample size N as well as one instance of the estimation of the truncation error tolerance δ . The implementation details are similar to those described in Section 4.1 for the case of $d = 14$.

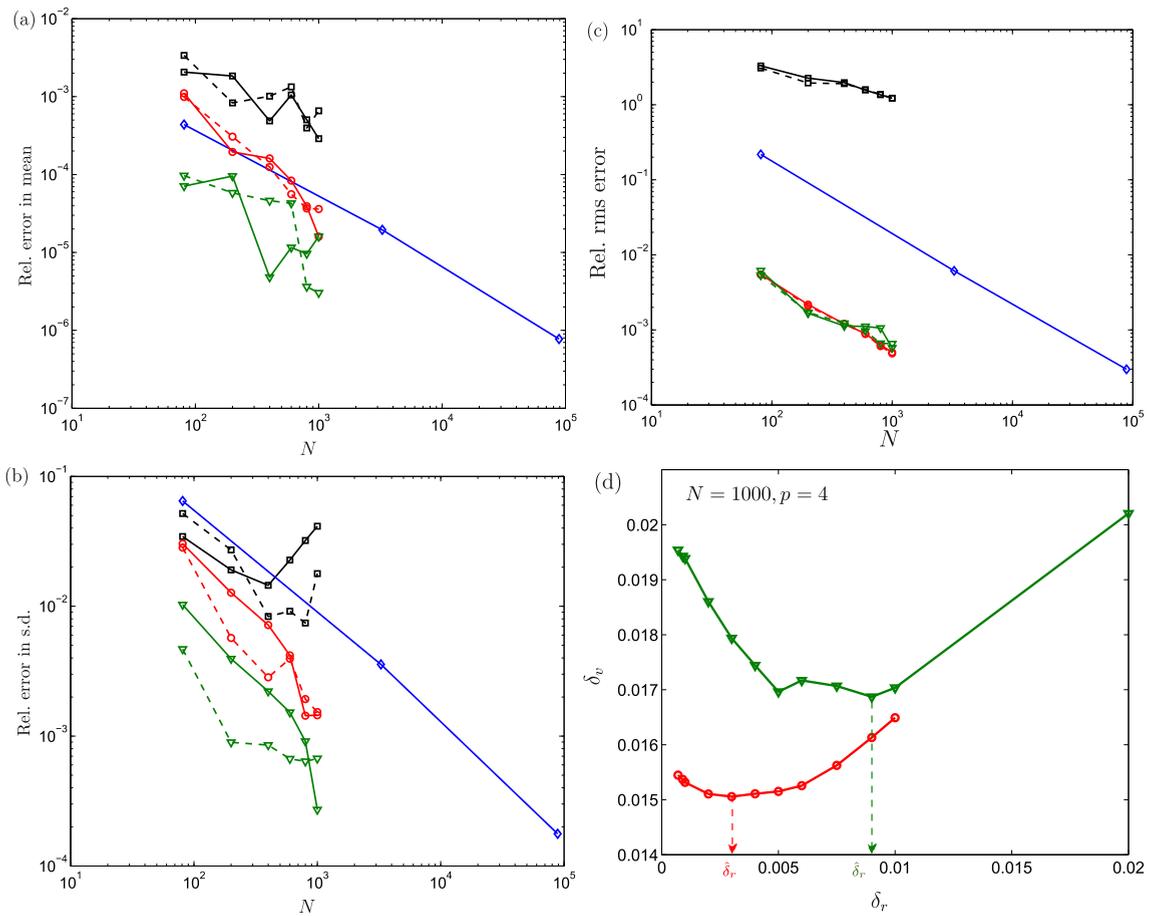


Fig. 5. Comparison of relative error in solution statistics at $x = 0.5$ for the Monte Carlo simulation, isotropic sparse grid stochastic collocation with the Clenshaw–Curtis abscissas, and the proposed sparse approximations (BPDN and OMP) for $d = 40$. Two sets of independent random realizations of $u(0.5, \mathbf{y})$ are generated first and are used for the Monte Carlo simulation, BPDN, and OMP. The solid and dashed lines correspond to the first and second sets of samples, respectively. (a) Relative error in mean; (b) relative error in standard deviation; (c) relative root-mean-squares (rms) error; and (d) estimation of δ using cross-validation: δ is computed from δ_r for which δ_v is minimum. (Monte Carlo simulation (\square); stochastic collocation (\diamond); BPDN (\circ); OMP (∇)).

It is worth highlighting that the computational saving of the present sparse approximations (in terms of the number of samples needed to achieve a certain accuracy) compared to the isotropic sparse grid collocation is even larger for the higher-dimensional ($d = 40$) problem. This is due to the fact that the number of samples needed to recover the solution using ℓ_1 - and ℓ_0 -minimization is dictated more by the number of dominant terms in the PC expansion compared to the total number of terms P , as in Theorem 3.1.

5. Conclusion

The present study proposes a *non-intrusive* and *non-adapted* approach based on the *compressive sampling* formalism for the approximation of sparse solution of stochastic PDEs. When sufficiently sparse in the polynomial chaos (PC) basis, the compressive sampling enables an accurate recovery of the solution using a number of random solution samples that is significantly smaller than the cardinality of the PC basis. Sparse PC approximations based on ℓ_0 - and ℓ_1 -minimization approaches have been introduced and implemented using the Basis Pursuit Denoising (BPDN) and the Orthogonal Matching Pursuit (OMP) algorithms, respectively. Probabilistic bounds based on the *concentration of measure* phenomenon have been derived to verify the convergence and the stability of the present sparse constructions. The performance and efficiency of the proposed techniques are explored through their application to a linear elliptic PDE with high-dimensional random diffusion coefficients where the sparsity of the solution with respect to the PC basis is guaranteed. The proposed formalism to recover the sparse PC expansion of stochastic functions is not restricted to the case of the elliptic PDEs, as its underlying applicability assumptions are universal. Although the discussions of this work have been focused on the particular case of the Legendre PC expansions, the proposed framework can be readily extended to other bases such as the Hermite PC expansion [38] (when the random variables \mathbf{y} are standard Gaussian) or multi-wavelet expansions [43].

Acknowledgments

The first author acknowledges the support of the United States Department of Energy under Stanford's Predictive Science Academic Alliance Program (PSAAP) for the preliminary stages of his work. The second author acknowledges the support of the National Science Foundation via NSF Grant CMMI-092600 and of the United States Department of Energy under Caltech's Predictive Science Academic Alliance Program (PSAAP).

References

- [1] I. Babuška, F. Nobile, R. Tempone, A stochastic collocation method for elliptic partial differential equations with random input data, *SIAM J. Numer. Anal.* 45 (3) (2007) 1005–1034.
- [2] I. Babuška, R. Tempone, G. Zouraris, Galerkin finite element approximations of stochastic elliptic partial differential equations, *SIAM J. Numer. Anal.* 42 (2) (2004) 800–825.
- [3] A. Beck, M. Teboulle, A fast iterative shrinkage-threshold algorithm for linear inverse problems, *SIAM J. Imaging Sciences* 2 (2009) 183–202.
- [4] S. Becker, J. Bobin, E.J. Candes, NESTA: a fast and accurate first-order method for sparse recovery, ArXiv e-prints, 2009. Available from <<http://arxiv.org/abs/0904.3367>>.
- [5] E. van den Berg, M.P. Friedlander, SPGL1: A solver for large-scale sparse reconstruction, June 2007. Available from <<http://www.cs.ubc.ca/labs/scl/spgl1>>.
- [6] M. Bieri, A sparse composite collocation finite element method for elliptic spdes, Technical Report Research Report No. 2009-08, Seminar für Angewandte Mathematik, SAM, Zürich, Switzerland, 2009.
- [7] M. Bieri, R. Andreev, C. Schwab, Sparse tensor discretization of elliptic spdes, Technical Report Research Report No. 2009-07, Seminar für Angewandte Mathematik, SAM, Zürich, Switzerland, 2009.
- [8] M. Bieri, C. Schwab, Sparse high order FEM for elliptic sPDEs, *Comput. Methods Appl. Mech. Eng.* 198 (2009) 1149–1170.
- [9] J.M. Bioucas-Dias, M.A.T. Figueiredo, A new TwIST: two-step iterative shrinking/thresholding algorithms for image restoration, *IEEE Trans. Image Process.* 16 (12) (2007) 2992–3004.
- [10] G. Blatman, Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis, Ph.D. Thesis, Université Blaise, Pascal – Clermont –Ferrand II, 2009.
- [11] G. Blatman, B. Sudret, An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis, *Probabilist. Eng. Mech.* 25 (2) (2010) 183–197.
- [12] P. Boufounos, M.F. Duarte, R.G. Baraniuk, Sparse signal reconstruction from noisy compressive measurements using cross validation, in: *SSP '07: Proceedings of the 2007 IEEE/SP 14th Workshop on Statistical Signal Processing*, IEEE Computer Society, 2007, pp. 299–303.
- [13] K. Bredies, D.A. Lorenz, Linear convergence of iterative soft-thresholding, *SIAM J. Sci. Comp.* 30 (2) (2008) 657–683.
- [14] A.M. Bruckstein, D.L. Donoho, M. Elad, From sparse solutions of systems of equations to sparse modeling of signals and images, *SIAM Rev.* 51 (1) (2009) 34–81.
- [15] E.J. Candes, J. Romberg, Quantitative robust uncertainty principles and optimally sparse decompositions, *Found. Comput. Math.* 6 (2) (2006) 227–254.
- [16] E.J. Candes, J. Romberg, Sparsity and incoherence in compressive sampling, *Inverse Probl.* 23 (3) (2007) 969–985.
- [17] E.J. Candes, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inform. Theory* 52 (2) (2006) 489–509.
- [18] E.J. Candes, T. Tao, Near optimal signal recovery from random projections: universal encoding strategies?, *IEEE Trans Inform. Theory* 52 (12) (2006) 5406–5425.
- [19] E.J. Candes, M.B. Wakin, S. Boyd, Enhancing sparsity by reweighted ℓ_1 minimization, *J. Fourier Anal. Appl.* 14 (5) (2008) 877–905.
- [20] S.S. Chen, D.L. Donoho, M. Saunders, Atomic decomposition by basis pursuit, *SIAM J. Sci. Comput.* 20 (1998) 33–61.
- [21] S.S. Chen, D.L. Donoho, M. Saunders, Atomic decomposition by basis pursuit, *SIAM Rev.* 43 (1) (2001) 129–159.
- [22] A. Cohen, W. Dahmen, R. DeVore, Compressed sensing and best k -term approximation, *J. Am. Math. Soc.* 22 (2009) 211–231.
- [23] P.L. Combettes, V.R. Wajs, Signal recovery by proximal forward-backward splitting, *Multiscale Model. Simul.* 4 (4) (2005) 1168–1200.
- [24] W. Dai, O. Milenkovic, Subspace pursuit for compressive sensing signal reconstruction, *IEEE Trans. Inform. Theory* 55 (5) (2009) 2230–2249.
- [25] I. Daubechies, M. DeFrise, C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, *Commun. Pure Appl. Math.* 57 (11) (2004) 1413–1457.
- [26] T. Blumensath, M.E. Davies, Iterative hard thresholding for compressed sensing, *Appl. Comp. Harm. Anal.* 27 (3) (2009) 265–274.
- [27] G. Davis, S. Mallat, M. Avellaneda, Adaptive greedy approximation, *J. Constr. Approx.* 13 (1997) 57–98.
- [28] M.K. Deb, I. Babuska, J.T. Oden, Solution of stochastic partial differential equations using Galerkin finite element techniques, *Comput. Methods Appl. Mech. Eng.* 190 (2001) 6359–6372.
- [29] D.L. Donoho, I. Drori, V.C. Stodden, Y. Tsaig, Available from <<http://www-stat.stanford.edu/~sparselab/>>.
- [30] D.L. Donoho, Compressed sensing, *IEEE Trans. Inform. Theory* 52 (4) (2006) 1289–1306.
- [31] D.L. Donoho, I. Drori, Y. Tsaig, J.L. Starck, Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit, Technical report, 2006.
- [32] D.L. Donoho, M. Elad, V.N. Temlyakov, Stable recovery of sparse overcomplete representations in the presence of noise, *IEEE Trans. Inform. Theory* 52 (1) (2006) 6–18.
- [33] A. Doostan, R. Ghanem, J. Red-Horse, Stochastic model reduction for chaos representations, *Comput. Methods Appl. Mech. Eng.* 196 (37–40) (2007) 3951–3966.
- [34] A. Doostan, G. Iaccarino, A least-squares approximation of partial differential equations with high-dimensional random inputs, *J. Comput. Phys.* 228 (12) (2009) 4332–4345.
- [35] A. Doostan, H. Owhadi, A. Lashgari, G. Iaccarino, Non-adapted sparse approximation of PDEs with stochastic inputs, Technical report, 2009. Center for Turbulence Research, Annual Research Brief, Stanford University (available from <http://ctr.stanford.edu/ResBriefs09/10_doostan.pdf>).
- [36] B. Efron, T. Hastie, L. Johnstone, R. Tibshirani, Least angle regression, *Ann. Stat.* 32 (2004) 407–499.
- [37] M.A. T Figueiredo, R.D. Nowak, S.J. Wright, Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems, *IEEE J. Sel. Top. Signal Process.* 1 (4) (2007) 586–597.
- [38] R. Ghanem, P. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Dover, 2002.
- [39] E.T. Hale, W. Yin, Y. Zhang, Fixed-point continuation for ℓ_1 -minimization: methodology and convergence, *SIAM J. Optim.* 19 (3) (2008) 1107–1130.
- [40] S. Hosder, R.W. Walters, R. Perez, A non-intrusive polynomial chaos method for uncertainty propagation in CFD simulations, in: *44th AIAA Aerospace Sciences Meeting and Exhibit*, AIAA-2006-891, Reno (NV), 2006.
- [41] M.A. Khajehnejad, W. Xu, A.S. Avestimehr, B. Hassibi, Improved sparse recovery thresholds with two-step reweighted ℓ_1 minimization, ArXiv e-prints, 2010. Available from <http://arxiv.org/abs/1004.0402>.
- [42] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, D. Gorinevsky, An interior-point method for large-scale ℓ_1 -regularized least squares, *IEEE J. Sel. Top. Signal Process.* 1 (4) (2007) 606–617.
- [43] O.P. Le Maitre, H. Najm, R. Ghanem, O. Knio, Multi-resolution analysis of Wiener-type uncertainty propagation schemes, *J. Comput. Phys.* 197 (2) (2004) 502–531.

- [44] X. Ma, N. Zabarar, An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations, *J. Comput. Phys.* 228 (2009) 3084–3113.
- [45] L. Mathelin, M.Y. Hussaini, A stochastic collocation algorithm for uncertainty analysis, Technical Report NAS 1.26:212153; NASA/CR-2003-212153, NASA Langley Research Center, 2003.
- [46] C. McDiarmid, On the method of bounded differences, in: *Surveys in Combinatorics*, Cambridge University Press, pp. 148–188.
- [47] D. Needell, Noisy signal recovery via iterative reweighted ℓ_1 -minimization, in: *Proceedings of the Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, 2009.
- [48] D. Needell, J.A. Tropp, CoSaMP: iterative signal recovery from incomplete and inaccurate samples, *Appl. Comput. Harmonic Anal.* 26 (3) (2008) 01–321.
- [49] D. Needell, R. Vershynin, Signal recovery from incomplete and inaccurate measurements via Regularized Orthogonal Matching Pursuit, *ArXiv e-prints*, 2007.
- [50] F. Nobile, R. Tempone, C.G. Webster, An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data, *SIAM J. Numer. Anal.* 46 (5) (2008) 2411–2442.
- [51] F. Nobile, R. Tempone, C.G. Webster, A sparse grid stochastic collocation method for partial differential equations with random input data, *SIAM J. Numer. Anal.* 46 (5) (2008) 2309–2345.
- [52] A. Nouy, A generalized spectral decomposition technique to solve a class of linear stochastic partial differential equations, *Comput. Methods Appl. Mech. Eng.* 196 (37–40) (2007) 4521–4537.
- [53] A. Nouy, Generalized spectral decomposition method for solving stochastic finite element equations: invariant subspace problem and dedicated algorithms, *Comput. Methods Appl. Mech. Eng.* 197 (2008) 4718–4736.
- [54] M.R. Osborne, B. Presnell, B. Turlach, A new approach to variable selection in least squares problems, *IMA J. Numer. Anal.* 20 (2000) 389–403.
- [55] Y.C. Pati, R. Rezaifar, P.S. Krishnaprasad, Orthogonal Matching Pursuit: Recursive function approximation with applications to wavelet decomposition, in: *Proceedings of the 27th Annual Asilomar Conference on Signals, Systems, and Computers*, 1993, pp. 40–44.
- [56] H. Rauhut, R. Ward, Sparse legendre expansions via ℓ_1 -minimization, *ArXiv e-prints*, 2010.
- [57] C. Schwab, R.A. Todor, Karhunen–Loève approximation of random fields by generalized fast multipole methods, *J. Comput. Phys.* 217 (1) (2006) 100–122.
- [58] M.A. Tatang, Direct incorporation of uncertainty in chemical and environmental engineering systems, PhD thesis, Massachusetts Institute of Technology, 1995.
- [59] R. Tibshirani, Regression shrinkage and selection via the Lasso, *J. Roy. Stat. Soc. Ser. B* 58 (1) (1996) 267–288.
- [60] R.A. Todor, C. Schwab, Convergence rates for sparse chaos approximations of elliptic problems with stochastic coefficients, *IMA J. Numer. Anal.* 27 (2) (2007) 232–261.
- [61] J.A. Tropp, S.J. Wright, Computational methods for sparse solution of linear inverse problems, in: *Proceedings of the IEEE*, in press.
- [62] E. van den Berg, M.P. Friedlander, Probing the Pareto frontier for basis pursuit solutions, *SIAM J. Sci. Comput.* 31 (2) (2008) 890–912.
- [63] X. Wan, G. Karniadakis, An adaptive multi-element generalized polynomial chaos method for stochastic differential equations, *J. Comput. Phys.* 209 (2005) 617–642.
- [64] X. Wan, B. Rozovskii, G.E. Karniadakis, A stochastic modeling methodology based on weighted Wiener chaos and Malliavin calculus, *Proc. National Acad. Sci.* 106 (34) (2009) 14189–14194.
- [65] R. Ward, Compressed sensing with cross validation, *IEEE Trans. Inform. Theory* 55 (12) (2009) 5773–5782.
- [66] D. Xiu, J.S. Hesthaven, High-order collocation methods for differential equations with random inputs, *SIAM J. Sci. Comput.* 27 (3) (2005) 1118–1139.
- [67] D. Xiu, G.M. Karniadakis, The Wiener–Askey polynomial chaos for stochastic differential equations, *SIAM J. Sci. Comput.* 24 (2) (2002) 619–644.
- [68] W. Xu, M.A. Khajehnejad, A.S. Avestimehr, B. Hassibi, Breaking through the thresholds: an analysis for iterative reweighted ℓ_1 minimization via the Grassmann Angle Framework, *ArXiv e-prints*, 2009. Available from <<http://arxiv.org/abs/0904.0994>>.
- [69] J. Yang, Y. Zhang, Alternating direction algorithms for ℓ_1 -problems in compressive sensing, *ArXiv e-prints*, 2009. Available from <<http://arxiv.org/abs/0912.1185>>.