

Sparse multiresolution stochastic approximation for uncertainty quantification

By D. Schiavazzi, A. Doostan AND G. Iaccarino

1. Motivation and objectives

Most physical systems are inevitably affected by uncertainties due to natural variabilities or incomplete knowledge about their governing laws. To achieve predictive computer simulations of such systems, a major task is, therefore, to study the impact of these uncertainties on response quantities of interest. Within the probabilistic framework, uncertainties may be represented in the form of random variables/processes. Several computational strategies may then be applied to propagate these uncertainties in order to obtain the statistics of quantities of interest.

Monte Carlo (MC) sampling methods have been widely used for this purpose. Besides their generally slow convergence, they offer a blend of simplicity, robustness, and efficiency for high-dimensional problems. Recently, there has been an increasing interest in developing more efficient alternative numerical methods. Most notably, stochastic approximation schemes based on Wiener-Askey polynomial chaos expansions (Ghanem & Spanos 2003; Xiu & Karniadakis 2002) or interpolations on sparse grids (Xiu & Hesthaven 2006) have been applied successfully to a variety of problems with random inputs. For sufficiently smooth solutions, these methods achieve as high as exponential convergence rate in the mean-squares sense. For non-smooth responses such as those exhibiting sharp gradients or discontinuities, however, these methods may result in poor approximations. To address this shortcoming, various methodologies including adaptive multiresolution (Le Maître *et al.* 2004), multi-element (generalized) polynomial chaos (Wan & Karniadakis 2005), adaptive sparse grid collocation (Ma & Zabarar 2009), rational approximation (Chantrasmī *et al.* 2009), and simplex stochastic collocation (Witteveen & Iaccarino 2012), have been suggested in recent years.

In this work, we present a multiresolution approach based on sparse multiwavelet expansions for uncertainty propagation. Unlike the work in Le Maître *et al.* (2004) where the multiwavelet coefficients are computed via Galerkin projections (typically requiring modifications of deterministic physics solvers), we propose a Compressive Sampling (CS) strategy in which physics solvers are treated as “black boxes”. CS is a new direction in signal processing that enables (up to) exact reconstruction of signals admitting sparse representations, in some suitable basis, using samples obtained at a sub-Nyquist rate (Candes & Tao 2005; Bruckstein *et al.* 2009). In Doostan & Owhadi (2011), the use of CS in the approximation of sparse polynomial chaos solutions to stochastic PDEs was first introduced. In the present study, we demonstrate the efficiency of CS in recovering stochastic functions having sparse expansions in multiwavelet bases. At the core of the proposed CS framework are the application of a greedy basis pursuit technique together with an Importance Sampling strategy that enable accurate reconstruction of non-smooth stochastic functions. To the best of the authors’ knowledge, no previous study is available in the literature where the combination of the three elements of multiresolution approxi-

mations, CS, and Importance Sampling have been applied for the purpose of uncertainty propagation.

1.1. Problem setup

Consider a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ in which Ω is the sample space, \mathcal{F} is the σ -algebra of events, and \mathcal{P} defines a probability measure on \mathcal{F} . A vector of independent and identically distributed random variables with joint probability density function $\rho(\mathbf{y}) : \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}$ is indicated by $\mathbf{y} = (y_1, \dots, y_d)$ with $y_i : \Omega \rightarrow \mathbb{R}$, $i = 1, \dots, d$, and $d \in \mathbb{N}$. We state our problem as an evolution in time of $\mathbf{u}(\mathbf{y}, t) : \Omega \times [0, T] \rightarrow \mathbb{R}^q$, $q \in \mathbb{N}$, such that

$$\frac{\partial}{\partial t} \mathbf{u}(\mathbf{y}, t) = \mathbf{f}(\mathbf{u}, t), \quad t \in [0, T], \mathbf{y} \in \Omega, \quad \text{with} \quad \mathbf{u}(\mathbf{y}, t = 0) = \mathbf{u}_0(\mathbf{y}) \quad (1.1)$$

hold \mathcal{P} -a.s. in Ω . Here we assume the well-posedness (in \mathcal{P} -a.s. sense) of (Eq. 1.1) with respect to the choices of the forcing and boundary functions \mathbf{f} and \mathbf{u}_0 , respectively. In the present study, we seek to construct a multiresolution representation of $\mathbf{u}(\mathbf{y}, t)$ at a fixed time $t_a \in [0, T]$ by using samples $\{\mathbf{u}(\mathbf{y}^{(k)}, t_a) : k = 1, \dots, M\}$ corresponding to M realizations $\{\mathbf{y}^{(k)} : k = 1, \dots, M\}$ of the random inputs \mathbf{y} . To simplify the notation and presentation, we henceforth drop the time variable t_a and describe our approach for a scalar, univariate solution $u(y)$, i.e., $q = d = 1$.

The remainder of this paper is organized as follows. In Section 2, we overview the multiresolution representation of a function $u(y) \in \mathbf{L}^2([0, 1])$ in multiwavelet bases. We describe our approach – within the Compressive Sampling framework – to obtain such a representation in Sections 3 and 4, respectively. Finally, in Section 5, we apply the proposed framework to a non-linear ODE problem.

2. Multiresolution analysis and multiwavelet approximation

A multiresolution approximation of $\mathbf{L}^2([0, 1])$ is expressed by means of a nested sequence of closed subspaces $\mathbf{V}_0 \subset \mathbf{V}_1 \subset \dots \subset \mathbf{V}_j \subset \dots \subset \mathbf{L}^2([0, 1])$, where each $\mathbf{V}_j = \text{span}\{\phi_{j,k}(y) : k = 0, \dots, 2^j - 1\}$ and

$$\phi_{j,k}(y) = 2^{j/2} \phi(2^j y - k) \quad (2.1)$$

are generated by dilations and translations of a *scaling* function $\phi(y) : [0, 1] \rightarrow \mathbb{R}$. The scaling function $\phi(y)$ is such that the closure of the union of \mathbf{V}_j , i.e., $\overline{\bigcup_{k=1}^{\infty} \mathbf{V}_k}$, is dense in $\mathbf{L}^2([0, 1])$. Let the *wavelet* subspace \mathbf{W}_j denote the orthogonal complement of \mathbf{V}_j in \mathbf{V}_{j+1} , that is $\mathbf{V}_{j+1} = \mathbf{V}_j \oplus \mathbf{W}_j$ and $\mathbf{V}_j \perp \mathbf{W}_j$. It can be shown that $\mathbf{W}_j = \text{span}\{\varphi_{j,k}(y) : k = 0, \dots, 2^j - 1\}$ where $\varphi_{j,k}(y)$ is generated from dilation and translation of a *mother* wavelet function $\varphi(y) : [0, 1] \rightarrow \mathbb{R}$, i.e.,

$$\varphi_{j,k}(y) = 2^{j/2} \varphi(2^j y - k). \quad (2.2)$$

By the construction of the wavelet spaces \mathbf{W}_j , it is straightforward to see that $\mathbf{V}_j = \mathbf{V}_0 \oplus (\bigoplus_{k=0}^j \mathbf{W}_k)$, and consequently $\mathbf{V}_0 \oplus (\bigoplus_{k=0}^{\infty} \mathbf{W}_k) = \mathbf{L}^2([0, 1])$. Therefore, any function $u(y) \in \mathbf{L}^2([0, 1])$ admits an orthogonal decomposition of the form

$$u(y) = \tilde{\alpha}_{0,0} \phi_{0,0}(y) + \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} \alpha_{j,k} \varphi_{j,k}(y), \quad (2.3)$$

where $\tilde{\alpha}_{0,0} = \langle u, \phi_{0,0} \rangle_{\mathbf{L}^2([0,1])}$ and $\alpha_{j,k} = \langle u, \varphi_{j,k} \rangle_{\mathbf{L}^2([0,1])}$. To simplify the notation, we

rewrite (Eq. 2.3) in the form

$$u(y) = \sum_{i=1}^{\infty} \alpha_i \psi_i(y), \quad (2.4)$$

in which we establish a one-to-one correspondence between elements of the basis sets $\{\psi_i : i = 0, \dots, \infty\}$ and $\{\phi_{0,0}, \varphi_{j,k} : k = 0, \dots, 2^j - 1, j = 0, \dots, \infty\}$.

In the present study we adopt the slightly more complicated multiresolution of Alpert (1993) where multiple scaling functions $\{\phi_i(y) : i = 0, \dots, m - 1\}$ are used to construct \mathbf{V}_0 . Specifically, we choose $\phi_i(y)$ as the Legendre polynomial of degree i defined on the interval $[0, 1]$. An orthonormal basis $\{\varphi_i(y) : i = 0, \dots, m - 1\}$ for \mathbf{W}_0 is also established. More precisely, let $\mathbf{U}_m = \{u(y) \in \mathbf{L}^2([0, 1]) : \int_{[0,1]} u(y) y^m dy = 0\}$ represent the subspace of functions in $\mathbf{L}^2([0, 1])$ with m vanishing moments. We then construct $\varphi_i \in \mathbf{U}_j$, $j = 0, \dots, i + m - 1$, with the orthonormality constraint $\langle \varphi_i, \varphi_j \rangle_{\mathbf{L}^2([0,1])} = \delta_{ij}$, $i, j = 0, \dots, m - 1$, where δ_{ij} is the Kronecker delta. The multiwavelet basis functions $\varphi_{j,k}$ are then generated by dilations and translations of $\{\varphi_i(y) : i = 0, \dots, m - 1\}$. The resulting basis is unique (up to the sign) and provides a generalization of Legendre and Haar representations. In particular, Legendre polynomials can be obtained by stopping the expansion at the resolution $j = 0$, while Haar wavelets are obtained for $m = 0$. If expanded in the Alpert multiwavelet basis, sparse representations are likely to be observed for piecewise smooth functions. Sharp gradients, bifurcations or discontinuities, for example in hyperbolic problems, motivate the use of such dictionaries as multiwavelet with the ability of capturing these local features for which global polynomials may not be adequate (see, e.g., Le Maître *et al.* 2004). In addition to several numerical advantages, the orthogonality property of Alpert multiwavelets is also desirable allowing first- and second-order statistics of u to be evaluated directly from the expansion coefficients. We refer the interested reader to Alpert (1993) for an in-depth derivation of the Alpert multiwavelet basis. The construction of a multiwavelet basis in $\mathbf{L}^2([0, 1]^d)$ with $d \geq 2$ – employed in the numerical example of Section 5 – is also discussed in Alpert (1993). Examples of one-dimensional Alpert multiwavelets with 3 and 4 vanishing moments, respectively, are shown in Figure 1.

3. Compressive Sampling of multiwavelet expansions

3.1. Rudiments of Compressive Sampling

Compressive Sampling (CS) is a new direction in signal processing that breaks the traditional limits of the Shannon-Nyquist sampling rate for reconstruction of sparse signals. Consider a vector of measurements $\mathbf{u} = (u(y^{(1)}), \dots, u(y^{(M)}))^T \in \mathbb{R}^M$ of $u \in \mathbf{L}^2([0, 1])$. Assuming that u admits a multiwavelet expansion of the form (Eq. 2.4) with some finite m and resolution j , \mathbf{u} can be represented as $\mathbf{u} = \mathbf{\Psi} \boldsymbol{\alpha}$, where the so-called measurement matrix $\mathbf{\Psi} \in \mathbb{R}^{M \times P}$ contains the realization of the multiwavelet basis $\{\psi_i(y)\}$ corresponding to \mathbf{u} and $\boldsymbol{\alpha} \in \mathbb{R}^P$ is the vector of unknown expansion coefficients. Here, P is the cardinality of the truncated multiwavelet basis. Then u has a sparse multiwavelet representation if $\|\boldsymbol{\alpha}\|_0 = \#\{\alpha_i : \alpha_i \neq 0\} \ll P$. For a sufficiently sparse u , CS recovers u exactly using some $M \ll P$ measurements by solving an optimization problem of the form

$$\min_{\boldsymbol{\alpha} \in \mathbb{R}^P} \|\boldsymbol{\alpha}\|_s \quad \text{subject to} \quad \mathbf{u} = \mathbf{\Psi} \boldsymbol{\alpha}. \quad (P_s)$$

The sparsest solution $\boldsymbol{\alpha}$ to (P_s) corresponds to $s = 0$, i.e., minimizing the ℓ_0 semi-norm

$\|\alpha\|_0$, which is generally NP-hard to compute. To break this complexity, several heuristics based on greedy pursuit, e.g., Orthogonal Matching Pursuit (OMP), and convex relaxation via ℓ_1 -minimization, i.e., $s = 1$, have been proposed, among other approaches. Moreover, several metrics such as the mutual coherence (Bruckstein *et al.* 2009), or the restricted isometry property (Candes & Tao 2005) have been introduced to provide guarantees on the uniqueness of the sparsest solution to (P_s) as well as the ability of the heuristic approaches in recovering the solution. In particular, the mutual coherence of Ψ (e.g., see Bruckstein *et al.* 2009) is defined as

$$\mu(\Psi) = \max_{i \neq j} \frac{|\psi_i^T \psi_j|}{\|\psi_i\|_2 \|\psi_j\|_2}, \quad (3.1)$$

where $\psi_i \in \mathbb{R}^M$ is the i -th column of Ψ . Note that $\mu(\Psi) \in [0, 1]$ in general, and that it is strictly positive for $M < P$. Depending on the sparsity level $\|\alpha\|_0$, the mutual coherence provides a sufficient condition on the number M of measurements for a successful recovery of α from P_s , as shown in Bruckstein *et al.* (2009).

3.2. Sparse multiwavelet approximation using Tree-based OMP

In the present study, we extend the greedy Tree-based OMP (TOMP) approach of La & Do (2006) and Duarte *et al.* (2005) to solve (P_0) for sparse multiwavelet coefficients α . Due to the local nature of the multiwavelet basis, it may occur, for example, that none of the realizations of the random input y – from which the measurement vector \mathbf{u} is generated – belong to the support of the i -th basis function ψ_i , resulting in $\psi_i = \mathbf{0}$. In the interest of brevity, only a short description of the adopted greedy algorithms is provided here; detailed expositions are provided in La & Do (2006) and Duarte *et al.* (2005). Both the TOMP and OMP algorithms are iterative two-staged procedures. In the first sensing stage, basis functions are selected based on the correlation of their corresponding columns in Ψ with the residual vector. While only the most correlated basis is selected in OMP, few additional candidates with high correlations are considered in TOMP. During the second reconstruction stage, each candidate in turn, with associated ancestors, is temporarily added to the active basis set and the corresponding coefficients are evaluated using a standard least squares. Only the candidate (and ancestors) generating the minimum residual is permanently added to the index set and the process is re-iterated. The TOMP algorithm in La & Do (2006) was originally developed for scalar tree representations. We, however, extend that work to account for vector multiwavelet trees, as needed in this work. Algorithm 1 summarizes the main steps of TOMP.

4. TOMP with Importance Sampling

A natural way to obtain the samples \mathbf{u} is to generate random realizations of the input y and evaluate the corresponding solution $u(y)$. However, for situations where u exhibits, for instance, sharp gradients or discontinuities, such a sampling strategy may not necessarily lead to accurate approximation because the higher-resolution basis functions needed to capture the local structure of u may not be sampled enough to constitute a well-conditioned measurement matrix Ψ . To achieve a local accumulation of samples, where needed, we here propose an Importance Sampling approach that is discussed next.

Importance Sampling strategy. Importance Sampling is a well-known variance reduction methodology in Monte Carlo estimations. Sampling is performed according to a modi-

Algorithm 1 Tree-based Orthogonal Matching Pursuit algorithm - TOMP

inputs:
The number of samples M , basis cardinality P .
Matrix Ψ
Vector \mathbf{u}
Maximum allowable iterations k_{max}
Tolerance δ .
Coefficients γ, ρ

outputs:
Solution vector α .

initialize:
 $k \leftarrow 0$.
Initialize residual $\mathbf{r}_0 \leftarrow \mathbf{u}$.
Initialize index set $\mathcal{I}_0 \leftarrow \{0\}$.
 $\mathbf{W} \leftarrow \text{diag}(\|\psi_i\|_2)$
Set to unitary columns $\tilde{\Psi} \leftarrow \Psi \mathbf{W}$.
Define \mathcal{I} as the set of all columns of $\tilde{\Psi}$.

while $\|\mathbf{r}_k\|_2 > \delta$ and $k < k_{max}$ and $\text{card}(\mathcal{I}_k) < \rho M$ **do**
 $c_k^i = |\psi_i^T \mathbf{r}_{k-1}|$ where $i \in \mathcal{I} \setminus \mathcal{I}_{k-1}$
 $c_k^{i,Max} = \max_i \{c_k^i : i \in \mathcal{I} \setminus \mathcal{I}_{k-1}\}$
 $S_k = \{i : c_k^i \geq \gamma c_k^{i,Max}\}$
for all $i \in S_k$ **do**
 $F_i \leftarrow \text{Anc}(i)$.
Assemble $\tilde{\Psi}_k$ with $\mathcal{I}_{k-1} \cup F_i$.
 $\alpha_i = (\tilde{\Psi}_k^T \tilde{\Psi}_k)^{-1} \tilde{\Psi}_k^T \mathbf{u}$
 $r(i) = \|\mathbf{u} - \tilde{\Psi}_k \alpha_i\|_2$
end for
 $i_k = \arg \min_{i \in S_k} r(i)$
 $\mathcal{I}_k = \mathcal{I}_{k-1} \cup F_{i_k}$.
 $\mathbf{r}_k = \mathbf{r}_{i_k}$.
end while
return:
 $\mathbf{W}^{-1} \alpha_k$

fied distribution, which promotes the important regions of the input variables and the quantity of interest whose expectation is sought.

An insight into the typical wavelet structure of piecewise smooth functions is given in Duarte *et al.* (2005). In particular, the wavelet coefficients of piecewise smooth functions tend to form connected subtrees within wavelet trees. Additionally, a large wavelet coefficient (in magnitude) generally indicates the presence of a local singularity or sharp gradient. The above considerations form the basis of our sampling strategy. The idea is to concentrate samples at locations where large multiwavelet coefficients are observed while preconditioning the basis to maintain orthogonality. The proposed importance sampling consists of a number of steps that are applied iteratively:

(a) A multiwavelet approximation up to a given m and resolution j is obtained by solving (P_0) .

(b) The coefficients α_i are sorted in decreasing order, based on $|\alpha_i|/|\text{supp}(\psi_i)|$, where $|\text{supp}(\psi_i)|$ is the size of the support of ψ_i .

(c) A sample is drawn in $\text{supp}(\alpha_i)$ according to a uniform distribution only if $|\alpha_i| > \alpha_{tol}$ ($\alpha_{tol} = 1.0 \times 10^{-3}$ is used in the present study).

Preconditioning the basis. Assuming y is uniformly distributed on $[0, 1]$, i.e., $\rho(y) : [0, 1] \rightarrow 1$, the direct application of the above modified sampling leads to measurement matrices Ψ with large mutual coherence $\mu(\Psi)$. This is because the multiwavelets are orthogonal with respect to the measure $\rho(y)$, i.e., $\int_0^1 \psi_i(y) \psi_j(y) dy = \delta_{ij}$. A correction, therefore, is needed to retain orthogonality for sufficiently large M . Let $\gamma(y) : [0, 1] \rightarrow \mathbb{R}_{>0}$ denote the density function according to which the (independent) modified samples $y^{(k)}$, $k = 1, \dots, M$, are distributed and let $\hat{\psi}_i(y) = \psi_i(y)/\sqrt{\gamma(y)}$ be the scaled multiwavelet basis. Then,

$$\frac{1}{M} \sum_{k=1}^M \hat{\psi}_i(y^{(k)}) \hat{\psi}_j(y^{(k)}) \xrightarrow{\text{a.s.}} \int_0^1 \frac{\psi_i(y)}{\sqrt{\gamma(y)}} \frac{\psi_j(y)}{\sqrt{\gamma(y)}} \gamma(y) dy = \delta_{ij}, \quad (4.1)$$

as a result of the strong law of large numbers.

In the CS framework, this translates in sampling according to $\gamma(y)$ and using a modified measurement matrix $\hat{\Psi} = \mathbf{W}\Psi$ and the data $\hat{\mathbf{u}} = \mathbf{W}\mathbf{u}$ with the preconditioner matrix $\mathbf{W} = \text{diag}(1/\sqrt{\gamma(y^{(i)})})$, $i = 1, \dots, M$.

We now discuss how a piecewise constant measure $\gamma(y)$ can be defined on partitions of $[0, 1]$ associated with a (truncated) multiwavelet representation. We focus on establishing a one-to-one relationship between a scalar wavelet tree and a partition of $[0, 1]$. Vector trees (whose vertices are arrays of numbers) are used to store multiwavelet representations whereas scalar trees are usually adopted for wavelets. A partition of $[0, 1]$ is built by first forming a scalar connected subtree \mathcal{T} , obtained by pruning all vertices with coefficients α_i with $|\alpha_i| < \alpha_{tol}$. The leaves (L in total) of \mathcal{T} are identified, and their supports form a set of disjoint intervals $\{\mathcal{B}_i : i = 1, \dots, L\}$ which result in the desired partition of $[0, 1]$.

Using the coefficient-driven sampling discussed above, $\gamma(y)$ is defined as a piecewise constant distribution. In particular, a set of probability masses $p_i = 2^{-j}(M_i/M)$ for every box \mathcal{B}_i is considered in which M_i is the number of samples within the interval \mathcal{B}_i (with $|\mathcal{B}_i| = 2^{-j}$), j is the resolution level of the associated leaf, and M is the total number of available samples. A graphical representation of the adaptive Importance Sampling measure $\gamma(y)$ is illustrated in Figure 1 for the Kraichnan-Orszag problem with one random initial condition. It can be seen how the measure evolves with increasing samples from a uniform distribution to having a major peak close to $y = 1/2$, where finer scales are observed in the stochastic response.

5. Numerical test: Kraichnan-Orszag problem

The Kraichnan-Orszag (KO) problem is derived from simplified inviscid Navier-Stokes equations (Kraichnan 1963), and is expressed as a coupled system of non-linear ODEs. We here adopt a rotated version of the original KO problem

$$\frac{du_1}{dt} = u_1 u_3, \quad \frac{du_2}{dt} = -u_2 u_3, \quad \frac{du_3}{dt} = -u_1^2 + u_2^2, \quad (5.1)$$

with initial conditions specified below.

In Wan & Karniadakis (2005), the KO problem is used as a benchmark and analytical solutions are provided in terms of Jacobi's elliptic functions. If the set of initial conditions is chosen such that the bifurcation point $(u_1, u_2, u_3) = (\sqrt{2}, 0, 1)$ is consistently crossed,

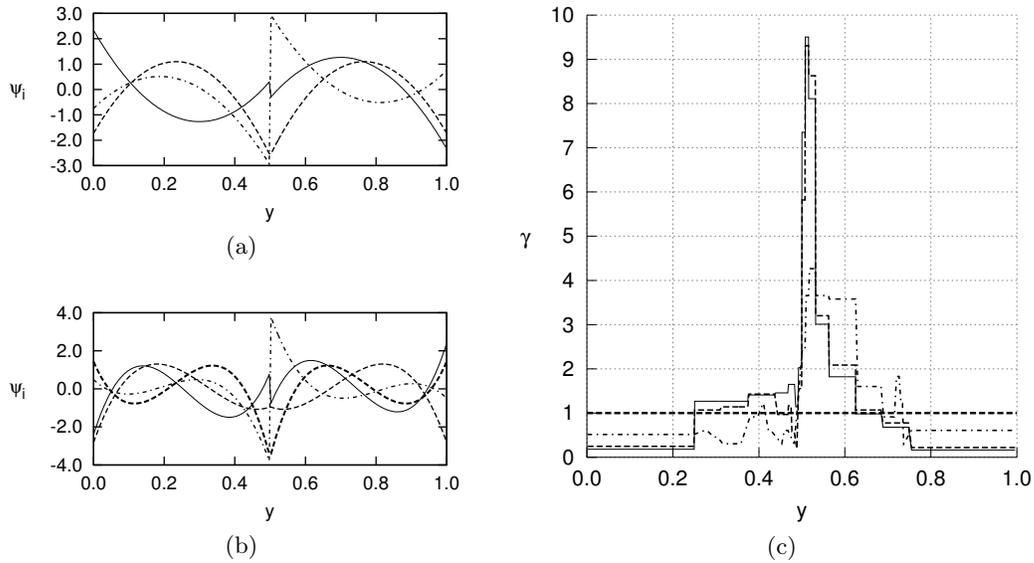


Figure 1: Examples of Alpert Multiwavelets with 3 (a) and 4 (b) vanishing moments, respectively. Adaptivity of the Importance Sampling measure with increasing number of samples (c) for the one-dimensional KO problem.

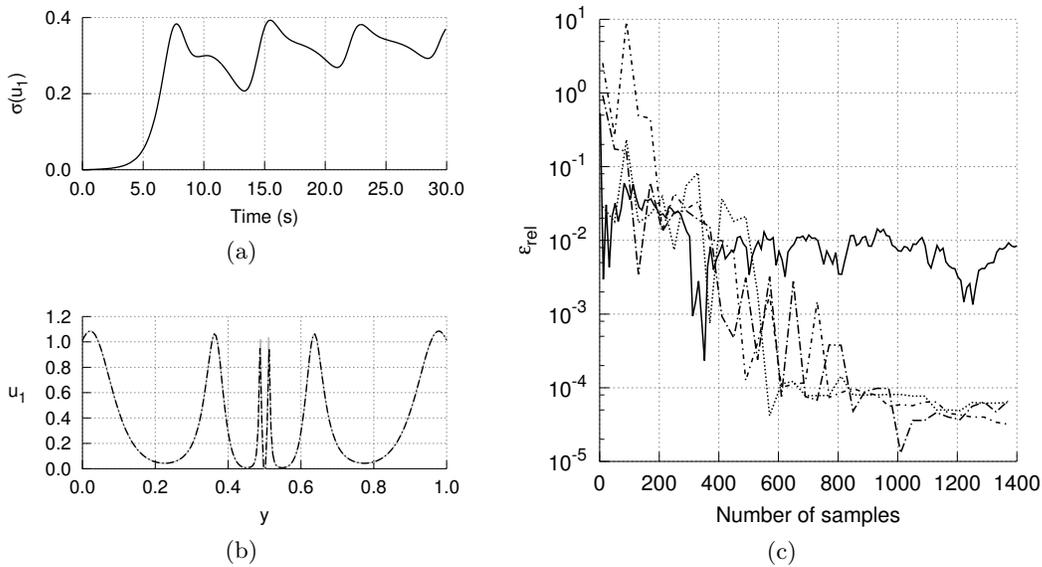


Figure 2: Results for the 1D KO problem. Evolution in time for the standard deviation of u_1 (a). Parametric response for u_1 at $t = 30s$ as obtained using 10^6 Monte Carlo realizations (—) and approximated with the proposed (CS-MW) approach (---)(b). The convergence of σ_{u_1} at $t = 30s$ using the Monte Carlo (MC) method (—) is compared to various independent runs of the proposed strategy (---,, -.-.-)(c).

it is shown that the accuracy of the global polynomial approximations (at the stochastic level) deteriorates rapidly with time.

5.1. Results for 1D KO problem

Initial conditions for (Eq. 5.1) are assumed to be uncertain and specified as

$$u_1(t=0) = 1; \quad u_2(t=0) = 0.2y - 0.1; \quad u_3(t=0) = 0, \quad (5.2)$$

where y is uniformly distributed on $[0, 1]$. The stochastic response is evaluated at $t = 20s$ and $t = 30s$ using a multiwavelet dictionary with $m = 3$ and a resolution up to $j = 7$. The OMP solver was used with a relative tolerance $\epsilon = 1.0 \times 10^{-4}$. The time history of the standard deviation for variable u_1 together with the reconstructed response at $t = 30s$ and convergence graphs are illustrated in Figure 2. The error metric $\epsilon_{rel} = |\sigma_{CSMW} - \hat{\sigma}|/\hat{\sigma}$ is also evaluated where σ_{CSMW} is the estimate for the standard deviation calculated with the CS-based multiwavelet expansion and $\hat{\sigma}$ the corresponding exact value, resulting from several millions of Monte Carlo simulations.

5.2. Results for the 2D KO problem at $t = 10s$

The initial conditions of the Kraichnan-Orszag problem are again assumed to be uncertain but this time are functions of two random variables

$$u_1(t=0) = 1; \quad u_2(t=0) = 0.2y_1 - 0.1; \quad u_3(t=0) = 2y_2 - 1, \quad (5.3)$$

where y_1 and y_2 are independent and uniformly distributed on $[0, 1]$. A two-dimensional multiwavelet measurement matrix is generated with $m = 2$ and resolution up to $j = 4$, resulting in a basis of cardinality $P = 9216$. Figure 3 shows results in terms of sampling distribution, multiwavelet coefficients, and convergence in statistics for the system's response. As noted from Figure 3, the expansion coefficients produced by the proposed strategy with about $M = 2400$ samples are comparable to those obtained using a multiwavelet least-squares approximation (where coefficients are evaluated as $\alpha_{LS} = (\Psi^T \Psi)^{-1} \Psi^T \mathbf{u}$ with 9×10^4 samples, demonstrating the efficiency of the CS-based reconstruction.

6. Conclusions

A novel framework for non-intrusive uncertainty propagation is proposed which combines the ability of a multiresolution approximation in capturing piecewise smooth stochastic responses via Compressive Sampling. Within this framework, an adaptive Importance Sampling strategy is applied in order to increase the efficiency in approximating responses with sharp gradients or discontinuities, for a given number of samples. Finally, the convergence of the method is demonstrated through its application to the Kraichnan-Orszag problem – a non-linear ODE system – with random initial conditions.

The proposed approach is particularly efficient under assumptions of mild dimensionality and sparsity of the response in the selected multiresolution basis. Further research will be devoted to progressively reducing the role played by the above assumptions, by exploring new representations and more general reconstruction paradigms.

Acknowledgments

The authors would like to thank Paul Constantine for kindly reviewing the paper, providing helpful suggestions and comments. G.I.'s and D.S.'s work is supported un-

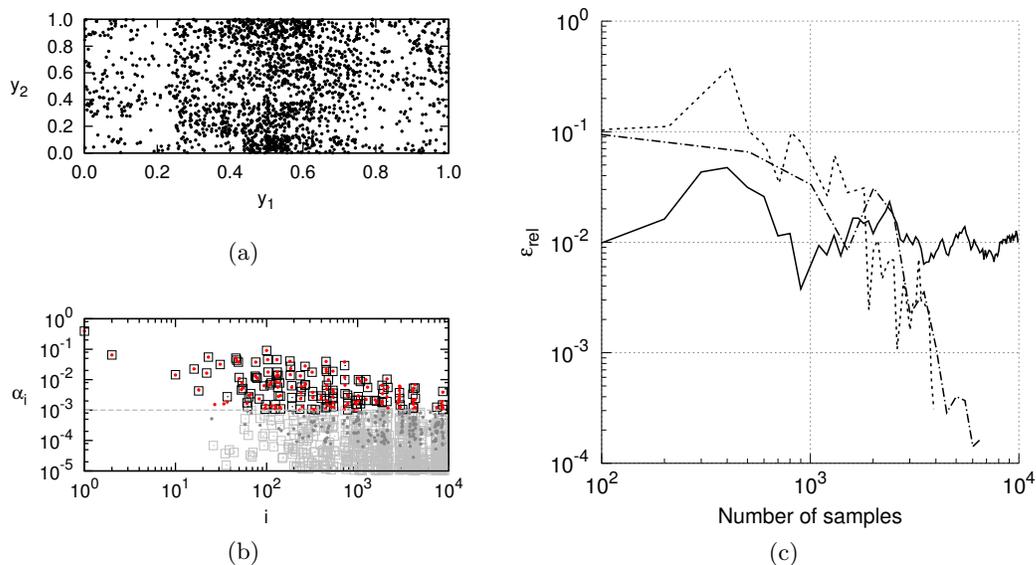


Figure 3: Results for the 2D KO problem. Distribution of 2410 adaptively sampled locations over the unitary square (a). Higher densities can be observed close to $y_1 = 1/2$, as expected. Expansion coefficients as obtained with CS-MW (\bullet) and using least-squares with 9×10^4 samples (\square) (b). Convergence profiles for the Monte Carlo (MC) method (—) and two independent runs of the proposed (CS-MW) approach (---,) (c).

der Subcontract No. B597952, with Lawrence Livermore National Security under Prime Contract No. DE-AC52-07NA27344 from the Department of Energy National Nuclear Security Administration for the management and operation of the Lawrence Livermore National Laboratory. A.D.'s work is supported by the Department of Energy under Advanced Scientific Computing Research Early Career Award DE- SC0006402.

REFERENCES

- ALPERT, B. 1993 A class of bases in l^2 for the sparse representation of integral operators. *Siam J. Math. Anal.* **24**, 246.
- BRUCKSTEIN, A., DONOHO, D. & ELAD, M. 2009 From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Review* **51** (1).
- CANDES, E. & TAO, T. 2005 Decoding by linear programming. *Information Theory, IEEE Transactions on* **51** (12), 4203–4215.
- CHANTRASMI, T., DOOSTAN, A. & IACCARINO, G. 2009 Padé–Legendre approximants for uncertainty analysis with discontinuous response surfaces. *J. Computational Physics* **228** (19), 7159–7180.
- DOOSTAN, A. & OWHADI, H. 2011 A non-adapted sparse approximation of pdes with stochastic inputs. *J. Computational Physics* **230** (8), 3015 – 3034.
- DUARTE, M., WAKIN, M. & BARANIUK, R. 2005 Fast reconstruction of piecewise smooth signals from incoherent projections. *SPARS05*.
- GHANEM, R. & SPANOS, P. 2003 *Stochastic finite elements: a spectral approach*. Dover Publ.

- KRAICHNAN, R. 1963 Direct-interaction approximation for a system of several interacting simple shear waves. *Physics of Fluids* **6**, 1603.
- LA, C. & DO, M. 2006 Tree-based orthogonal matching pursuit algorithm for signal reconstruction. In *Image Processing, 2006 IEEE International Conference on*, pp. 1277–1280. IEEE.
- LE MAITRE, O., NAJM, H., GHANEM, R. & KNIO, O. 2004 Multi-resolution analysis of wiener-type uncertainty propagation schemes. *J. Computational Physics* **197** (2), 502–531.
- MA, X. & ZABARAS, N. 2009 An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations. *Journal of Computational Physics* **228** (8), 3084–3113.
- WAN, X. & KARNIADAKIS, G. 2005 An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. *J. Computational Physics* **209** (2), 617–642.
- WITTEVEEN, J. & IACCARINO, G. 2012 Simplex stochastic collocation with random sampling and extrapolation for nonhypercube probability spaces. *SIAM J. Sci. Comput.* **34**, A814.
- XIU, D. & HESTHAVEN, J. 2006 High-order collocation methods for differential equations with random inputs. *SIAM J. Sci. Comput.* **27** (3), 1118.
- XIU, D. & KARNIADAKIS, G. 2002 The wiener–askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.* **24** (2), 619–644.