Multi-Resolution Analysis of Wiener-Type Uncertainty Propagation Schemes ⋆

O.P. Le Maître [a], H.N. Najm [b], R.G. Ghanem [c], O.M. Knio [d,*]

[a] *Université d'Evry Val d'Essonne, CEMIF and Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur, LIMSI-CNRS, BP 133, F-91 403 Orsay, France*

[b] *Combustion Research Facility, Sandia National Laboratories, Livermore, CA 94550, USA*

[c] *Department of Civil Engineering, The Johns Hopkins University, Baltimore, MD 21218, USA*

[d] *Department of Mechanical Engineering, The Johns Hopkins University, Baltimore, MD 21218, USA*

Abstract

A Multi-Resolution Analysis (MRA) is applied to an uncertainty propagation scheme based on a generalized Polynomial Chaos (PC) representation. The MRA relies on an orthogonal projection of uncertain data and solution variables onto a multi-wavelet basis, consisting of compact piecewise-smooth polynomial functions. The coefficients of the expansion are computed through Galerkin procedure. The MRA scheme is applied to the simulation of the Lorenz system having a single random parameter. The convergence of the solution with respect to the resolution level and expansion order is investigated. In particular, results are compared to two Monte Carlo sampling strategies, demonstrating the superiority of the MRA. For more complex problems, however, the MRA approach may require excessive CPU times. Adaptive methods are consequently developed in order to overcome this drawback. Two approaches are explored: the first is based on adaptive refinement of the multi-wavelet basis, while the second is based on adaptive block-partitioning of the space of random variables. Computational tests indicate that the latter approach is better suited for large problems, leading to a more efficient, flexible and parallelizable scheme.

Key words: Polynomial Chaos, multi-resolution analysis, multi-wavelets, uncertainty quantification, adaptive scheme

1 Introduction

Spectral representation schemes based on the Polynomial Chaos (PC) decomposition [5, 30] provide an efficient alternative to Monte Carlo methods [6] for uncertainty propagation and quantification. Following the work of Ghanem and Spanos [14], this approach has been applied to different fields of engineering sciences. Recently, the development of complex physical and numerical models and the availability of large-scale parallel platforms have increased the need for efficient uncertainty quantification schemes to be used for validation, design, and desicion making analysis. For instance, spectral representations have been applied to the analysis of stochastic elastic structures [14, 23]; to flow through porous media [12, 13], the Navier-Stokes equations [18, 20], thermal engineering [16], combustion and reacting flows [26], and electrochemical microfluidic applications [10]. Although successful for solving many problems, the PC representations have some limitations in particular when the solutions are complex [7].

One specific limitation, considered in a previous effort [19], appears when the solutions are smooth, but exhibit sharp variation or even discontinuity with respect to random data. Such a situation occurs for instance when one or more of the random parameters is in the neighborhood of a critical point. In [19], it was shown that the classical Wiener-Hermite (WHe) spectral expansions, or generalized forms based on global smooth polynomials such as Wiener-Legendre (WLe), may fail to describe steep or discontinuous dependence of the solution on a random parameter. Specifically, aliasing errors and Gibbs-type phenomena make such expansions impractical. In [19], a wavelet-based PC expansion was developed in order to overcome this limitation. The expansion, which combined concepts of generalized PC expansions [31–33] and of using of piecewise functions in stochastic Galerkin methods [9], relied on the decomposition of the random data and solution process into Haar wavelets (see *e.g.* [8, 25, 28, 29]), allowing for localized decomposition and resulting in a more robust method. Numerical tests conducted on problems involving critical

⋆ This work was supported by the Laboratory Directed Research and Development Program at Sandia National Laboratories, funded by the US Department of Energy. Support was also provided by the Defense Advanced Research Projects Agency (DARPA) and Air Force Research Laboratory, Air Force Materiel Command, USAF, under agreement number F30602-00-2-0612. The U.S. government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Computations were performed at the National Center for Supercomputer Applications.

* Corresponding author. Tel: 1-410-516-7736; fax: 1-410-516-7254.
Email addresses: olm@cemif.univ-evry.fr (O.P. Le Maître), hnnajm@ca.sandia.gov (H.N. Najm), ghanem@jhu.edu (R.G. Ghanem), knio@jhu.edu (O.M. Knio).

random parameters demonstrated the robustness of the Wiener-Haar (WHa) scheme, but showed that for smooth problems the latter exhibits a lower rate of convergence than that observed for spectral approximations. The present work generalizes the methodology introduced in [19], to arbitrary polynomial order expansions according to the framework proposed by Alpert in [3], the WHa expansion corresponding to the zero order case. The goal of the present generalization is to explore the possibility of combining the advantages of a higher order convergence rate resulting from higher-order polynomial expansion (p-convergence), and of the robustness of local decompositions. To this end, a multi-resolution analysis scheme is constructed, allowing for refinement of the expansion by increasing the number of resolution levels (h-refinement) and/or the polynomial order (p-refinement). The present concept of $h-p$ refinement resembles but differs that introduced in [4], where a stochastic Galerkin method combining spatial (h) refinement with order (p) refinement in the space of random data. In the present case, finite difference methods are used for spatial discretization, and both the h- and p-refinements characterize the multi-wavelet discretization of the space of random data.

The paper is organized as follows. In section 2 a brief outline is provided of the PC representation. The construction of the MRA scheme is then described in section 3. In section 4, the resulting scheme is tested based on application to the Lorentz system having a single random parameter. In particular, the convergence of the solution is analyzed with respect to the number of refinement levels and the expansion order, and the convergence rate of the MRA scheme is also contrasted with that of two Monte-Carlo (MC) sampling strategies. The results underscore the importance of seeking the right balance between accuracy, computational cost, and complexity. They also motivate the development of adaptive strategies that may be ideally suited for larger problems. Initial steps in the latter direction are taken in section 5, which describes the construction of an adaptive multi-wavelet expansion scheme and its application to a critical Rayleigh-Bénard problem [19]. An alternative approach is considered in section 5.2, where a block-partioning strategy is developed and tested using a surface reaction model [22] with multiple random parameters. Major conclusions are summarized in section 6.

2 Spectral Stochastic Representation

We consider the following generic equation governing a stochastic process $\mathcal{P}(\boldsymbol{\theta})$:

$$\mathcal{O}(\mathcal{P}(\boldsymbol{\theta}), \boldsymbol{\theta}) = 0, \tag{1}$$

where $\mathcal{O}$ is an non-linear operator and $\boldsymbol{\theta}$ is random coefficient vector. The spectral stochastic representation considered here relies on an orthogonal Fourier-like decomposition of the solution process according to:

$$\mathcal{P}(\boldsymbol{\theta}) = \sum_i \mathcal{P}_i \Psi_i(\boldsymbol{\xi}(\boldsymbol{\theta})), \tag{2}$$

where the $\mathcal{P}_i$'s are (yet-to-be-determined) spectral coefficients, and the $\{\Psi_i, i = 0, \ldots, \infty\}$ denote the basis functions. The expansion (2) is an orthogonal basis function representation of $\mathcal{P}$. This approach has lead to the concepts of Homogeneous Chaos and Polynomial Chaos (PC) expansions [5, 14, 30] when $\boldsymbol{\xi}$ is a Gaussian vector with independent components. It can also be extended to non-Gaussian measures [11, 31]. To simplify the notation, we shall hereafter drop the dependence on $\boldsymbol{\theta}$, and write $\mathcal{P}(\boldsymbol{\xi})$ instead of $\mathcal{P}(\boldsymbol{\theta})$ or $\mathcal{P}(\boldsymbol{\xi}(\boldsymbol{\theta}))$ Due to the orthogonality of the basis functions, that is $< \Psi_i \Psi_j >= 0$ for $i \neq j$, the spectral coefficients satisfy:

$$\mathcal{P}_i = \frac{\langle \mathcal{P} \Psi_i \rangle}{\langle \Psi_i \Psi_i \rangle}. \tag{3}$$

In order to determine the solution $\mathcal{P}$ it is sufficient to determine the spectral coefficients, $\mathcal{P}_i$.

Several approaches can be used to determine the $\mathcal{P}_i$'s. A simple approach is to use Monte Carlo sampling (e.g. [26]) or alternatively rely on quadrature formulae (e.g. [20]). In both cases, the problem is reduced to repeated solutions of a deterministic problem corresponding to different realizations of $\boldsymbol{\xi}$. The ensemble of realizations is then used to reconstruct the spectral coefficients. Thus, we refer to this approach as a non-intrusive spectral projection (NISP) because it enables immediate application of a deterministic solution scheme.

An alternative to MC and quadrature schemes is the Galerkin approach. In the latter case, the expansion (2) is substituted into the governing equation (1) and orthogonal projections onto the basis functions are formed, resulting in the following system for the basis function coefficients:

$$\left\langle \mathcal{O}\left(\sum_i \mathcal{P}_i \Psi_i(\boldsymbol{\xi}), \boldsymbol{\xi} \right), \Psi_j \right\rangle = 0, \quad \forall j. \tag{4}$$

Below, we shall focus exclusively on combining the Galerkin approach with a multi-wavelet (MW) basis function expansion.

3 Multi-Resolution Analysis and Multi-Wavelet Basis

As mentioned earlier, if $\mathcal{P}(\boldsymbol{\xi})$ exhibits a fast or discontinuous dependence on the random data, a large expansion order No may be necessary when using an orthogonal basis of smooth, global polynomials. In many situations, this may render such classical expansions impractical. In this section, an alternative representation of $\mathcal{P}(\boldsymbol{\xi})$ is introduced based on a multi-wavelet expansion. To simplify the presentation, we shall first focus on the case of a single uncertain parameter, ξ; generalization to the multi-dimensional case is then addressed in subsection 3.4.

3.1 Change of Variable

Let $\mathrm{p}(\eta)$ denote the Cumulative Distribution Function, giving the probability that $\xi \leq \eta$. As in [19], we assume that $\mathrm{p}(\eta)$ is a continuous, monotically increasing function η over the interval (a, b), $-\infty \leq a < b \leq \infty$, and that $\mathrm{p}(a) = 0$ and $\mathrm{p}(b) = 1$. Based on the assumed properties of $\mathrm{p}(\eta)$, it follows that for all $r \in [0, 1]$ there is a unique $\eta \in [a, b]$ such that $\mathrm{p}(\eta) = r$. In addition, if x is a uniformly distributed random variable on $[0, 1]$, then $p^{-1}(x)$ is a random variable on (a, b) having the same distribution as ξ [15]. Consequently, instead of expanding the random process in terms of ξ, we alternatively develop a representation with respect to x.

3.2 Multi-Resolution Analysis

In this section, we recall some properties of the multi-wavelet bases introduced by Alpert in [3] (see also [2]). The application to the representation of random processes is considered in subsection 3.3.

3.2.1 Vector spaces

For $No = 0, 1, \ldots$ and $k = 0, 1, 2, \ldots$, we define the space $\mathbf{V}_k^{No}$ of piecewise-continuous polynomials, according to:

$$\begin{aligned} \mathbf{V}_k^{No} \equiv \Big\{ f : \ & \text{the restriction of } f \text{ to the interval } (2^{-k}l, 2^{-k}(l+1)) \\ & \text{is a polynomial of degree} \leq No, \text{ for } l = 0, \ldots, 2^k - 1 \qquad (5) \\ & \text{and } f \text{ vanishes outside the interval [0,1]} \Big\}. \qquad (6) \end{aligned}$$

Thus, $\mathbf{V}_k^{No}$ has dimension $(No+1)(2^k)$ and $\mathbf{V}_0^{No} \subset \mathbf{V}_1^{No} \subset \ldots \subset \mathbf{V}_k^{No} \subset \ldots$. Denoting $\mathbf{V}^{No}$ the union all of spaces $\mathbf{V}_k^{No}$, $\mathbf{V}^{No} = \overline{\bigcup_{k\geq 0} \mathbf{V}_k^{No}}$, [3], we remark that $\mathbf{V}^{No}$ is dense in $L^2([0,1])$ with respect to the norm $\|f\| = \langle f, f\rangle^{1/2}$ where

$$\langle f, g\rangle = \int_0^1 f(x)g(x)dx. \tag{7}$$

The Multi-Wavelets (MW) subspace $\mathbf{W}_k^{No}$, $k = 0, 1, 2, \ldots$ is defined as the orthogonal complement of $\mathbf{V}_k^{No}$ in $\mathbf{V}_{k+1}^{No}$; we write:

$$\mathbf{V}_k^{No} \oplus \mathbf{W}_k^{No} = \mathbf{V}_{k+1}^{No}, \quad \mathbf{W}_k^{No} \perp \mathbf{V}_k^{No}. \tag{8}$$

From this construction, we have:

$$\mathbf{V}_0^{No} \bigoplus_{k\geq 0} \mathbf{W}_k^{No} = L^2([0,1]). \tag{9}$$

3.2.2 Multi-wavelet basis

An orthonormal basis, $\{\psi_0, \psi_1, \ldots, \psi_{No}\}$, of $\mathbf{W}_0^{No}$ is introduced. The ψ_i's are piecewise polynomial functions of degree less than or equal to No. From the orthonormality condition, we have:

$$\langle \psi_i(x), \psi_j(x)\rangle = \delta_{ij}. \tag{10}$$

Since $\mathbf{W}_0^{No} \perp \mathbf{V}_0^{No}$, the first $No+1$ moments of the ψ_i vanish, i.e.

$$\left\langle \psi_j, x^i \right\rangle = 0, \quad 0 \leq i, j \leq No. \tag{11}$$

Equations (10) and (11) result in a system of polynomial equations which is solved, yielding the $(No+1)$ ψ_i functions (see Appendix A).

The space $\mathbf{W}_k^{No}$, whose dimension is $(No+1)(2^k)$, is spanned by the multi-wavelets, ψ_{jl}^k, which are translated and dilated version of the ψ_i's. The ψ_{jl}^k are given by:

$$\psi_{jl}^k(x) = 2^{k/2}\psi_j(2^k x - l), \quad j = 0, \ldots, No, \text{ and } l = 0, \ldots, 2^k - 1. \tag{12}$$

and their support is $\text{Supp}(\psi_{jl}^k) = [2^{-k}l, 2^{-k}(l+1)]$. Due to the orthonormality of the ψ_i's, we have:

$$\left\langle \psi_{il}^k, \psi_{jm}^{k'} \right\rangle = \delta_{ij}\delta_{lm}\delta_{kk'}. \tag{13}$$

A basis $\{\phi_0, \ldots, \phi_{k-1}\}$ for $\mathbf{V}_0^k$ is also constructed. Rescaled Legendre polynomials are used for this purpose. Letting $\mathcal{L}e_i$ denote the Legendre polynomial [1] of degree i, defined over $[-1, 1]$, we set:

$$\phi_i(x) = \frac{\mathcal{L}e_i(2x-1)}{L_i}, \quad i = 0, 1, \ldots, No, \tag{14}$$

where L_i is a normalization factor selected such that:

$$\langle \phi_i(x), \phi_j(x) \rangle = \delta_{ij} \text{ for } i, j = 0, \ldots, No. \tag{15}$$

The space $\mathbf{V}_k^{No}$, whose dimension is $2^k(No+1)$, is spanned by the polynomials ϕ_{il}^k,

$$\phi_{il}^k(x) = 2^{k/2}\phi_i(2^k x - l), \quad i = 0, \ldots, No \text{ and } l = 0, \ldots, 2^k - 1. \tag{16}$$

which are translated and dilated versions of the ϕ_i's.

3.2.3 MW expansion

A function $f(x) \in L^2([0,1])$ can be arbitrary well approximated using the MRA scheme constructed above. We denote $f^{No,Nr}$ the projection of f on $\mathbf{V}_{Nr}^{No}$; we have:

$$\begin{aligned} f^{No,Nr} \equiv \mathbb{P}_{Nr}^{No}[f] &= \sum_{l=0}^{2^{Nr}-1} \sum_{i=0}^{No} \left\langle \phi_{il}^{Nr}(x), f(x) \right\rangle \phi_{il}^{Nr}(x) \\ &= \sum_{l=0}^{2^{Nr}-1} \sum_{i=0}^{No} \overline{f}_{il}^{Nr} \phi_{il}^{Nr}(x). \end{aligned} \tag{17}$$

An alternative expression for $f^{No,Nr}$, valid for all $Nr \geq 1$, in terms of multiwavelets is:

$$f^{No,Nr}(x) \equiv \mathbb{P}_0^{No}[f(x)] + \sum_{k=0}^{Nr-1} \sum_{l=0}^{2^k-1} \left(\sum_{i=0}^{No} df_{il}^k \psi_{il}^k(x) \right). \tag{18}$$

The MW coefficients df_{il}^k appearing in Eq. (18) are given by:

$$df_{il}^k = \left\langle \left\{ \mathbb{P}_{k+1}^{No}[f] - \mathbb{P}_k^{No}[f] \right\}, \psi_{il}^k \right\rangle \tag{19}$$

Denoting by $\delta_{No,Nr}$ the L^2-norm of the approximation error of f on $\mathbf{V}_{Nr}^{No}$, $\delta_{No,Nr} \equiv \left\langle f - f^{No,Nr}, f - f^{No,Nr} \right\rangle$, convergence is characterized by the decay

of δ with increasing polynomial order No (p convergence) or with increasing resolution levels Nr (h convergence).

3.3 Expansion of the random process

Consider a random process $\mathcal{P}(\xi(\theta))$, where $\xi(\theta)$ is a random variable satisfying the assumption of section 3.1. Further, we assume that $\mathcal{P}(\xi)$ is a second order process, i.e.

$$\int_a^b \mathcal{P}(\xi)^2 \mathrm{pdf}(\xi) d\xi < \infty. \tag{20}$$

Using the change of variables introduced in section 3.1, we express $\mathcal{P}(\xi)$ in terms of $x = \mathrm{p}(\xi)$ as:

$$\mathcal{P}(\xi) = \mathcal{P}(\mathrm{p}^{-1}(x)) = \tilde{\mathcal{P}}(x). \tag{21}$$

Introducing this change of variable in Eq. (20) gives

$$\int_a^b \mathcal{P}(\xi)^2 \mathrm{pdf}(\xi) d\xi = \int_0^1 \tilde{\mathcal{P}}(x)^2 dx = \left\langle \tilde{\mathcal{P}}, \tilde{\mathcal{P}} \right\rangle < \infty, \tag{22}$$

showing that $\tilde{\mathcal{P}}(x) \in L^2([0,1])$. Thus, $\tilde{\mathcal{P}}(x)$ can be expanded according to Eq. (18).

Let us denote by $\triangle$ the set of index integers λ concatenating the scale index k, support index l and MW index i:

$$\begin{aligned} \triangle \equiv \{\lambda : \lambda = (No+1)(2^k + l - 1) + i; k = 0, \ldots, \infty; \\ l = 0, \ldots, 2^k - 1; i = 0, No\}. \end{aligned} \tag{23}$$

The resolution level k for any $\lambda \in \triangle$ will be denoted by $|\lambda|$. Using this convention, the MW expansion of $\tilde{\mathcal{P}}(x)$ can be expressed as:

$$\tilde{\mathcal{P}}(x) = \mathbb{P}_0^{No}[\tilde{\mathcal{P}}] + \sum_{\lambda \in \triangle} \tilde{\mathcal{P}}_\lambda \psi_\lambda(x) = \sum_{i=0}^{No} \overline{\tilde{\mathcal{P}}}^0_{i0} \phi_i(x) + \sum_{\lambda \in \triangle} \tilde{\mathcal{P}}_\lambda \psi_\lambda(x). \tag{24}$$

Letting $\triangle_0 = \{-No-1, -No, \ldots, -1\}$ and $\diamondsuit = \triangle \cup \triangle_0$, we can rewrite Eq. (24) as:

$$\tilde{\mathcal{P}}(x) = \sum_{\lambda \in \diamondsuit} \tilde{\mathcal{P}}_\lambda W_\lambda(x), \tag{25}$$

where

$$\begin{cases} W_\lambda(x) = \psi_\lambda(x) & \text{for } \lambda \in \triangle, \\ \\ W_\lambda(x) = \phi_{1-\lambda}(x) & \text{for } \lambda \in \triangle_0. \end{cases} \tag{26}$$

Thus, the process can be expanded as:

$$\mathcal{P}(\xi) = \tilde{\mathcal{P}}(x) = \sum_{\lambda \in \diamondsuit} \tilde{\mathcal{P}}_\lambda W_\lambda(x).$$

3.4 The Multi-dimensional case

Extension of the 1D MW expansion to the N-dimensional case is now considered. For simplicity, we focus on a vector $\boldsymbol{\xi}$ with random components $\{\xi_1(\theta), \ldots, \xi_N(\theta)\}$. We assume that the components of $\boldsymbol{\xi}$ are uncorrelated and independent, so that:

$$\text{pdf}(\boldsymbol{\xi}) = \prod_{d=0}^{N} \text{pdf}_d(\xi_d).$$

where pdf_d denotes the probability density function of the ξ_d. Consistent with the 1D case, we also assume that:

$$\text{pdf}_d(\xi_d) = \begin{cases} > 0 & \text{if } \xi \in (a_d, b_d) \\ 0 & \text{if } \xi \notin (a_d, b_d) \end{cases}, \tag{27}$$

so that $\forall x_d \in [0,1]$ there is a unique $\xi_d \in (a_d, b_d)$ such that

$$\text{p}_d(\xi_d) \equiv \int_{a_d}^{\xi_d} \text{pdf}_d(\xi') d\xi' = x_d$$

We now consider the multi-index $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)$, and define the set

$$\Lambda_k = \left\{ \boldsymbol{\lambda} : \sum_{d=1}^{N} |\lambda_d| = k \right\} \tag{28}$$

Let $M_k = \mathrm{Card}(\Lambda_k)$, and define the set

$$\mathcal{W}_k \equiv \left\{ \prod_{d=1}^{N} W_{\lambda_d}(x_d) \ : \boldsymbol{\lambda} \equiv (\lambda_1, \ldots, \lambda_N) \in \Lambda_k \right\} \tag{29}$$

of multidimensional multi-wavelets having resolution level k. The MW expansion of $\tilde{\mathcal{P}}(\boldsymbol{x})$ can now be formally written as:

$$\begin{aligned} \tilde{\mathcal{P}}(\boldsymbol{x}) = & \sum_{i=1}^{M_0} c_i^0 \Gamma_i^0(x_1, \ldots, x_N) \\ & + \sum_{i=1}^{M_1} c_i^1 \Gamma_i^1(x_1, \ldots, x_N) \\ & + \sum_{i=1}^{M_2} c_i^2 \Gamma_i^2(x_1, \ldots, x_N) \\ & + \ldots \end{aligned} \tag{30}$$

where $\Gamma^k(\boldsymbol{x}) \in \mathcal{W}_k$ denotes a multidimensional wavelet of resolution k. In practice, the MW expansion has to be truncated. Here, we choose to retain all multi-indices $\boldsymbol{\lambda}$ such that $|\boldsymbol{\lambda}| \equiv \sum_{d=1}^{N} |\lambda_d| \leq Nr$, where Nr is a prescribed resolution level. After truncation, the finite expansion may be rewritten in a single index form as:

$$\tilde{\mathcal{P}}(\boldsymbol{x}) \approx \sum_{i=0}^{P} \tilde{\mathcal{P}}_i \mathcal{M}^w{}_i(x_1, \ldots, x_N), \tag{31}$$

where $P + 1 = \sum_{i=0}^{Nr} M_i$ is the dimension of the truncated basis $\{\mathcal{M}^w{}_i, i = 0, \ldots, P\}$. Further, we use the convention that the indexing is performed in such a way that the first element of the basis is $\mathcal{M}^w{}_0 = 1$.

3.4.1 Mean and variance

The mean of $\mathcal{P}(\boldsymbol{\xi})$ is, by definition, given by:

$$\begin{aligned} \langle \mathcal{P} \rangle &= \int_{a_1}^{b_1} d\xi_1 \ldots \int_{a_N}^{b_N} d\xi_N \left[\mathcal{P}(\boldsymbol{\xi}) \prod_{d=1}^{N} \mathrm{pdf}_d(\xi_d) \right] \\ &= \int_0^1 dx_1 \ldots \int_0^1 dx_N \left[\tilde{\mathcal{P}}(x_1, \ldots, x_N) \right]. \end{aligned} \tag{32}$$

Introducing the MW expansion (31) in the above one immediately obtains:

$$\langle \mathcal{P} \rangle = \tilde{\mathcal{P}}_0.$$

Expression for the variance of $\mathcal{P}$ may be obtained in an analogous fashion. By definition, we have:

$$\begin{aligned} \sigma^2(\mathcal{P}) &= \int_{a_1}^{b_1} d\xi_1 \ldots \int_{a_N}^{b_N} d\xi_N \left[\mathcal{P}(\xi) - \langle \mathcal{P} \rangle\right]^2 \prod_{i=1}^{N} \mathrm{pdf}(\xi_i) \\ &= \int_0^1 dx_1 \ldots \int_0^1 dx_N \left[\tilde{\mathcal{P}}(x_1, \ldots, x_N) - \tilde{\mathcal{P}}_0\right]^2. \end{aligned} \tag{33}$$

Introducing the expansion (31) in the previous expression, and taking into account the orthonormality of the basis functions, we get:

$$\sigma^2(\mathcal{P}) = \sum_{i=1}^{P} \tilde{\mathcal{P}}_i^2.$$

4 Application to Lorenz system

The MRA scheme above is first applied to the stochastic Lorenz system:

$$\begin{cases} \dfrac{\partial X}{\partial t} = \rho(Y - X), \\ \dfrac{\partial Y}{\partial t} = \mathrm{Ra}(\theta) X - Y - XZ, \\ \dfrac{\partial Z}{\partial t} = -aZ + XY, \end{cases} \tag{34}$$

where $\rho = 10$, $a = 8/3$ and $\mathrm{Ra}(\theta)$ is a random parameter with uniform distribution over $[15, 21]$. Deterministic initial conditions are used, according to $X(t = 0) = Y(t = 0) = Z(t = 0) = 1$. For the present setting, the Lorenz system exhibits damped oscillations for all possible values of Ra, leading to an asymptotically steady solution as $t \to \infty$. However, the damping time-scale exhibits sharp dependence on Ra. Below, we shall focus on the statistics of X at time $t = 25$. At this value of t, the solution for the lower values of Ra has nearly achieved a steady state, while the solution for higher values of Ra still

exhibits large-amplitude oscillations. Such sharp variation with Ra requires a refined discretization along the random dimension in order to properly represent the local dynamics. In section 4.1 the $h-$ and $p-$ convergence of the computed expectation and standard deviation of X are investigated. Then in section 4.2, the same problem is solved using two MC sampling techniques in order to assess the efficiency of the MRA scheme.

4.1 $h-p$ convergence of the MW expansion

4.1.1 Solution method

For given expansion order and resolution level, the solution is approximated using:

$$(X,Y,Z)(t,\xi) = (\tilde{X},\tilde{Y},\tilde{Z})(t,x) \approx \sum_{\beta=0}^{P} (\tilde{X},\tilde{Y},\tilde{Z})_\beta(t)\mathcal{M}^w{}_\beta(x), \tag{35}$$

where we have suppressed, for clarity, the explicit dependence on θ. Introducing these expansions into the Lorenz equations, and using the Galerkin approach outlined in section 2, one obtains the following coupled system for the MW coefficients:

$$\begin{cases} \dfrac{\partial \tilde{X}_\beta}{\partial t} = \rho(\tilde{Y}_\beta - \tilde{X}_\beta), \\ \dfrac{\partial \tilde{Y}_\beta}{\partial t} = \widetilde{(\mathrm{Ra}X)}_\beta - \tilde{Y}_\beta - \widetilde{(XZ)}_\beta, \\ \dfrac{\partial \tilde{Z}_\beta}{\partial t} = -a\tilde{Z}_\beta + \widetilde{(XY)}_\beta, \end{cases} \tag{36}$$

which need to be solved for all $\beta = 0,\ldots,P$. The initial conditions are also obtained by application of the Galerkin approach, resulting in:

$$\begin{cases} \tilde{X}_\beta(t=0) = \tilde{Y}_\beta(t=0) = \tilde{Z}_\beta(t=0) = 1 & \text{for } \beta = 0 \\ \tilde{X}_\beta(t=0) = \tilde{Y}_\beta(t=0) = \tilde{Z}_\beta(t=0) = 0 & \text{for } \beta = 1,\ldots,P \end{cases} \tag{37}$$

The time integration of the $3 \times \mathrm{Card}(Nr, No) = 3(No+1)2^{Nr}$ unknown MW coefficients is performed using a fourth-order Runge-Kutta scheme, with a time step $\Delta t = 0.005$. This was selected based on successive refinement until the

solution became essentially independent of Δt. Note that in order to integrate the above system, one needs to evaluate the MW expansion of the products of two stochastic quantities, as in the quadratic term $\widetilde{XZ}$. To this end, we rely on an "exact" Galerkin procedure, according to:

$$\widetilde{(XZ)}_\beta = \left\langle \tilde{X}\tilde{Z}, \mathcal{M}^w{}_\beta \right\rangle = \left\langle \sum_{\lambda=0}^{P}\sum_{\gamma=0}^{P} (\tilde{X}_\lambda \tilde{Z}_\gamma)\mathcal{M}^w{}_\lambda \mathcal{M}^w{}_\gamma, \mathcal{M}^w{}_\beta \right\rangle$$
$$= \sum_{\lambda=0}^{P}\sum_{\gamma=0}^{P} \tilde{X}_\lambda \tilde{Z}_\gamma \left\langle \mathcal{M}^w{}_\lambda \mathcal{M}^w{}_\gamma, \mathcal{M}^w{}_\beta \right\rangle . \qquad (38)$$

Since the multiplication tensor $\langle \mathcal{M}^w{}_\alpha \mathcal{M}^w{}_\beta, \mathcal{M}^w{}_\gamma \rangle$ is solution independent, it is computed and stored in a pre-processing stage. As further discussed below, this tensor is sparse and thus only the non-zero entries are stored and actually used in the computations [10, 17]. This minimizes both storage and CPU requirements.

4.1.2 Convergence results

Figure 1 shows the computed values of $X(t, \mathrm{Ra})$ for $20 \leq t \leq 25$; plotted are results obtained with $Nr = 1, \ldots, 4$ and $No = 0, 4$. In all the frames, the same contour levels are used. The results illustrate the convergence of the predictions as Nr and No increase. The plots appear to indicate that convergence is faster for the lower values of Ra, where the solution has essentially decayed, than for the higher values, where X still exhibits large oscillations. This reveals an interesting feature of the MW representation (which has also been observed for WHa expansions [19]) namely that errors can be localized to regions where insufficient resolution is provided. This is in contrast to classical spectral representations, where under-resolution typically results in global breakdown of the solution.

Additional quantitative assessment of the $h - p$ convergence of the MRA scheme is obtained by analyzing the computed solution statistics at $t = 25$. Table 1 shows the computed values of $< X >$ at $t = 25$, for expansion orders $0 \leq No \leq 6$ and resolution levels $0 \leq Nr \leq 6$. The results indicate that the predictions convergence as Nr and No are increased. For $Nr = 0$, however, $p-$convergence is very slow and accuracy up to two digits is not achieved even for $No = 8$ (not shown). In particular, the results also indicate that in order to achieve accuracy up to the 3rd digit, a fifth-order expansion is needed for $Nr = 1$, while using $No = 0$ with $Nr = 6$ provides similar accuracy. Similar trends are also observed when analyzing the predictions of the standard deviation $\sigma(X)$, which are reported in Table 2.

The results above indicate that in order to improve the accuracy of the predic-

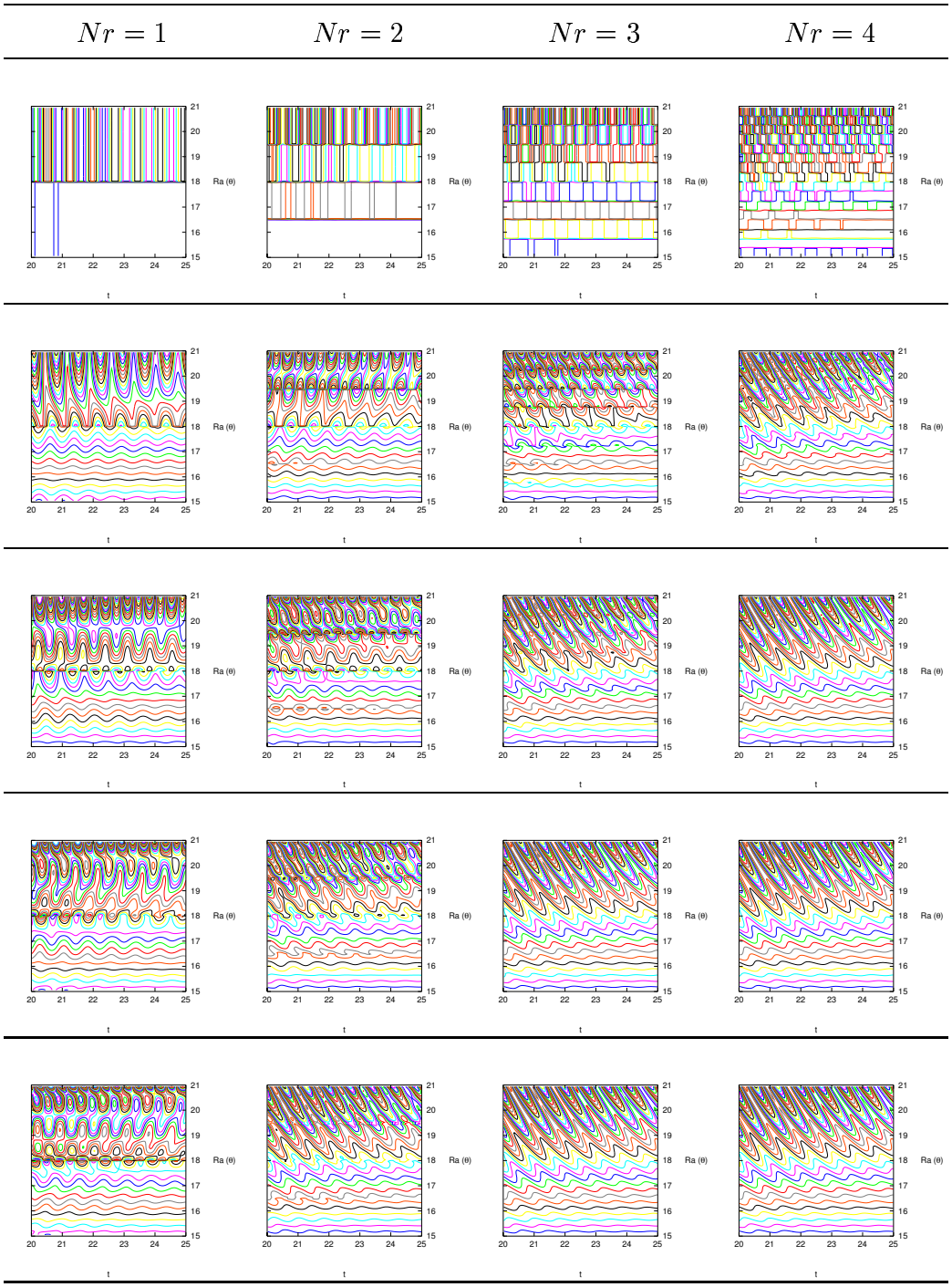

Fig. 1. Iso-lines of $X(t \in [20, 25], \mathrm{Ra})$ of the stochastic Lorenz problem, for $Nr = 1, \ldots, 4$ (left to right) and $No = 0, \ldots, 4$ (top to bottom). Time increases along the horizontal axis, while the vertical axis corresponds to $\mathrm{Ra} \in [15, 21]$. The same contour levels are used for all plots.

tions, one may either increase the number of resolution levels ($h-$refinement) or the expansion order ($p-$refinement). Since for complex problems it is usually desirable to maintain a moderate expansion order, typically $No \leq 4$ (see

$-\langle X\rangle$	$No=0$	$No=1$	$No=2$	$No=3$	$No=4$	$No=5$	$No=6$
$Nr=0$	6.739074	6.702924	3.657155	5.305756	6.844937	6.575842	6.669102
$Nr=1$	6.754146	6.742118	6.702877	6.731783	6.741929	6.730446	6.730405
$Nr=2$	6.716436	6.740495	6.733795	6.731213	6.730252	6.730173	6.730114
$Nr=3$	6.738296	6.729341	6.730230	6.730139	6.730108	6.730105	-
$Nr=4$	6.731381	6.730081	6.730106	6.730105	6.730105	-	-
$Nr=5$	6.730400	6.730104	6.730105	6.730105	-	-	-
$Nr=6$	6.730177	6.730105	6.730105	-	-	-	-

Table 1
Computed values of $\langle X\rangle$ at $t=25$ for the stochastic Lorenz system with different values of Nr and No.

$\sigma(X)$	$No=0$	$No=1$	$No=2$	$No=3$	$No=4$	$No=5$	$No=6$
$Nr=0$	0.000000	0.316249	5.660838	3.312410	0.518508	0.220683	0.287331
$Nr=1$	0.327423	0.375252	0.312143	0.381295	0.371085	0.360037	0.360215
$Nr=2$	0.329662	0.373340	0.361711	0.360578	0.357961	0.357668	0.357666
$Nr=3$	0.371072	0.355050	0.357989	0.357654	0.357668	0.357667	-
$Nr=4$	0.358382	0.357673	0.357667	0.357668	0.357667	-	-
$Nr=5$	0.357853	0.357666	0.357667	0.357667	-	-	-
$Nr=6$	0.357713	0.357667	0.357667	-	-	-	-

Table 2
Computed values of $\sigma(X)$ at $t=25$ for the stochastic Lorenz system with different values of Nr and No.

section 5), increasing Nr provides suitable means to enhance the predictions. This raises the question concerning how Nr and No should be selected so as to achieve target accuracy at the lowest possible CPU cost. For the present simple setting, this question may be addressed by relating CPU cost to discretization parameters. A rough initial estimate is the size of the ODE system, which is proportional to $\mathrm{Card}(Nr,No)=(No+1)2^{Nr}$ (Table 3). However, since the computational load is in fact dominated by the evaluation of quadratic Galerkin products, a sharper estimate is based on the number of operations actually performed in these products, which is proportional to the number of non-zero entries in the multiplication tensor $\langle \mathcal{M}^{w}{}_{\lambda}\mathcal{M}^{w}{}_{\gamma},\mathcal{M}^{w}{}_{\beta}\rangle$. This number, which is used as surrogate measure for CPU, is denoted by $\mathcal{C}(Nr,No)$ and reported in Table 4 for different values of Nr and No. Comparison of Tables 3 and 4 shows that with increasing refinement $\mathcal{C}(Nr,No)$ increases much quickly than $\mathrm{Card}(Nr,No)$.

Next, we seek a relationship between CPU and accuracy by plotting in Fig-

Card(Nr, No)	$No = 0$	$No = 1$	$No = 2$	$No = 3$	$No = 4$	$No = 5$	$No = 6$
$Nr = 0$	1	2	3	4	5	6	7
$Nr = 1$	2	4	6	8	10	12	16
$Nr = 2$	4	8	12	16	20	24	28
$Nr = 3$	8	16	24	32	40	48	56
$Nr = 4$	16	32	48	64	80	96	112
$Nr = 5$	32	64	96	128	160	-	-
$Nr = 6$	64	128	192	256	-	-	-

Table 3
Values of Card$(\overline{\diamondsuit})$ for different order and number of resolution levels.

$\mathcal{C}(Nr, No)$	$No = 0$	$No = 1$	$No = 2$	$No = 3$	$No = 4$	$No = 5$	$No = 6$
$Nr = 0$	1	4	11	29	42	69	106
$Nr = 1$	4	22	69	166	308	531	880
$Nr = 2$	16	154	611	1 538	3 084	5 457	9 034
$Nr = 3$	52	706	2 883	7 354	14 936	26 541	43 254
$Nr = 4$	148	2 578	10 625	27 434	56 040	99 561	160 018
$Nr = 5$	388	8 242	34 305	89 094	181 216	-	-
$Nr = 6$	964	24 178	101 501	262 826	-	-	-

Table 4
Values of $\mathcal{C}(Nr, No)$ for different order and number of resolution levels.

ure 2 the absolute errors in $\langle X \rangle$ and $\sigma(X)$ at $t = 25$ against $\mathcal{C}(Nr, No)$. For the purpose of estimating errors, the solution obtained using $Nr = 6$ and $No = 2$ has been used as surrogate for the exact solution (MC predictions below support this approximation). The results indicate that for the present parameter range the "errors" on the first two statistical moments of X decay roughly as $\mathcal{C}(Nr, No)^{-2}$. While the details of this relationship may be dependent on the problem and parameter range, the collapse of the data clearly indicates that in the present case the "complexity estimate" $\mathcal{C}(Nr, No)$ provides a good measure for both the CPU and the accuracy of the predictions. In addition, the results provide a convenient means for selecting Nr and No, and indicate that multiple alternatives may exist for achieving a target accuracy level.

4.2 Comparison with Monte Carlo sampling

In this section, we compare the efficiency of the MRA scheme above with that of MC sampling strategies. The latter rely on computing deterministic

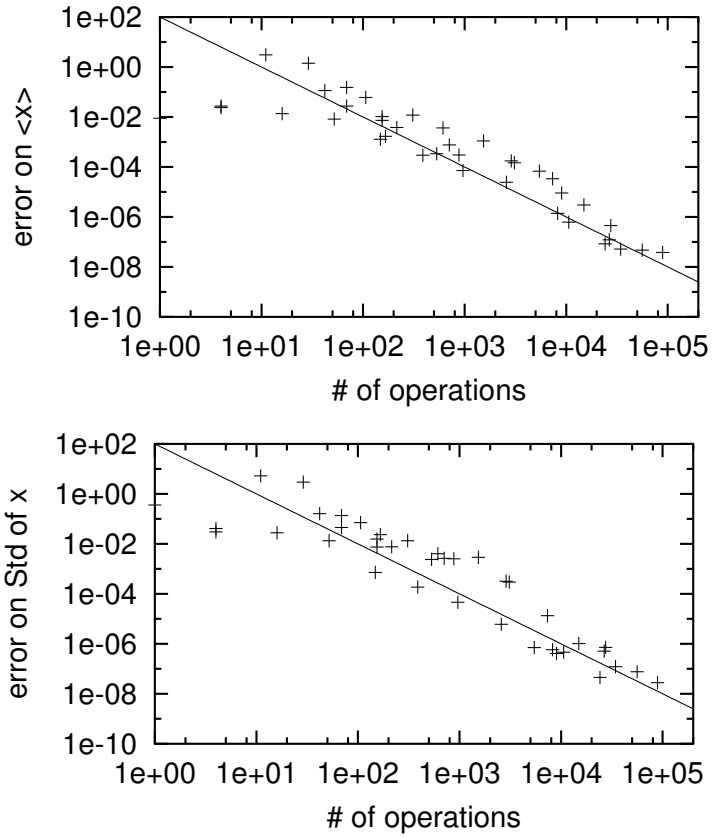

Fig. 2. Absolute errors on the expectation and standard deviation of X, at $t = 25$, as a function of the number of operations in the spectral product : Card(Nr, No). The exact solution is taken to be the computed values for $Nr = 6$ and $No = 2$, as reported in Tables 1 and 2. The solid lines are $\sim x^{-2}$.

solutions for specific values of the random parameters, and then evaluating the desired moments through a collocation approach [15, 21]. Since individual solutions are deterministic, the evaluation of the product XY of X and Y requires a single operation, so that the "complexity" of obtaining a single MC realization is taken to be 1. This provides a consistent approach for comparing the efficiency of MRA and MC. In other words, the CPU load in MRA is gauged using $\mathcal{C}(Nr, No)$ while in MC it is gauged using the number, m, of realizations.

4.2.1 Classical sampling strategy

A classical MC sampling strategy is first applied. The approach is based on using a random number generator to generate m independent realizations of the stochastic parameter, Ra_i, $i = 1, \ldots, m$, that are uniformly distributed over the interval $[15, 21]$. For each realization, the corresponding deterministic Lorenz system is integrated up to $t = 25$, resulting in particular in a set of predictions, X_i, $i = 1, \ldots, m$. The first two statistical moments of X are then

estimated using:

$$\langle X \rangle_m = \frac{1}{m} \sum_{i=1}^{m} X_i$$

and

$$\sigma_m^2(X) \approx \left\langle (X - \langle X \rangle_m)^2 \right\rangle^2 = \frac{1}{m-1} \sum_{i=1}^{m} X^2 - \langle X \rangle_m^2 .$$

In Figure 3, the absolute errors in $\langle X \rangle$ and $\sigma(X)$ plotted against the number of MC realizations, $1 \leq m \leq 10^7$. As in the analysis above, the MRA solution with $Nr = 6$, $No = 2$ is used as surrogate for the exact solution for the definition of the errors. The results indicate that the errors in $\langle X \rangle$ and $\sigma(X)$ decay as $m^{-1/2}$, as expected for an unbiased sampling strategy. (This behavior in fact justifies the use of the resolved MRA estimate as a substitute of the exact solution). The low convergence rate of the present sampling scheme is contrasted with that of the MRA scheme, where errors would scale as m^{-2}. Comparison of Figures (2,3) also shows that the number of operations needed to achieve error levels smaller than 10^{-2} is much smaller for MRA than for MC. Thus, for the present stochastic Lorenz problem, the MRA scheme appears to be substantially more efficient than conventional MC.

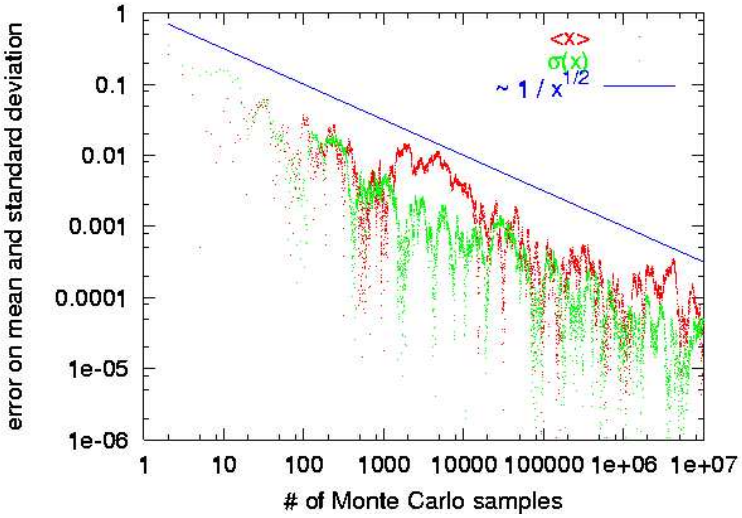

Fig. 3. Absolute errors on $\langle X \rangle$ and $\sigma(X)$ at $t = 25$ versus the number of MC samples.

4.2.2 Latin Hypercube Sampling

A Latin Hypercube Sampling (LHS) strategy [24] is considered in this section in order to improve the convergence of the MC approach. The essential feature of LHS is to divide the probability domain into a set, (Nbin), of bins having equal probability, and then to perform classical MC sampling on each of the bins. This strategy enhances the effectiveness of the MC computations, especially at low values of m, by forcing the the sampling scheme to visit all the bins, before hitting a given bin a second time.

The convergence of LHS scheme is illustrated in Figure 4, where the errors in $\langle X \rangle$ and $\sigma(X)$ at $t = 25$ are plotted against the total number of samples, m. Shown are results obtained using Nbin $= 10,\ 100$ and 1000. The results indicate that LHS enhances the convergence of the MC simulations, especially for low values m ($m < 10,000$). On the other hand, when m/Nbin becomes large, the asymptotic convergence rate of classical MC method, in $1/\sqrt{m}$ is recovered. Despite the improvement over the standard MC approach, the performance of the LHS is significantly lower than that of the MRA.

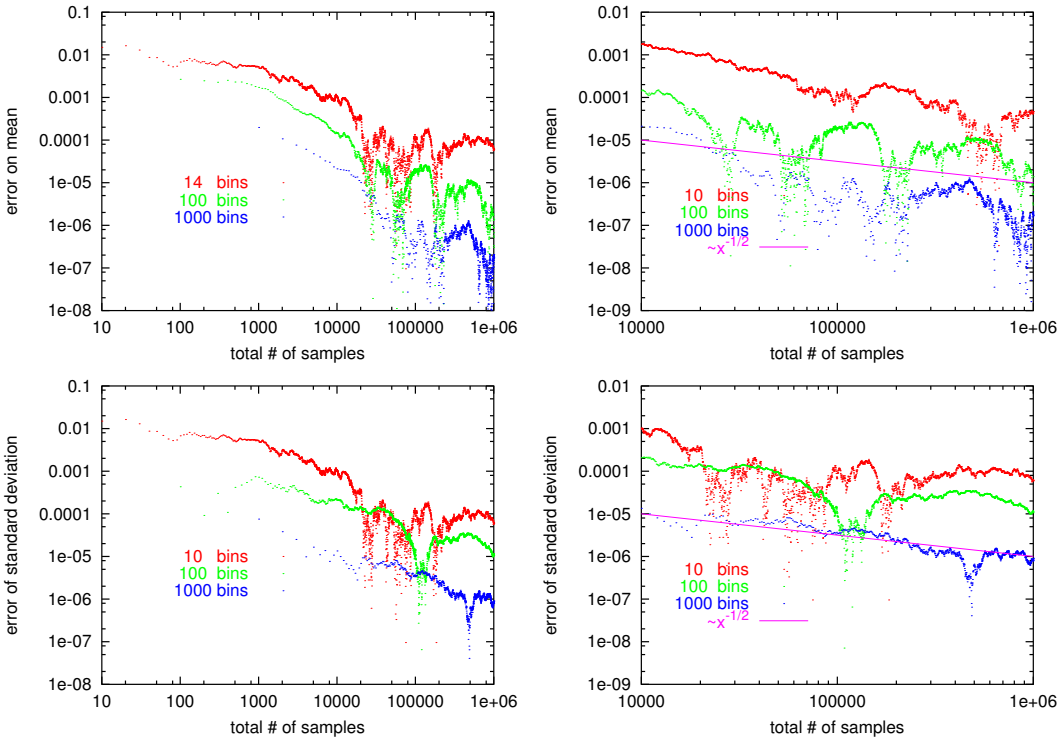

Fig. 4. Absolute errors on $\langle X \rangle$ (top) and $\sigma(X)$ (bottom) for LHS at $t = 25$. The right plots show enlarged views for $10^4 \leq m \leq 10^6$.

4.3 Discussion

The numerical tests of the previous section illustrated the convergence of the MW expansion with increasing order No and resolution levels Nr. The tests also revealed that for the present setting, the MRA scheme offers substantial improvement in efficiency over MC approaches. On the other hand, the tests also indicate that the computational overheads of MRA (as measured using $\mathcal{C}$) rapidly increase as the MW parameters (Nr, No) are refined. This points to a potentially severe limitations in the extension of MRA to stochastic problems with multiple random parameters. Means to overcome such limitations are addressed in the section below, where adaptive strategies are considered in single- and multi-dimensional settings.

5 Adaptive Strategies

Adaptive schemes generally aim at reducing CPU cost by adjusting the quality of the representation locally where needed. In this section, we explore the possibility of locally refining both the local expansion order and the resolution level of the MW expansion. As a first attempt toward such strategy, we shall restrict ourselves in this work to adaptation of the **resolution** level. Adaptive strategies based on both expansion order and resolution will be considered in elsewhere. Within the present framework, two means are attempted to reduce CPU cost. A first possibility (section 5.1) is to disregard the basis components corresponding to low magnitude MW coefficients. This approach is expected to bring substantial saving, since at the highest resolution levels the MW coefficients have non-vanishing values only in areas where the solution exhibits steep dependence on the random data. A second approach (section 5.2) is developed based on partitioning the space of random data, yet maintaining the original expansion at all levels considered.

5.1 Adaptive MW expansion

5.1.1 Refinement strategy

As outlined above, the present approach aims at reducing the CPU cost by adaptively refining the MW expansion. For simplicity, discussion is restricted to the one-dimensional case. The essential concept, which is similar to compression in image processing, is outlined as follows. Given the highest resolution level allowed, Nr, we denote $\overline{\diamondsuit}$ the restriction of $\diamondsuit$ to MW corresponding to resolution levels $\leq Nr$. Then, the solution $\tilde{s}(x)$ is expanded using a reduced basis:

$$\tilde{s}(x) \approx \sum_{\lambda \in \overline{\diamondsuit}_a \subseteq \overline{\diamondsuit}} \tilde{s}_\lambda W_\lambda(x), \tag{39}$$

where $\overline{\diamondsuit}_a \subseteq \overline{\diamondsuit}$ is a reduced set of MW indices such that

$$\overline{\diamondsuit}_a \equiv \{\lambda \in \overline{\diamondsuit} : |\tilde{s}_\lambda| > \varepsilon_a\}. \tag{40}$$

Here, ε_a is a prescribed threshold parameter that may be a function of the resolution level, i.e. $\varepsilon_a = \varepsilon_a(|\lambda|)$. Clearly, for this scheme, the expansion retains only details with significant "energy". Obviously, since $\tilde{s}(x)$ is yet to be determined, it is not possible to reduce $\overline{\diamondsuit}$ *a priori*. To overcome this difficulty, an iterative scheme is constructed, where the representation is successively

refined, by locally increasing the level of resolution only where needed. The algorithm below is used for this purpose:

(1) **Initialization.** We start be computing an initial coarse approximation of $\tilde{s}(x)$, denoted $\tilde{s}^{(l=0)}(x)$, on $\overline{\diamondsuit}_a^{(l=0)} \equiv \Delta_0$.

(2) **Analysis.** Let us denote $\delta\overline{\diamondsuit}^{(l)}$ the set of integer indices defined as

$$\delta\overline{\diamondsuit}^{(l)} \equiv \overline{\diamondsuit}_a^{(l)} \cap \{\beta : |\beta| = l, |\tilde{s}_\beta^{(l)}| > \varepsilon_a(l)\}. \tag{41}$$

We consider on the following two situations:

- If $\delta\overline{\diamondsuit}^{(l)} = \emptyset$ the adaptive scheme has converged, $\tilde{s}^{(l)}(x)$ is the final solution and the iterations are stopped.
- Otherwise, we have $\delta\overline{\diamondsuit}^{(l)} = \{\beta_0, \ldots, \beta_q\}$ then the solution $\tilde{s}^{(l)}$ needs refinement over the union of the supports of the elements of $\delta\overline{\diamondsuit}^{(l)}$,

$$\mathrm{Supp}[\delta\overline{\diamondsuit}^{(l)}] \equiv \cup_{\beta \in \delta\overline{\diamondsuit}^{(l)}} \mathrm{Supp}(\mathrm{W}_\beta).$$

(3) **Refinement.** We set

$$\overline{\diamondsuit}_a^{(l+1)} = \overline{\diamondsuit}_a^{(l)} \cup \delta\overline{\diamondsuit}_a^{(l)},$$

where

$$\delta\overline{\diamondsuit}_a^{(l)} \equiv \left\{\lambda \in \triangle : |\lambda| = l+1, \text{ and } \mathrm{Supp}(\mathrm{W}_\lambda) \subseteq \mathrm{Supp}[\delta\overline{\diamondsuit}^{(l)}]\right\}.$$

(4) **Computation.** Compute $\tilde{s}^{(l+1)}$ the approximation of s spanned by $\{W_\lambda, \lambda \in \overline{\diamondsuit}_a^{(l+1)}\}$. If $l < Nr$, where Nr is a prescribed maximal resolution level, set $l = l+1$ and go back to step 2.

Note that the iterative scheme above generates a continuous cascade of details at successive resolution levels. Specifically, for any index $\lambda \in \overline{\diamondsuit}_a^{(l>0)}$ with $|\lambda| \geq 1$, $\exists \beta \in \overline{\diamondsuit}_a^{(l)}$ such that $|\beta| = |\lambda| - 1$ and $\mathrm{Supp}(\mathrm{W}_\lambda) \subset \mathrm{Supp}(\mathrm{W}_\beta)$. In other words, any MW of resolution level $k > 1$ of the adaptive basis has parents. Thus, one may not be able to guarantee that the iterations give $\lim_{l\to\infty} \overline{\diamondsuit}_a^{(l)} = \diamondsuit_a$, for every function ε_a and process $s(\xi)$. Moreover, while more elaborate versions of the scheme above can be conceived, the simple approach above may still be well-suited for most applications. Below, it is applied to a stochastic Rayleigh-Bénard problem.

5.1.2 Stochastic Rayleigh-Bénard problem

To test the adaptive refinement scheme, we consider the problem initially treated in [19]. The setup consists of a 2D rectangular cavity with insulated

vertical walls. The upper boundary is maintained at a deterministic cold temperature, T_c, while the bottom wall has a spatially-uniform random temperature $T_h(\xi)$. The system is characterized by three non-dimensional parameters: the Rayleigh number $\mathrm{Ra} = \rho g \beta \Delta T H^3/\mu\kappa = 2,150$; the Prandtl number $\mathrm{Pr} = \mu C_p/\kappa = 0.71$, and the aspect ratio $A = L/H = 2$; where ρ is the fluid density, g the gravity, β the thermal expansion coefficient, $\Delta T = \langle T_h \rangle - T_c$, μ the fluid viscosity, κ the thermal conductivity, Cp the heat capacity, L and H the cavity length and height respectively. Denoting by $\Theta = (T - T_{ref})/T_{ref}$ the scaled temperature, where $2T_{ref} = \langle T_h \rangle + T_c$, we assume that $\Theta_h(\xi)$, is uniformly distributed on the interval $[0.3, 0.7]$. The difficulty in the present setup arises from the presence of a critical point, corresponding to $\Theta_h = 0.4301$ [19]. For $\Theta_h < 0.4301$ the system has a stable solution corresponding to a pure conductive regime with vanishing fluid velocity, while for $\Theta_h > 0.4301$ two steady fluid recirculations exist, enhancing the heat transfer across the cavity compared to the conductive regime.

The flow inside the cavity is modeled using the Navier-Stokes equations in the Boussinesq limit [18, 20]. The governing equations are solved on 60×30 computational grid using a second-order scheme. The stochastic projection method (SPM) is used for this purpose [20]. It was shown in [19] that a WLe expansion (corresponding to $Nr = 0$ in the present MRA construction) fails to correctly represent the bifurcation in the uncertainty range, while a WHa expansion (corresponding to $No = 0$ in the MRA scheme) provides robust estimates. The adaptive scheme is now employed using $No = 1$, 2 and 3, and a maximal resolution level $Nr = 6$. For the purpose of the analysis step of the adaptive scheme, a stochastic observable s has to be selected. In the computations below, we rely on the overall heat transfer rate, i.e. we set

$$\tilde{s}^{(l)}(x) = \sum_{\lambda \in \overline{\diamondsuit}_a^{(l)}} W_\lambda(x) \int_0^A \frac{\partial \tilde{\Theta}_\lambda^{(l)}}{A \partial z} dy, \tag{42}$$

where z and y are the normalized vertical and horizontal coordinates. For the refinement threshold, we set $\epsilon_a(\lambda) = 0.03 \times 2^{|\lambda|}$. When the set of indices is increased from $\overline{\diamondsuit}_a^{(l)}$ to $\overline{\diamondsuit}_a^{(l+1)}$, the newly created MW coefficients are initialized to zero, while the others are kept at their previously computed values. Then the flow equations are time-integrated up to the steady state, before being further analyzed for refinement. Note that the simulation time needed to reach the steady state increases with the iteration index l, since the newly added MW are increasingly localized around the critical point, where the growth rate of the instability vanishes. As in [19], the normalized heat transfer enhancement, defined as

$$\delta \tilde{\mathrm{N}}\mathrm{u}^{(l)}(x) \equiv \tilde{s}^{(l)}(x)/(\tilde{\Theta}_h(x) - \Theta_c) - 1,$$

is monitored for the purpose of analyzing the predictions.

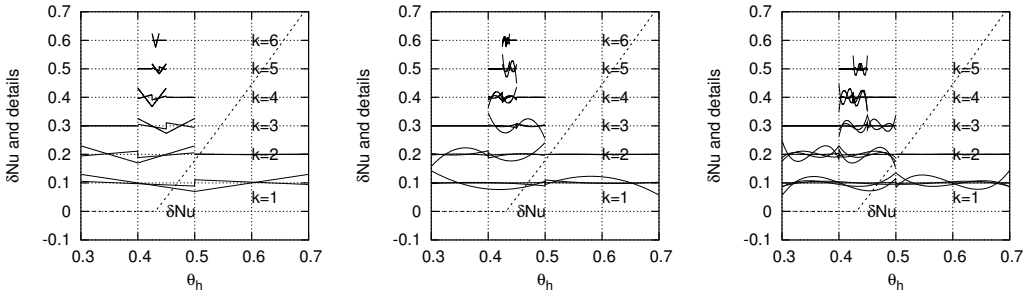

Fig. 5. Computed values δNu versus $\theta_{\rm h}$, for $No = 1$ (left), 2 (middle) and 3 (right). The plots also depict the details $\tilde{s}^{(l)}$ at different resolution levels, k, as indicated. The details are shifted in the vertical direction for clarity.

In Figure 5 shows the values of δNu obtained the end of the iterations are plotted against Θ_h. As for the WHa solution in [19], the discontinuity at the critical temperature is well captured. Also plotted in Fig. 5 are the (rescaled) details of $\tilde{s}^{(l)}$ at different resolution levels ($k > 1$), as obtained at the end of the refinement. The results illustrate how the solution is refined in the neighborhood of the critical point, and only there. In fact, at each iteration l, refinement occurs within the support of the current-level MWs that overlap the critical point. Moreover, for the selected threshold function ϵ_a, it is observed that the iterative scheme proceeds up to the maximal resolution level allowed ($Nr = 6$) for $No = 1$ and 2, but stops at level $k = Nr - 1$ for $No = 3$. Also note that the predictions for the three expansion orders considered are in excellent agreement with each other, suggesting that in all cases accurate predictions are obtained.

For a better appreciation of the efficiency of the adaptive scheme, we provide in Table 5 an analysis of the size of the "adapted" basis. The first three columns provide, respectively, the highest resolution level at the end of the iterations, the dimension of the basis and the corresponding value of $\mathcal{C}_a$. The efficiency of the adaptive refinement is estimated by the "compression" ratios, $\mathrm{Card}_a/\mathrm{Card}(Nr_f, No)$ and $\mathcal{C}_a/\mathcal{C}(Nr_f, No)$, between the adaptively reduced basis and the full basis. As shown in Table 5, these ratios are quite small, and thus reveal a substantial reduction in CPU cost.

Order	Nr_f	Card_a	$\mathcal{C}_a$	$\mathrm{Card}_a/\mathrm{Card}(Nr_f, No)$	$\mathcal{C}_a/\mathcal{C}(Nr_f, No)$
$No = 1$	6	24	2,602	0.1875	0.1076
$No = 2$	6	36	10,855	0.1875	0.1070
$No = 3$	5	40	17,174	0.3125	0.1928

Table 5
Properties of the reduced basis obtained by adaptive refinement. Provide are the maximal refinement level, N_{rf} at the end of the iterations, the dimension of the basis, and the corresponding value of $\mathcal{C}_a$. Also shown are the "compression" ratios, $\mathrm{Card}_a/\mathrm{Card}(Nr_f, No)$ and $\mathcal{C}_a/\mathcal{C}(Nr_f, No)$.

5.1.3 Remarks

The experiences above suggest that in situations requiring a high level of local refinement, lower-order expansions are preferable to higher-order expansions as the former are more likely to have lower compression ratios than the latter. While very low compression ratios may be expected in problems with large number N of stochastic dimensions, extension of the present adaptive refinement scheme to the multidimensional case may prove difficult to implement. In addition, since as noted earlier the dimension of basis and the number of Galerkin operation increase rapidly with N, small compression ratios may not be sufficient to overcome the added complexity in multi-dimensional problems. These observations in part motivate the alternative approach below.

5.2 Adaptive partitioning of random parameter space

The experiences of the previous adpative section indicate that it is possible to minimize the CPU load of MRA of stochastic problems by limiting the number of overlapping MW supports. However, one drawback of the scheme above is that the complexity of the spectral product, which reflects the coupling between MW coefficients, is quickly increasing. This is due to the generation of a **continuous** cascade of details at successive scales. This observation raises the question whether a truly local analysis, that decouples the representation at different scales, can lead to even more efficient computations.

In this section, we develop a local refinement scheme based on the expansion in Eq. (17). Comparison of the two expansions in Eqs. (17) and (18) shows that the former, in terms of ϕ functions, does not involve any summation over the scale indices, contrary to the expansion in terms of details, ψ. This difference stems from the fact that the basis functions ϕ^k_{il}, $i = 0, \ldots, No$ and $l = 0, \ldots, 2^k - 1$, couple with with only No other components, namely those having the same sliding index l. In contrast, the ψ^k_{il} couple with many other components. This suggests an adaptive strategy based on successive partitions of the random parameter space, through the determination of a **local** resolution level. These considerations lead us to the second adaptive scheme below.

5.2.1 Partition of the random parameter space

Let $\Omega = [a_1, b_1] \times \ldots \times [a_N, b_N]$ be space of random parameters. Let Ω^m, $m = 1, \ldots, Nb$ be a finite partition of Ω in Nb non-overlapping subdomains :

$$\begin{cases} \Omega^m = [a_1^m, b_1^m] \times \ldots [a_N^m, b_N^m], \\ \Omega = \bigcup_{m=1}^{Nb} \Omega^m, \\ \Omega^m \cap \Omega^{m'} = \varnothing \quad \text{if } m \neq m'. \end{cases} \tag{43}$$

On each of the subdomains Ω^m we define the local probability density function of $\boldsymbol{\xi}$, denoted by $\mathrm{pdf}^m(\boldsymbol{\xi})$. Since the components of $\boldsymbol{\xi}$ are independent, we have:

$$\mathrm{pdf}^m(\boldsymbol{\xi}) = \prod_{d=1}^{N} \mathrm{pdf}_d^m(\xi_d), \tag{44}$$

where $\mathrm{pdf}_d^m(\xi)$ is defined according to

$$\mathrm{pdf}_d^m(\xi) \equiv \frac{\mathrm{pdf}_d(\xi)}{\mathrm{p}_d(b_d^m) - \mathrm{p}_d(a_d^m)}. \tag{45}$$

Clearly, we have

$$\mathrm{pdf}^m(\boldsymbol{\xi}) > 0 \quad \text{for } \boldsymbol{\xi} \in \Omega^m \text{ and } \int_{\Omega^m} \mathrm{pdf}^m(\boldsymbol{\xi}) d\boldsymbol{\xi} = 1. \tag{46}$$

Moreover, defining

$$\mathrm{p}_d^m(\xi \in [a_d^m, b_d^m]) \equiv \int_{a_d^m}^{\xi} \mathrm{pdf}_d^m(\xi') d\xi' = r_d^m, \tag{47}$$

we see that $r_d^m \in [0, 1]$. Thus, if x_d^m is uniformly distributed over $[0, 1]$, then the random variable $(\mathrm{p}_d^m)^{-1}(x_d^m) \in [a_b^m, b_d^m]$ and has the same distribution as ξ_d^m. Thus, a second-order stochastic process can be locally expanded on Ω^m, in terms of the random vector $\boldsymbol{x}^m$, having independent components $(x_1^m, \ldots, x_N^m) \in [0, 1]^N$.

5.2.2 Local expansion basis

Let $\boldsymbol{\gamma}$ be the set of multidimensionnal indices

$$\boldsymbol{\gamma} = \left\{ (\gamma_1, \ldots, \gamma_N) : \sum_{d=1}^{N} \gamma_d \leq No \right\}.$$

For $\boldsymbol{\xi} \in \Omega^m$, we build the local projection basis as

$$\mathcal{B}_p(\Omega^m) \equiv \left\{ \Phi^m_{\boldsymbol{\lambda} \in \boldsymbol{\gamma}}(x_1^m, \ldots, x_N^m) \equiv \prod_{d=1}^{N} \phi_{\lambda_d}(x_d^m) \right\},$$

and the details directional basis $\mathcal{B}_a^d$, $d = 1, \ldots, N$ as

$$\mathcal{B}_a^d(\Omega^m) \equiv \{\psi_i(x_d^m), i = 0, \ldots, No\}.$$

The complete local expansion basis will be the union of $\mathcal{B}_p$ and $\mathcal{B}_a^d$:

$$\mathcal{B}(\Omega^m) = \mathcal{B}_p(\Omega^m) + \bigcup_{d=1}^{N} \mathcal{B}_a^d(\Omega^m).$$

Finally, the multidimensional process $\mathcal{P}(\boldsymbol{\xi})$ will have for local expansion on Ω^m:

$$\mathcal{P}(\boldsymbol{\xi}) \approx \sum_{\boldsymbol{\lambda} \in \boldsymbol{\gamma}} \tilde{\mathcal{P}}^m_{\lambda_1, \ldots, \lambda_N} \Phi_{\lambda_1, \ldots, \lambda_N}(x_1^m, \ldots, x_N^m) + \sum_{d=1}^{N} \sum_{i=0}^{No} \tilde{\mathcal{P}}^m_{d,i} \psi_i(x_d^m). \tag{48}$$

Note that the local basis $\mathcal{B}(\Omega^m)$, spanning the local expansion of $\mathcal{P}$ according to Eq. (48), is in fact the rescaled Legendre polynomials basis ($\mathcal{B}_p(\Omega^m)$) augmented with the first-level detail basis, $\mathcal{B}_a^d$, $d = 1, \ldots, N$. Thus, $\mathcal{P}(\boldsymbol{\xi})$, for $\boldsymbol{\xi} \in \Omega^m$, approximated by Eq. (48) is the local Wiener-Legendre projection of order No, plus one dimensional details. The utility of the one dimensional details will be made clear soon.

The local expectation of $\mathcal{P}$ is given by $\langle \mathcal{P} \rangle_{\Omega^m} = \mathcal{P}^m_{0,\ldots,0}$, and its local variance is

$$\sigma^2_{\Omega^m}(\mathcal{P}) \approx \left(\hat{\sigma}_{\Omega^m}\right)^2 + \sum_{d=1}^{N} \left(\sigma^d_{\Omega^m}\right)^2,$$

where we have denoted

$$\left(\hat{\sigma}_{\Omega^m}\right)^2 \equiv \sum_{\boldsymbol{\lambda} \in \boldsymbol{\gamma}_p} (\mathcal{P}^m_{\boldsymbol{\lambda}})^2, \quad \left(\sigma^d_{\Omega^m}\right)^2 \equiv \sum_{i=0}^{No} \left(\mathcal{P}^m_{d,i}\right)^2,$$

with $\boldsymbol{\gamma}_p = \boldsymbol{\gamma} - \{0, \ldots, 0\}$.

5.2.3 Global statistics

The total expectation of the process is given by the volume-weighted summation of the local expectations:

$$\langle \mathcal{P} \rangle = \sum_{m=1}^{Nb} \langle \mathcal{P} \rangle_{\Omega^m} \, \mathrm{Vol}^m, \tag{49}$$

where Vol^m is the volume of the m-th subdomain of the random parameter space:

$$\mathrm{Vol}^m = \prod_{d=1}^{N} (\mathrm{p}_d(b_d^m) - \mathrm{p}_d(a_d^m)).$$

Finally the total variance of the process is given by

$$\sigma^2(\mathcal{P}) = \sum_{m=1}^{Nb} \left[\sigma^2_{\Omega^m}(\mathcal{P}) + \left(\mathcal{P}^m_{0,\ldots,0} - \langle \mathcal{P} \rangle \right)^2 \right] \mathrm{Vol}^m. \tag{50}$$

5.2.4 Adaptive strategy

Assume that the current partition of Ω involves Nb blocks, *i.e.* $\Omega = \bigcup_{m=1}^{Nb} \Omega^m$. On each subdomain Ω^m, the process is expanded on the local basis $\mathcal{B}(\Omega^m)$, the spectral coefficients being computed through Galerkin projection methods as in the previous test problems. To decide if a given block m needs more refinement, and to determine which stochastic directions need such refinement, we consider the following test:

$$\text{if } \frac{\sigma^d_{\Omega_m}}{\sigma_{\Omega_m}} \geq \varepsilon_2(\mathrm{Vol}^m). \tag{51}$$

If the inequality is satisfied, the subdomain is refined along the d-th dimension. Here $\varepsilon_2(\mathrm{Vol}^m) < 1$ is a prescribed threshold function. Note that the test compares the "energy" of the one-dimensional details along the d-th stochastic direction with the local variance of the solution. In other words, the one-dimensional details coefficients are used as indicators of the quality of the representation along their respective stochastic direction. A new partition of Ω is then constructed, by splitting Ω^m into smaller subdomains. Specifically, if we assume the inequality (51) is satisfied for a single dimension d, then refinement of $\Omega^m = [a_1^m, b_1^m] \times \ldots \times [a_N^m, b_N^m]$ will give birth to two new subdomains $\Omega^{m'}$

and $\Omega^{m''}$, defined by:

$$\begin{cases} \Omega^{m'} &= [a_1^{m'}, b_1^{m'}] \times \ldots \times [a_N^{m'}, b_N^{m'}] \\ &= [a_1^m, b_1^m] \times \ldots \times [a_d^m, (a_d^m + b_d^m)/2] \times \ldots \times [a_N^m, b_N^m] \\ \Omega^{m''} &= [a_1^{m''}, b_1^{m''}] \times \ldots \times [a_N^{m''}, b_N^{m''}] \\ &= [a_1^m, b_1^m] \times \ldots \times [(a_d^m + b_d^m)/2, b_d^m] \times \ldots \times [a_N^m, b_N^m]. \end{cases} \tag{52}$$

Then, local expansions of the process on the newly created subdomains are computed, before being analyzed to determine whether additional refinement is needed. This sequence of analysis and refinement steps is repeated up to convergence. It is emphasized that, during refinement, computations are performed in newly created subdomains only, since the local solutions over other subdomains are unaffected, the expansions being local. Note also that this methodology is well suited for parallel implementation, since local computations are independent of each other.

5.3 Adaptive partitioning of the random parameter space

5.3.1 Test problem

We consider the following test problem:

$$\begin{cases} \dfrac{d\rho}{dt} = \alpha(1-\rho) - \gamma\rho - \beta(\rho-1)\rho^2 \\ \rho(t=0) = \rho_0 \end{cases} \tag{53}$$

which models the time-evolution of the surface coverage $\rho \in [0,1]$ for a given species [22]. α is the surface absorption rate γ is the desorption rate, and β is the recombination rate (the exponent 2 is due to the need of two sites for a recombination) [22]. This problem has one or two fixed points according to the value of β and it exhibits smooth dependence on the other parameters. In section 5.3.2, the statistics of the solution at $t = 1$ is investigated considering uncertainties in the initial coverage ρ_0 and in the reaction parameter β. Next in Section 5.3.3, uncertainty in α is also considered, increasing the number of stochastic dimensions to 3. We shall consider that ρ_0 is uniformly distributed in the range $[0,1]$; the distributions of other random parameters are specified later. In order to propagate the uncertainty and determine its impact on the solution, the Galerkin scheme is applied to the governing equation (53). The resulting coupled system of ODEs is integrated in order to determine the evolution of the expansion coefficients. A fourth order Runge-Kutta scheme is used for this purpose, with a time step $\Delta t = 0.01$.

5.3.2 Two-dimensional problem

In this section, deterministic values of the absorption and desorption rates are used, respectively $\alpha = 1$ and $\gamma = 0.01$. Meanwhile, β is assumed to be uniformly distributed in the interval $[0, 20]$. Thus the problem has two stochastic dimensions, $\boldsymbol{\xi} = \{\xi_1, \xi_2\}$ with ξ_1 uniformly distributed in $[0, 1]$ and ξ_2 uniformly distributed in $[0, 20]$.

Computations are performed with an expansion order $No = 3$, and the results are used to analyze the statistics of the surface coverage, ρ, at time $t = 1$. Figure 6 illustrates the computed solution, obtained with different values of the threshold ε_2. The latter is expressed as $\varepsilon_2(\text{Vol}) = C/\sqrt{\text{Vol}^m}$, where C is a prescribed constant. Four values are considered, namely $C = 0.5$, 0.1, 0.01, and 0.001. Shown in Fig. 6 are the predicted response surface of ρ plotted against the random input data, the partitions of random parameter space as determined by the adaptive scheme, and the pdf of $\rho(t = 1)$. The latter is computed by direct sampling of the spectral expansions using 10^6 samples.

The results show that with $\varepsilon_2 = 0.5/\sqrt{\text{Vol}^m}$ only one block is generated, i.e. no subdivisions are performed. The corresponding response surface and pdf reflect a poor approximation, particularly since unphysical realizations with $\rho < 0$ and $\rho > 1$ are predicted. When the threshold parameter is decreased to $0.1/\sqrt{\text{Vol}^m}$, one observes that the solution involves $Nb = 5$ subdomains. Note that refinement is concentrated in areas where the solution exhibits the steepest dependence on random data, as one would expect based on the construction of the adaptive scheme. Also note that no unphysical realizations are predicted for this threshold function, as all predicted value of ρ fall between 0 and 1. As ε_2 is further decreased (last two rows of Fig. 6), the partition becomes increasingly more refined, particularly along directions of sharp variation with respect to the random data. At the lower values of C, the insensitivity of the pdf to the selected value of C can be easily appreciated.

For a better appreciation of the convergence of the adaptive scheme, we report in Tables 6-8 the computed mean and standard deviation of $\rho(t = 1)$ for decreasing values of C, respectively for third, second, and first-order expansions. The tables also provide the number of subdomains generated during adaptive refinement, as well as estimates of the CPU load and the memory requirement. The CPU load is estimated as the product of $\mathcal{C}(1, No)$ with the *total* number of subdomains computed in order for the adaptive refinement to converge. The latter includes the number of subdomains that are discarded due to refinement. The memory requirement is based on the number of MW coefficients of the most refined solution, equal to the product of the final number of subdomains with the dimension of the expansion basis over a single block.

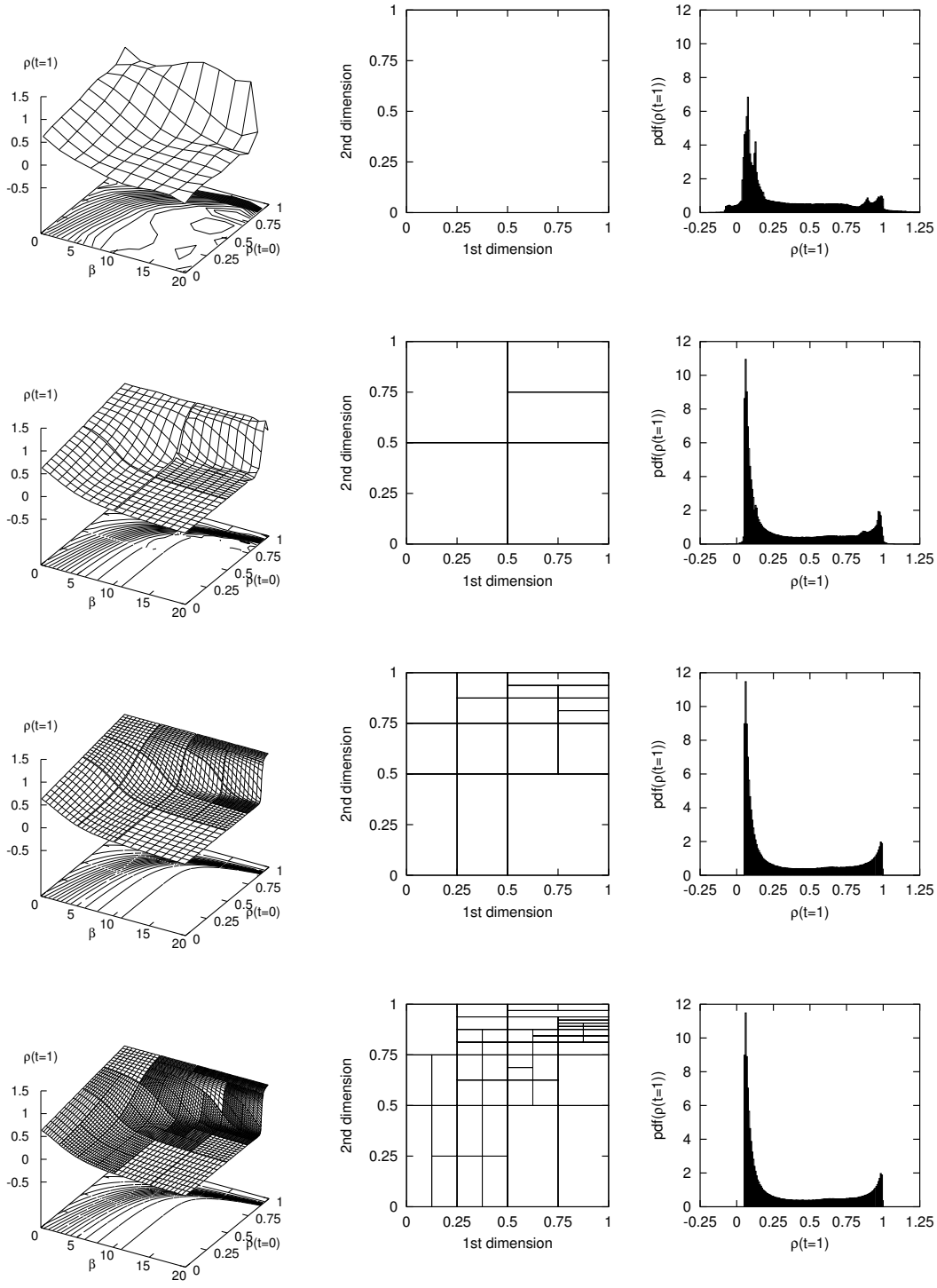

Fig. 6. Surface response (left), partition of the random parameter space (middle) and estimated probability density function (right) of the solution at $t = 1$. Results are obtained for $\varepsilon_2(\mathrm{Vol}^m) = 0.5/\sqrt{\mathrm{Vol}^m}$, $0.1/\sqrt{\mathrm{Vol}^m}$, $0.01/\sqrt{\mathrm{Vol}^m}$ and $0.001/\sqrt{\mathrm{Vol}^m}$, arranged from top to bottom. The expansion order $No = 3$.

$\varepsilon_2(\mathrm{Vol}^m)$	Nb	$\langle\rho\rangle$	$\sigma(\rho)$	CPU	Memory
$5.10^{-1}/\sqrt{\mathrm{Vol}^m}$	1	0.338,366,540	0.329,585,455	8 301	48
$1.10^{-1}/\sqrt{\mathrm{Vol}^m}$	5	0.350,449,337	0.332,070,259	58 107	240
$1.10^{-2}/\sqrt{\mathrm{Vol}^m}$	16	0.350,410,081	0.332,172,422	224 127	768
$1.10^{-3}/\sqrt{\mathrm{Vol}^m}$	47	0.350,410,331	0.332,171,412	622 575	2 256
$1.10^{-4}/\sqrt{\mathrm{Vol}^m}$	151	0.350,410,331	0.332,171,421	1 884 327	7 248
$1.10^{-5}/\sqrt{\mathrm{Vol}^m}$	439	0.350,410,331	0.332,171,421	5 486 961	21 072

Table 6
Computed values of $< \rho >$ and $\sigma(\rho)$ for $No = 3$ and different threshold functions. Also provided are the number of blocks Nb at the end of the refinement, and estimates of the CPU load and memory requirement.

$\varepsilon_2(\mathrm{Vol}^m)$	Nb	$\langle\rho\rangle$	$\sigma(\rho)$	CPU	Memory
$5.10^{-1}/\sqrt{\mathrm{Vol}^m}$	1	0.346,230,683	0.335,764,259	1 667	27
$1.10^{-1}/\sqrt{\mathrm{Vol}^m}$	6	0.350,244,115	0.332,167,461	15 003	162
$1.10^{-2}/\sqrt{\mathrm{Vol}^m}$	34	0.350,410,968	0.332,173,480	85 017	918
$1.10^{-3}/\sqrt{\mathrm{Vol}^m}$	129	0.350,410,326	0.332,171,421	325 065	3 483
$1.10^{-4}/\sqrt{\mathrm{Vol}^m}$	514	0.350,410,330	0.332,171,421	1 291 925	13 878
$1.10^{-5}/\sqrt{\mathrm{Vol}^m}$	1 823	0.350,410,331	0.332,171,421	4 322 531	49 221

Table 7
Computed values of $< \rho >$ and $\sigma(\rho)$ for $No = 2$ and different threshold functions. Also provided are the number of blocks Nb at the end of the refinement, and estimates of the CPU load and memory requirement.

One observes from Tables 6-8 that for all expansion orders considered the adaptive predictions appear to converge to same values of mean and standard deviation, and that the solution for $No = 3$ appears most efficient, as the stopping criterion is achieved with the smallest number of subdomains, as well as smaller CPU and memory requirement. This behavior suggests that the solution is essentially smooth everywhere except on localized areas which are well captured and refined by the adaptive refinement scheme, as illustrated in Fig. 7. Comparison of the results in Fig. 7 also suggests that further enhancement of the present scheme could be achieved based on generalizing the refinement strategy so that the local expansion order is also adapted (reduced) during refinement. Excessive refinement in smooth regions observed with $No = 1$ also indicates that the threshold function may need to be related to the expansion order. Such generalizations will be considered in future work.

In the numerical experiments below, the threshold parameter is decreased before the refinement is stopped. Specifically, starting from $\varepsilon_2(\mathrm{Vol}^m) = C/\sqrt{\mathrm{Vol}^m}$, with $C = 0.5$, the adaptive refinement is carried out until the criterion is sat-

$\varepsilon_2(\mathrm{Vol}^m)$	Nb	$\langle\rho\rangle$	$\sigma(\rho)$	CPU	Memory
$5.10^{-1}/\sqrt{\mathrm{Vol}^m}$	1	0.333,617,359	0.318,197,117	190	12
$1.10^{-1}/\sqrt{\mathrm{Vol}^m}$	8	0.349,221,435	0.330,668,616	2 470	96
$1.10^{-2}/\sqrt{\mathrm{Vol}^m}$	103	0.350,410,367	0.332,170,754	30 970	1 236
$1.10^{-3}/\sqrt{\mathrm{Vol}^m}$	889	0.350,410,309	0.332,171,494	265 050	10 668
$(1.10^{-4}/\sqrt{\mathrm{Vol}^m})^*$	3301	0.350,410,329	0.332,171,426	892 050	39 612

Table 8

Computed values of $< \rho >$ and $\sigma(\rho)$ for $No = 1$ and different threshold functions. Also provided are the number of blocks Nb at the end of the refinement, and estimates of the CPU load and memory requirement. *With $\varepsilon_2 = 1.10^{-4}/\sqrt{\mathrm{Vol}^m}$, stopping criterion was not reached; further refinement would have required a larger number of subdomains than allowed (5000).

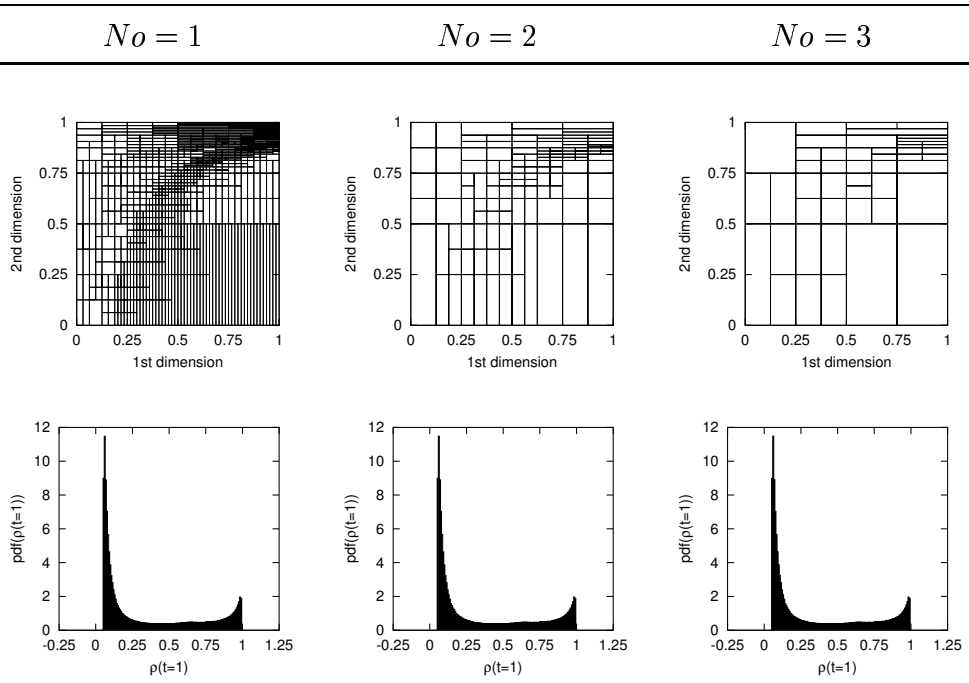

Fig. 7. Partition of random parameter space (top row) and pdf of $\rho(t = 1)$ (bottom row), for $No = 1$, 2 and 3 (left to right). Results were obtained with $\varepsilon_2(\mathrm{Vol}^m) = 10^{-3}/\sqrt{\mathrm{Vol}^m}$.

isfied over all subdomains. Then, the constant C is multiplied by a factor of 0.8, and the analysis is repeated. We monitor the convergence of the solution by computing the errors in the expectation and standard deviation of ρ at $t = 1$, using the solution obtained at the end of the refinement with $No = 3$ as surrogate for the exact solution. In Fig. 8, we plot the absolute values of $< \rho > - < \rho >_{ex}$ and of $\sigma(\rho) - \sigma_{ex}(\rho)$ as a function of the number of blocks Nb. These plots allow us to estimate how quickly the solutions converge as

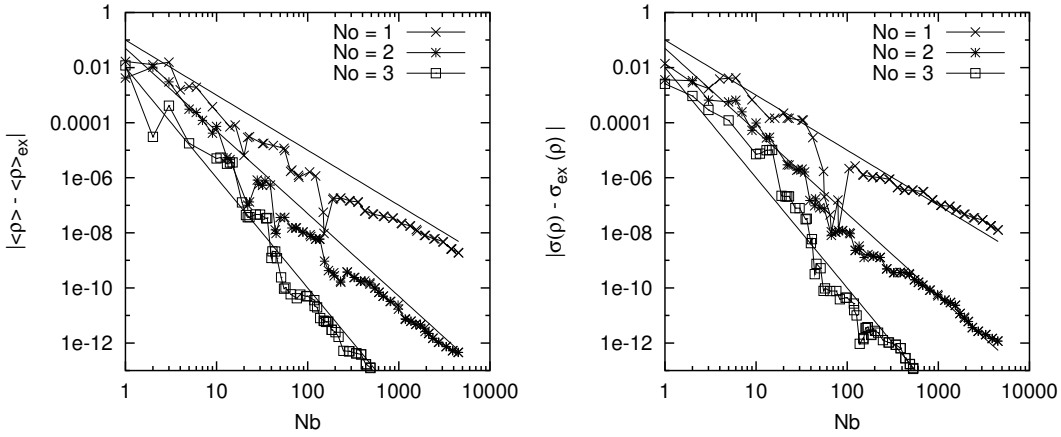

Fig. 8. Estimates of the error in $\langle\rho\rangle$ (left) and $\sigma(\rho)$ (right) at $t = 1$ versus the the number Nb of subdomains. Plotted are results obtained with $No = 1$, 2 and 3. The refinement criterion is progressively made more stringent, as described in the text. Solid lines illustrate decay rates proportional to $\sim Nb^{-(No+1)}$.

the partition is refined. The results reveal that the errors decay as $\sim Nb^{-No-1}$ for the three tested cases. Note that if the partition were uniform, then the number of blocks needed to achieve a similar level of accuracy would be much larger than that computed by the adaptive scheme. In particular, if the MRA of section 4 were to be applied without refinement at the highest refinement level reached in the adaptive scheme, the computations would require excessive CPU and memory requirements and thus would prove impractical. This highlights the efficiency of the adaptive computations.

Another assessment of the efficiency of the adaptive scheme is performed by estimating the CPU cost needed to achieve a given level of accuracy. Since actual CPU cost is generally dependent on the computing platform and implementation details, the analysis below are based on the CPU load; as in previous sections, the latter is estimated as the product of $\mathcal{C}$ with the total number of blocks that are computed. This enables us direct comparison with MC approach, for which the CPU load is simply the number of realizations. Figure 9 shows the error in the mean and standard deviation of $\rho(t = 1)$ as a function of the CPU-load, denoted CPU, of the spectral adaptive computations using $No = 1,\ 2$, and 3. Also shown is the error in the MC simulation, plotted against the sample size m. The results indicate that for the present setup the adaptive computations result in very small error estimates. In contrast, the small convergence rate of MC simulations practically limits the accuracy of the corresponding predictions. The figures show that in the intermediate range of CPU, lower errors are achieved with lower expansion order; as the CPU increases, however, accuracy improves with increasing expansion order. This trend supports earlier suggestion that a more efficient approach may be constructed in which the expansion order is also adapted during refinement.

Additional insight into the behavior of the errors is gained by repeating the

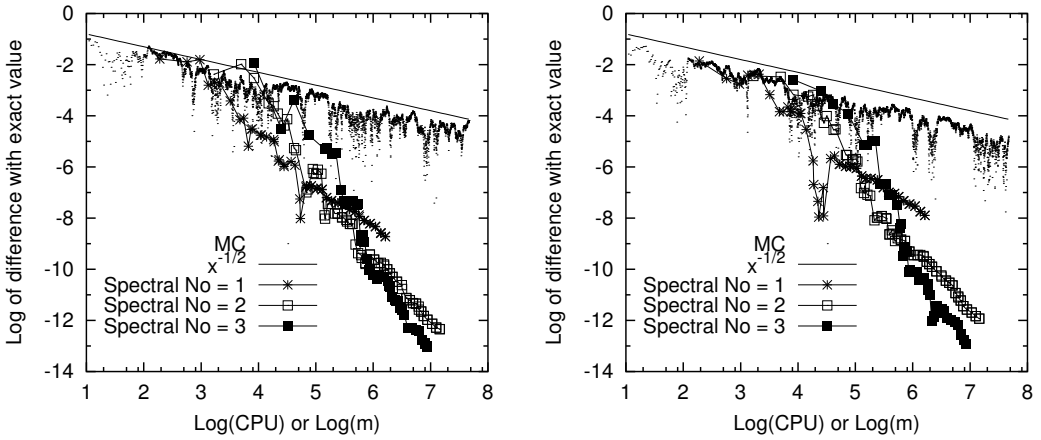

Fig. 9. Estimates of the error in $\langle\rho\rangle$ (left) and $\sigma(\rho)$ (right) at $t = 1$ versus the the CPU load in the adaptive computations. Plotted are results obtained with $No = 1$, 2 and 3. The refinement criterion is progressively made more stringent, as described in the text. Also shows are errors in MC simulations, plotted against the number of samples, m. The solid line illustrates a decay rate proportional to $m^{-1/2}$.

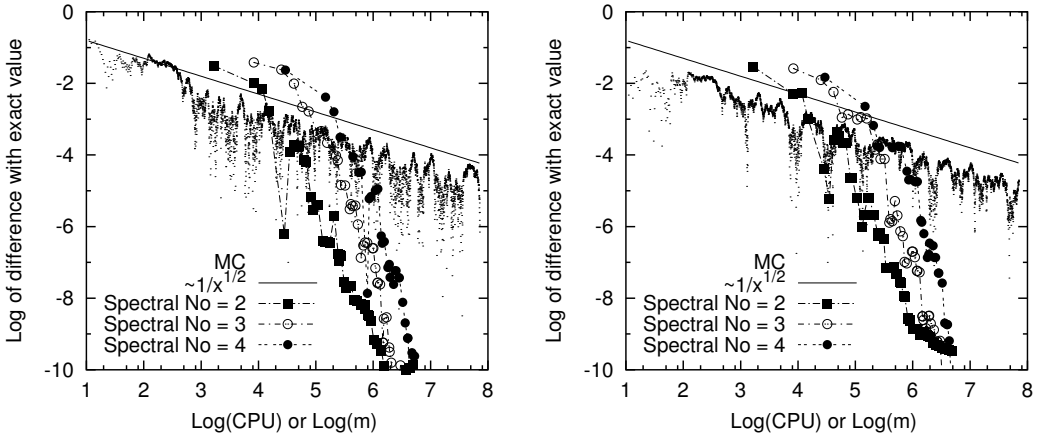

Fig. 10. Estimates of the error in $\langle\rho\rangle$ (left) and $\sigma(\rho)$ (right) at $t = 2$ versus the the CPU load in the adaptive computations. Plotted are results obtained with $No = 2$, 3 and 4. The refinement criterion is progressively made more stringent, as described in the text. Also shows are errors in MC simulations, plotted against the number of samples, m. The solid line illus-rates a decay rate proportional to $m^{-1/2}$.

same experiment, using second, third, and fourth-order expansion, and extending the simulation up to $t = 2$. As shown in [22], as time increases the solution develops sharper fronts, eventually becoming discontinuous at large time. Thus, one would expect that the convergence of higher-order expansions deteriorates as t increases. This is in fact observed in Fig. 10, where errors in $< \rho >$ and $\sigma(\rho)$ at $t = 2$ are plotted against CPU load. In particular, the results show that in most of the CPU range considered, second-order predictions have smaller errors than third- and fourth-order computations. Also note that regardless of the expansion order, the adaptive computations clearly outperform MC simulations.

5.3.3 Higher dimensional problems

We conclude this section by briefly demonstrating the application of the adaptive scheme to higher dimensional problem. Specifically, results are presented for the case of three random dimensions. The same surface reaction model is used, but uncertainty in the reaction rate is now considered, with α assumed to be uniformly distributed in the range $[0.1, 2]$. In Fig. 11, we plot the partition generated during refinement with different threshold functions. The figure shows that the adaptive can accommodate multiple dimensions, and illustrates how refinement is selectively applied along different directions in the space of random data. Also plotted in Fig. 11 is the pdf of the edge-size of the subdomains along the different dimensions; the results are generated using the finest partition ($C = 10^{-3}$). The plot illustrates that more refinement has been applied along the second direction (β) than along the first (initial condition, $\rho(t = 0)$), and that more refinement has been applied on the first direction than the third (α). This last result is not surprising since the dependence of ρ on α, in the range considered, is everywhere smoother than along the other two directions.

6 Conclusions

A MRA is applied to an uncertainty propagation scheme based on a generalized PC representation. The MRA is based on an orthogonal projection of uncertain data and solution variables onto a MW basis consisting of piecewise-smooth polynomials. A Galerkin procedure is then used to determine equations of motion for the MW expansion coefficients. Integration of the resulting coupled system yields the solution of the stochastic problem.

The behavior of the MRA is first examined through computations of a stochastic Lorenz system. The computations indicate that the MW expansion converges as one increases the number of resolution levels and/or the order of the expansion. Convergence is verified by comparing the results with MC predictions, and the comparison indicates has faster convergence rate and is significantly more efficient than either MC or LH sampling.

Analysis of the Lorenz system also indicates that overheads of the MRA rapidly increase as the parameters of MW expansion are refined. To address this potential limitation, two adaptive strategies are presented. The first strategy is based on local refinement of the MW expansion. The refinement is based on computing the "energy" content in the highest resolution levels, locally refining the expansion where this measure exceeds a prescribed threshold. Implementation of this adaptive refinement strategy is illustrated through computations of a stochastic Rayleigh-Bénard problem. The case of a stochastic

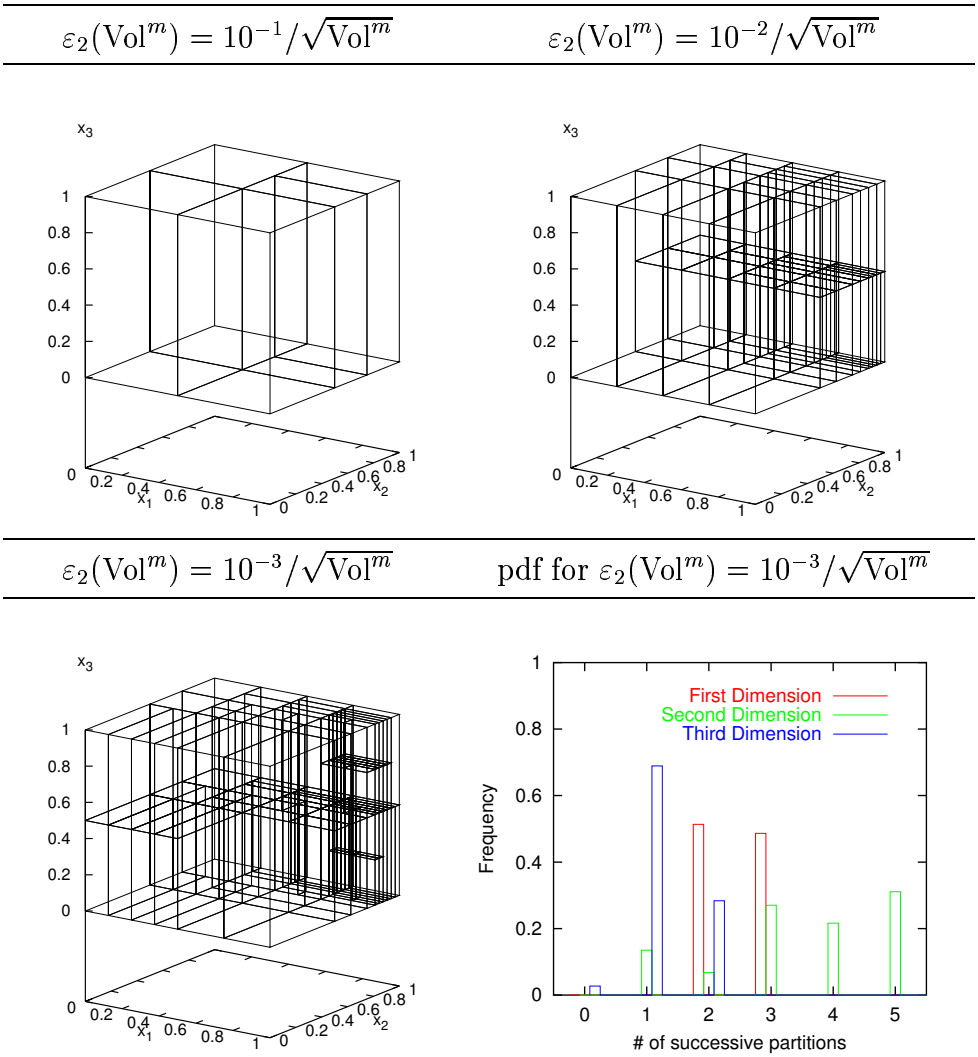

Fig. 11. Partition of 3D random parameter space using the adaptive scheme. The refinement threshold is indicated. Also shown in the pdf of the edge-size of the subdomains along the three directions.

Rayleigh number is considered; the latter is taken to be uniformly distributed in a finite interval containing the critical value. Simulations indicate that the adaptive scheme can effectively handle such complex situations, as the refinement is localized in the neighborhood of the bifurcation. In particular, the present experience suggests that in situations requiring a high level of local refinement, lower-order adaptive expansions may be preferable to higher-order ones, as they are likely to result in more efficient predictions. On the other hand, the analysis also indicates that for problems with a large number of stochastic dimensions, the improvements of the local refinement strategy may not be sufficient to overcome the added complexity of multi-dimensional problems.

An alternative adaptive strategy is finally considered in order to overcome

potential limitations in multi-dimensional problems. This second approach is based on adaptive block-partioning scheme of the space of random data. On each block, the solution is expanded in terms of a MW expansion consisting of smooth global functions and 1D details. The blocks are refined by division along individual dimensions whenever the contribution to the local variance of the corresponding 1D details exceeds a prescribed threshold. The solution is then recomputed on the newly generated blocks. Implementation of the resulting adaptive scheme is illustrated based on computations of a surface-kinetic problem having stochastic initial conditions and rate constants. Analysis of the behavior of the scheme indicates that the refinement is naturally concentrated in areas of steep variation of the the solution with respect to the random data. The analysis also indicates that errors decay rapidly as the number of blocks increases, and that the rate of decay increases with increasing order. For the present setup, the computations indicate that when the desired level of accuracy is not very large, lower-order expansions may prove more efficient than higher-order expansions.

The present experiences indicate that adaptive refinement provides an attractive means for tackling complex, multidimensional stochastic problems. Specifically, it offers the possibility of constructing efficient and robust schemes that are able to effectively tackle situations exhibiting steep or discontinuous dependence on random data. The computations also highlight several areas where substantial enhacement may be achieved. These include the development of more elaborate strategies in which the order of the expansion is increased or reduced simultaneously with "spatial" refinement, and the construction of more efficient refinement criteria that reduces the overheads of the local analysis. These generalizations are the focus of ongoing work.

References

[1] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions.* Dover, 1970.

[2] B. Alpert, G. Beylkin, D. Gines, and L. Vozovoi. Adaptive solution of partial differential equations in multiwavelet bases. *Journal of Computational Physics*, 182:149–190, 2002.

[3] B.K. Alpert. A class of bases in L_2 for the sparse representation of integral operators. *SIAM J. Math. Anal.*, 24:246–262, 1993.

[4] I. Babuska, R. Tempone, and G.E. Zouraris. Galerkin finite element approximations of stochastic elliptic differential equations. Report 02-38, TICAM, 2002.

[5] R.H. Cameron and W.T. Martin. The orthogonal development of nonlinear

functionals in series of Fourier-Hermite functionals. *Ann. Math.*, 48:385–392, 1947.

[6] A.J. Chorin. Hermite expansions in Monte Carlo computation. *Journal of Computational Physics*, 8:472–482, 1971.

[7] A.J. Chorin. Gaussian fields and random flow. *Journal of Fluid Mechanics*, 63:21–32, 1974.

[8] I. Daubechies. *Ten Lectures on Wavelets.* SIAM, 1992.

[9] M.K. Deb, I.M. Babuska, and J.T. Oden. Solution of stochastic partial differential equations using Galerkin finite element techniques. *Computer Methods in Applied Mechanics and Engineering*, 190:6359–6372, 2001.

[10] B. Debusschere, H.N. Najm, A. Matta, O.M. Knio, R.G. Ghanem, and O.P. Le Maître. Protein labeling reactions in electrochemical microchannel flow: Numerical prediction and uncertainty propagation. *Physics of Fluids*, 15(8):2238–2250, 2003.

[11] R. Ghanem. Ingredients for a general purpose stochastic finite elements formulation. *Computer Methods in Applied Mechanics and Engineering*, 168:19–34, 1999.

[12] R. Ghanem and S. Dham. Stochastic finite element analysis for multiphase flow in heterogeneous porous media. *Transport in Porous Media*, 32:239–262, 1998.

[13] R.G. Ghanem. Probabilistic characterization of transport in heterogeneous media. *Computer Methods in Applied Mechanics and Engineering*, 158:199–220, 1998.

[14] R.G. Ghanem and P.D. Spanos. *Stochastic Finite Elements: A Spectral Approach.* Springer Verlag, 1991.

[15] M. Grigoriu. *Stochastic Calculus, Applications in Sciences and Engineering.* Birkhäuser, 2002.

[16] T.D. Hien and M. Kleiber. Stochastic finite element modeling in linear transient heat transfer. *Computer Methods in Applied Mechanics and Engineering*, 144:111–124, 1997.

[17] O.M. Knio and R.G. Ghanem. Polynomial Chaos product and moment formulas : A user utility. Technical report, The Johns Hopkins University, Baltimore, MD, to appear.

[18] O.P. Le Maître, O.M. Knio, H.N. Najm, and R.G. Ghanem. A stochastic projection method for fluid flow. i. basic formulation. *Journal of Computational Phyics*, 173:481–511, 2001.

[19] O.P. Le Maître, H.N. Najm, R.G. Ghanem, and O.M. Knio. Uncertainty propagation using Wiener-Haar expansions. *Journal of Computational Phyics*, 2003. (in press).

[20] O.P. Le Maître, M.T. Reagan, H.N. Najm, R.G. Ghanem, and O.M. Knio. A stochastic projection method for fluid flow. ii. random process. *Journal of Computational Phyics*, 181:9–44, 2002.

[21] J.S. Liu. *Monte Carlo Strategies in Scientific Computing*. Springer Verlag, 2001.

[22] A. Makeev, D. Maroudas, and I. Kevrekidis. Coarse stability and bifurcation analysis using stochastic simulators : Kinetic Monte Carlo examples. *J. Chem. Phys.*, 116:10083–10091, 2002.

[23] H.G. Matthies, C.E. Brenner, C.G. Bucher, and C.G. Soares. Uncertainties in probabilistic numerical analysis of structures and solids - stochastic finite-elements. *Structural Safety*, 19(3):283–336, 1997.

[24] M. McKay, R. Beckman, and W. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979.

[25] R. T. Ogden. *Essentials Wavelets for Statistical Applications and Data Analysis*. Birkhäuser, 1997.

[26] M.T. Reagan, H.N. Najm, R.G. Ghanem, and O.M. Knio. Uncertainty quantification in reacting flow simulations through non-intrusive spectral projection. *Combustion and Flame*, 132:545–555, 2003.

[27] G. Strang. *Introduction to Applied Mathematics*. Wellesley-Cambridge Press, 1986.

[28] D. F. Walnut. *An Introduction to Wavelets Analysis*. Applied and Numerical Harmonic Analysis. Birkhäuser, 2002.

[29] G. G. Walter. *Wavelets and Other Orthogonal Systems with Applications*. CRC Press, 1994.

[30] S. Wiener. The Homogeneous Chaos. *Amer. J. Math.*, 60:897–936, 1938.

[31] D. Xiu and G.E. Karniadakis. Modeling uncertainty in steady state diffusion problems via generalized Polynomial Chaos. *Comput. Methods Appl. Mech. Engrg.*, 191:4927–4948, 2002.

[32] D.B. Xiu and G.E. Karniadakis. The Wiener-Askey Polynomial Chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24:619–644, 2002.

[33] D.B. Xiu and G.E. Karniadakis. Modeling uncertainty in flow simulations via generalized Polynomial Chaos. *Journal of Computational Physics*, 187:137–167, 2003.

A CONSTRUCTION OF THE ψ_j FUNCTIONS

We describe the methodology for the construction of the polynomial functions $\psi_i(x)$, $i = 0, \ldots, k-1$ satisfying Eqs. (10,11). To this end, the methodology

proposed by Alpert in [3] is briefly summarized in this appendix. The starting point is two sets of polynomial functions $p_i(x)$, and $\tilde{q}_j(x)$, defined for $x \in [0,1]$:

$$\begin{cases} p_i(x) = x^i & \text{for } i = 0, \ldots, k-1 \\ \\ \tilde{q}_i(x) = \begin{cases} p_i(x) & \text{if } x \geq 1/2 \\ -p_i(x) & \text{if } x < 1/2 \end{cases} & \text{for } i = 0, \ldots, k-1 \end{cases} \tag{A.1}$$

Step 1 :

In a first step, each function of the set $\{q_i(x), i = 0, \ldots, k-1\}$ is orthogonalized with respect to all the functions $p_i(x)$ for $i = 0, \ldots, k-1$. This is equivalent to the determination of the set of coefficients α_{ij} such that

$$q_j(x) = \tilde{q}_j(x) + \sum_{i=0}^{k-1} \alpha_{ij} p_i(x), \tag{A.2}$$

solves

$$< q_j(x) p_i(x) >= 0, \quad \text{for } i, j = 0, 1, \ldots, k-1. \tag{A.3}$$

At the end of this step, the functions $\{q_j(x), j = 0, \ldots, k-1\}$ have k vanishing moments (so satisfy Eq.(11)), but are not orthogonal : $\langle q_i q_j \rangle \neq 0$ for $i \neq j$.

Step 2 :

The orthogonalization of the set of functions $q_i(x)$ is enforced through the following procedure (Gram-Schmidt, see for instance [27]) :

(1) $r_{k-1}(x) = q_{k-1}(x)$.
(2) $j = k-1$
(3) $j \leftarrow j-1$.
(4) Orthogonalization of $q_j(x)$ with respect to the set of functions $\{r_{j+1}, \ldots, r_{k-1}\}$. This is achieved by determining the coefficients β_l so that

$$r_j(x) = q_j(x) + \sum_{l=j+1}^{k-1} \beta_l r_l(x), \tag{A.4}$$

satisfies $< r_j, r_l >= 0$ for $l = j+1, \ldots, k-1$.
(5) If $j > 0$ continue at 3, else terminate.

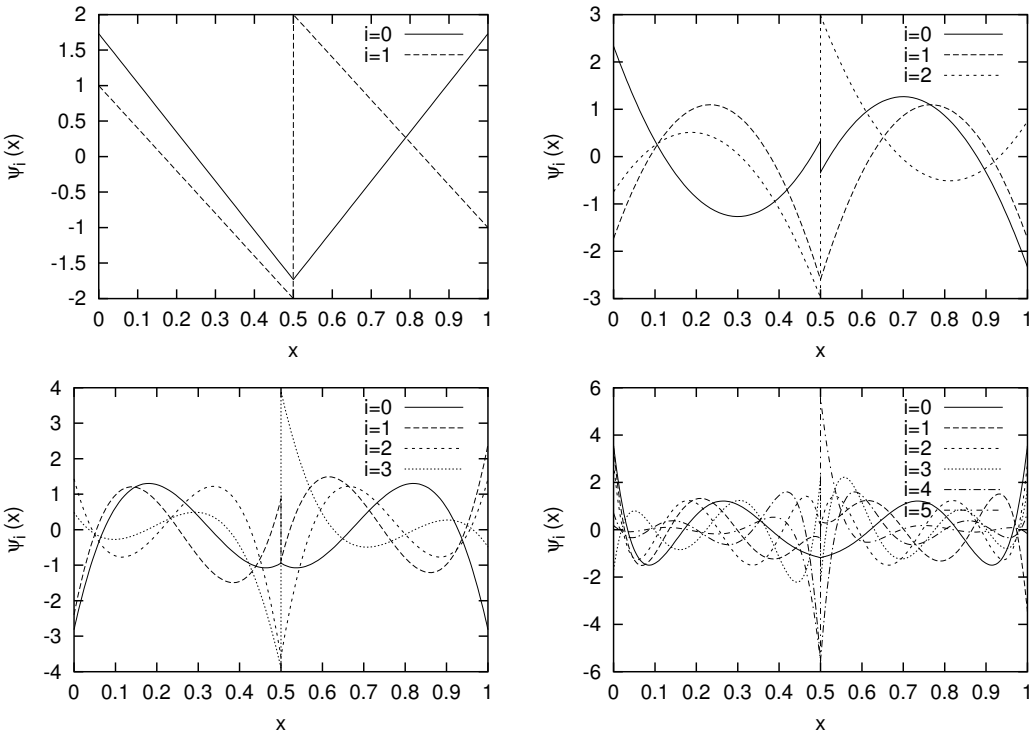

Fig. A.1. $\psi_{i=0,\ldots,No}(x)$ for expansion orders $No = 1$, 2, 3 and 5.

Since the polynomial functions r_j, $j = 0, \ldots, k-1$ are simply linear combinations of the $q_l(x)$ functions, they also have k vanishing moments but they are now mutually orthogonal.

Step 3 :

Finally, the desired functions $\psi_i(x)$ are obtained by normalizing the $r_i(x)$,

$$\psi_i(x) = \frac{r_i(x)}{\langle r_i, r_i \rangle}. \tag{A.5}$$

In this work, the computation of the ψ_i functions makes use of Gauss-Legendre quadrature rules (see [1]), allowing for exact estimate of the inner products $\langle f, g \rangle$ up to round-off errors which are negligible for moderate k. In Figure A.1, the functions $\psi_{j=0,\ldots,k-1}(x)$ are plotted for $k = 2, 3, 4$ and 6. Note that for $k = 1$, one obtains the Haar mother wavelet.