



Published in final edited form as:

*Comput Methods Appl Mech Eng.* 2017 February 1; 314: 196–221. doi:10.1016/j.cma.2016.09.024.

## A generalized multi-resolution expansion for uncertainty propagation with application to cardiovascular modeling

D.E. Schiavazzi<sup>a</sup>, A. Doostan<sup>b</sup>, G. Iaccarino<sup>c</sup>, and A.L. Marsden<sup>d</sup>

<sup>a</sup>Department of Applied and Computational Mathematics and Statistics, University of Notre Dame, IN 46556, USA

<sup>b</sup>Aerospace Engineering Sciences, University of Colorado Boulder, CO 80309, USA

<sup>c</sup>Department of Mechanical Engineering and ICME, Stanford University, CA 94305, USA

<sup>d</sup>Department of Pediatrics, Bioengineering and ICME, Stanford University, CA 94305, USA

### Abstract

Computational models are used in a variety of fields to improve our understanding of complex physical phenomena. Recently, the realism of model predictions has been greatly enhanced by transitioning from deterministic to stochastic frameworks, where the effects of the intrinsic variability in parameters, loads, constitutive properties, model geometry and other quantities can be more naturally included. A general stochastic system may be characterized by a large number of arbitrarily distributed and correlated random inputs, and a limited support response with sharp gradients or event discontinuities. This motivates continued research into novel adaptive algorithms for uncertainty propagation, particularly those handling high dimensional, arbitrarily distributed random inputs and non-smooth stochastic responses.

In this work, we generalize a previously proposed multi-resolution approach to uncertainty propagation to develop a method that improves computational efficiency, can handle arbitrarily distributed random inputs and non-smooth stochastic responses, and naturally facilitates adaptivity, i.e., the expansion coefficients encode information on solution refinement. Our approach relies on partitioning the stochastic space into elements that are subdivided along a single dimension, or, in other words, progressive refinements exhibiting a binary tree representation. We also show how these binary refinements are particularly effective in avoiding the exponential increase in the multi-resolution basis cardinality and significantly reduce the regression complexity for moderate to high dimensional random inputs. The performance of the approach is demonstrated through previously proposed uncertainty propagation benchmarks and stochastic multi-scale finite element simulations in cardiovascular flow.

### Keywords

Uncertainty quantification; Multi-resolution stochastic expansion; Cardiovascular simulation; Sparsity-promoting regression; Relevance vector machines; Multi-scale models for cardiovascular flow

## 1. Introduction

Computational models have become indispensable tools to improve our understanding of complex physical phenomena. Recent developments of these tools enable simulation of complex multi-physics systems at a cost that is, in many cases, negligible compared to setting up a physical experiment. Recent trends have led to a transition from deterministic to stochastic simulation approaches that better account for the intrinsic variability in parameters, material constant, geometry, and other input quantities. This improved approach boosts the predictive capability of models, allowing one to statistically characterize the outputs and therefore to quantify the confidence associated with the predictions.

While promising examples of this transition can be found in a variety of application fields, in this study we focus on the hemodynamic analysis of rigid or compliant vessels in the cardiovascular system. This is a rich application field where a stochastic solution requires multiple conceptual steps, from model reduction to data assimilation, and from design of effective parameterizations (e.g., for geometry, material properties, etc.) to efficient uncertainty propagation. In this study we focus on the propagation step, where the development of a general approach raises several challenges. First, our methods must handle arbitrary random inputs, from independent inputs with non-identical distribution, to correlated samples assimilated using Markov chain Monte Carlo (MCMC) from observations of the output quantities of interest. Second, there is a need for effective adaptive algorithms that are also easy to implement. Third, it is preferable to have the ability to reuse an existing library of model solutions. And finally, there is a need for an effective approach to reconstruct a stochastic response of interest requiring a minimal number of model evaluations.

Numerous approaches for uncertainty propagation have been proposed in the literature, many of which have been designed with specific applications in mind. As a comprehensive review of these approaches is outside the scope of this paper, here we mainly focus on methodologies supporting adaptivity, efficient reconstruction of sparse stochastic responses, and use of multi-resolution representations relevant to cardiovascular simulation.

The foundations for uncertainty propagation were laid in the first half of the last century [1,2] and re-proposed in [3] in the context of intrusive stochastic finite elements. An extension of Wiener chaos expansion to non-Gaussianly distributed random inputs, was introduced in [4] associating families of orthogonal polynomials in the Askey schemes with commonly used probability measures. An analysis of the convergence properties of this scheme is also proposed in [5]. Non intrusive stochastic collocation on tensor quadrature grids was formalized in [6] for random elliptic differential equations and extended in [7,8] to isotropic and anisotropic sparse tensor quadrature rules, respectively.

Adaptivity was introduced in [9] for non-intrusive uncertainty propagation schemes using a multi-element approach, while adaptive hierarchical sparse grids were proposed in [10]. An adaptive approach based on stochastic simplex collocation was proposed in [11] with the ability of supporting random input samples characterized by non regular domains. Use of sparsity-promoting approaches to identify the polynomial chaos coefficients was proposed in

[12] using standard  $\ell_2$  minimization, while a re-weighted  $\ell_1$  minimization strategy was proposed in [13]. Relevance vector machine regression in the context of adaptive uncertainty propagation was proposed in [14]. Use of multi-resolution expansion was first introduced in [15] for intrusive uncertainty propagation, and in [16] for the non-intrusive case. Finally, applications of stochastic collocation to cardiovascular simulation were proposed in [17] and in [18] in the context of robust optimization. Application to a human arterial tree with 37 parameters is also discussed in [19] using a sparse grid stochastic collocation method based on generalized polynomial chaos. Combination of data assimilation and uncertainty propagation was proposed in [20] in the context of virtual simulation of single ventricle palliation surgery.

In this paper we propose a generalized multi-resolution approach to uncertainty propagation, as an extension of the approach presented in [16]. In this multi-resolution approach, the range of the random inputs is partitioned into multiple *elements* where independent approximations of the local stochastic response are computed. These follow an expansion in an orthonormal multi-wavelet basis constructed with respect to the probability measure defined on each single element and generalize the approach in [21] which is limited to a uniform underlying measure. Expansion coefficients are computed using a Bayesian approach that has a number of advantages over previously proposed greedy heuristics for sparse regression. Based on the computed coefficients, element refinement is performed along one single dimension, leading to a significant reduction in the cardinality of multi-wavelet basis. The relevance of the proposed approach lies in its generality, ability to cope with steep gradients in the stochastic response, arbitrarily distributed random inputs and unstructured, e.g., random, solution samples.

The formulation of the uncertainty propagation problem for sPDEs is given in Section 2 and includes both cases of independent inputs and inputs sampled from the stationary joint posterior distribution through Markov chain Monte Carlo. A multi-resolution expansion is presented in Section 3, where a modification of the procedure in [21] is proposed to create a set of one-dimensional bases that are orthonormal with respect to an arbitrary distribution function, followed by the construction of a tensor product basis. Relevance vector machine regression is discussed in Section 4, while Section 5 presents the features of the proposed uncertainty propagation approach. Application to several benchmark problems is considered in Section 6 while Section 7 presents applications to patient-specific multi-scale cardiovascular computation. Finally, conclusions are presented in Section 8.

## 2. Problem formulation

Let  $(\Omega, \mathcal{F}, \mathcal{P})$  denote a complete probability space in which  $\Omega$  is the sample space,  $\mathcal{F}$  is the Borel  $\sigma$ -algebra of possible events, and  $\mathcal{P}$  is a probability measure on  $\mathcal{F}$ . We consider a vector of random inputs  $\mathbf{y} = (y_1, \dots, y_d)$ ,  $d > 0 \in \mathbb{N}$ ,  $y_i : \Omega \rightarrow \Sigma_{y_i} \subset \mathbb{R}$ , characterized by a joint probability density function (PDF)  $\rho(\mathbf{y})$ . Let  $\Gamma \subset \mathbb{R}^D$ ,  $D > 0 \in \mathbb{N}$ , be the spatial domain with boundary  $\Gamma$  and  $t \in [0, T]$  represents the temporal variable. We consider approximating the solution  $\mathbf{u}(\mathbf{x}, t, \mathbf{y}) : \Gamma \times [0, T] \times \Sigma_{\mathbf{y}} \rightarrow \mathbb{R}^h$ ,  $h > 0 \in \mathbb{N}$ , to the problem

$$\begin{aligned}
 \mathcal{L}(\mathbf{x}, t, \mathbf{y}; \mathbf{u}) &= \mathbf{f}(\mathbf{x}, t, \mathbf{y}) && \text{on } \Gamma \\
 \mathbf{u}(\mathbf{x}, t, \mathbf{y}) &= \mathbf{u}_b(\mathbf{x}, t, \mathbf{y}) && \text{on } \partial\Gamma \\
 \mathbf{u}(\mathbf{x}, t, \mathbf{y}) &= \mathbf{u}_0(\mathbf{x}, \mathbf{y}) && \text{at } t=0
 \end{aligned} \tag{1}$$

which holds  $\mathcal{P}$ -a.s. in  $\Omega$  for well posed (in  $\mathcal{P}$ -a.s. sense) forcing, boundary, and initial data  $\mathbf{f}$ ,  $\mathbf{u}_b$  and  $\mathbf{u}_0$ . We focus on approximating  $\mathbf{u}(\mathbf{x}, t, \mathbf{y})$  at a fixed location  $\mathbf{x} = \mathbf{x}^* \in \Gamma$  and time instant  $t = t^* \in [0, T]$  from  $M$  random realizations  $\{\mathbf{u}(\mathbf{x}^*, t^*, \mathbf{y}^{(i)}) : i = 1, \dots, M\}$ . In what follows, we drop the space–time dependence of  $\mathbf{u}$  and focus on estimating the  $p$ -statistical moment for the single-component response vector  $u(\mathbf{y}) : \mathbb{R}^d \rightarrow \mathbb{R}$ , i.e.,

$$\tilde{\nu}_p = \int_{\Omega} [u(\mathbf{y})]^p \rho(\mathbf{y}) d\mathbf{y}, \quad \tilde{\nu}_p = \int_{\Omega} [u(\mathbf{y}) - \tilde{\nu}_1]^p \rho(\mathbf{y}) d\mathbf{y}, \tag{2}$$

where order  $p > 0$  moments and centered moments are denoted by  $\tilde{\nu}_p$  and  $\tilde{\nu}_p$ , respectively.

Consider a truncated expansion of  $u(\mathbf{y})$  in terms of a family  $\psi_{\mathbf{i}}(\mathbf{y}) = \prod_{j=1}^d \psi_{i_j}(y_j)$ ,  $\mathbf{i} = (i_1, i_2, \dots, i_d) \in \mathcal{Q}$ , of multi-variate polynomials (or multi-wavelets, see Section 3), such that:

$$u(\mathbf{y}) \approx \sum_{\mathbf{i} \in \mathcal{Q}} \alpha_{\mathbf{i}} \psi_{\mathbf{i}}(\mathbf{y}), \tag{3}$$

where the cardinality of the basis is  $P = |\mathcal{Q}|$ , i.e., the *size* of the set  $\mathcal{Q}$ . Also, the family  $\psi_{\mathbf{i}}(\mathbf{y})$  is orthonormal to the product of the marginals  $\prod_{i=1}^d \rho_i(y_i)$ . In the next sections, we discuss how to compute the statistics of  $u(\mathbf{y})$  both for independent and dependent inputs  $\mathbf{y}$ .

### 2.1. Uncertainty Propagation Of Independent Inputs

If we assume the random inputs  $\mathbf{y}$  to be independent and characterized by a joint probability distribution  $\rho(\mathbf{y}) = \prod_{i=1}^d \rho_i(y_i)$ , the orthonormality of  $\psi_{\mathbf{i}}(\mathbf{y})$  assures a simple expression for the first two moments of  $u(\mathbf{y})$  in terms of the coefficients that can be computed from the regression of  $u(\mathbf{y})$ , i.e.,

$$\tilde{\nu}_1 = \alpha_0, \quad \text{and} \quad \nu_2 = \sum_{\mathbf{i} \in \mathcal{Q} / \{0\}} \alpha_{\mathbf{i}}^2. \tag{4}$$

### 2.2. Input samples from arbitrary joint distributions

In practical applications, random inputs are often not independent. Specifically, we consider the case where  $M$  samples  $\mathbf{y}^{(k)}$ ,  $k = 1, \dots, M$ , are generated from the joint distribution  $\rho(\mathbf{y})$  of  $\mathbf{y}$ , for example, through an MCMC-based solution of an inverse problem. In this case, constructing an orthonormal set  $\{\psi_{\mathbf{i}}(\mathbf{y}), \mathbf{i} \in \mathcal{Q}\}$  from tensor products of univariate basis functions is significantly more difficult. This is dealt with, in this study, by a change of measure in the expectations  $\tilde{\nu}_p$  or  $\nu_p$  to obtain simple formulas similar to (4) (for further reference see, e.g., [22,6]).

When the input vector  $\mathbf{y}$  has dependent components, we can multiply and divide the integrand of  $\tilde{\nu}_p$  by the product of the marginals  $\prod_{i=1}^d \rho_i(y_i)$  to obtain (see, e.g., [22]):

$$\tilde{\nu}_p = \int_{\Omega} \left[ u(\mathbf{y})^p \frac{\rho(\mathbf{y})}{\prod_{i=1}^d \rho_i(y_i)} \right] \prod_{i=1}^d \rho_i(y_i) d\mathbf{y} = \int_{\Omega} g(\mathbf{y}) \prod_{i=1}^d \rho_i(y_i) d\mathbf{y}, \quad \text{where } g(\mathbf{y}) = u(\mathbf{y})^p \frac{\rho(\mathbf{y})}{\prod_{i=1}^d \rho_i(y_i)}. \tag{5}$$

This means that the coefficients  $\{\alpha_{\mathbf{i}}, \mathbf{i} \in \mathcal{Q}\}$  that determine the stochastic expansion of  $g(\mathbf{y})$  according to the basis  $\psi_{\mathbf{i}}(\mathbf{y})$ , will encode the statistics of  $u(\mathbf{y})$ . Note also that regression of  $f(\mathbf{y})$  from  $\{u(\mathbf{y}^{(i)}) : i = 1, \dots, M\}$  requires the computation of  $\{\beta^{(k)} = \rho(\mathbf{y}^{(k)}) / \prod_{i=1}^d \rho_i(y_i^{(k)})\}$ ,  $k = 1, \dots, M$ . For random input samples generated through MCMC, the *stationary posterior*  $\tilde{\rho}(\mathbf{y}^{(k)})$  is known up to the integration constant  $c = 1 / \int_{\Sigma_{\mathbf{y}}} \tilde{\rho}(\mathbf{y}) d\mathbf{y}$ , or, in other words,  $c \tilde{\rho}(\mathbf{y}) = \rho(\mathbf{y})$ . Here we consider a family  $\psi_{\mathbf{i}}(\mathbf{y})$  of multi-variate functions orthonormal with respect to  $\rho(\mathbf{y}) = 1$  and expand the stationary posterior as:

$$\tilde{\rho}(\mathbf{y}) \approx \sum_{\mathbf{i} \in \mathcal{J}} \alpha_{\mathbf{i}} \psi_{\mathbf{i}}(\mathbf{y}), \quad \text{and therefore } c = \frac{1}{\alpha_0 |\Sigma_{\mathbf{y}}|}, \tag{6}$$

where  $|\Sigma_{\mathbf{y}}|$  denotes the *size* of  $\Sigma_{\mathbf{y}}$ . Using the expansion in (6), the generic marginal can be computed as:

$$\begin{aligned} \rho_i(\mathbf{y}_i) &= \alpha_0 |\Sigma_{\mathbf{y}}| \int_{\Sigma_{\mathbf{y}} / \Sigma_{y_i}} \tilde{\rho}(y_1, \dots, y_d) dy_1 \dots dy_{i-1} dy_{i+1} \dots dy_d \\ &= \alpha_0 |\Sigma_{\mathbf{y}}| \sum_{k=1}^K \int_{\Sigma_{\mathbf{y}} / \Sigma_{y_i}} \tilde{\rho}^k(y_1, \dots, y_d) dy_1 \dots dy_{i-1} dy_{i+1} \dots dy_d \\ &= \alpha_0 |\Sigma_{\mathbf{y}}| \sum_{k \in \mathcal{K}(y_i)} \sum_{\mathbf{j} \in \mathcal{J}} \alpha_{\mathbf{j}} \int_{\Sigma_{\mathbf{y}} / \Sigma_{y_i}} \psi_{\mathbf{j}}(y_1, \dots, y_d) dy_1 \dots dy_{i-1} dy_{i+1} \dots dy_d, \end{aligned} \tag{7}$$

where the set  $\mathcal{K}(y_i)$  contains the elements that include the  $y_i$  coordinate and  $\tilde{\rho}^k(\mathbf{y})$  is the restriction of the MCMC joint posterior distribution to the  $k$ th element. Due to the orthonormal property of the selected basis, it is easy to see that the above expression is not zero only when the multi-index  $\mathbf{j} = (j_1, \dots, j_d)$  belongs to the set  $\mathcal{Q}_i$  containing zeros in all components except  $i$ . This leads to

$$\rho_i(y_i) = \alpha_0 |\sum_{\mathbf{y}}| \sum_{k \in \mathcal{K}(y_i)} |\sum_{\mathbf{y}}^k / \sum_{y_i}^k| \sum_{\mathbf{j} \in \mathcal{Q}_i} \alpha_{\mathbf{j}} \psi_{\mathbf{j}}(y_i), \tag{8}$$

and therefore the coefficient  $\beta^{(l)}$ ,  $l = 1, \dots, M$  can be obtained as

$$\beta^{(l)} = \frac{\rho(\mathbf{y}^{(l)})}{\prod_{i=1}^d \rho_i(y_i^{(l)})} = \frac{\alpha_0 |\sum_{\mathbf{y}}| \sum_{\mathbf{j} \in \mathcal{Q}} \alpha_{\mathbf{j}} \psi_{\mathbf{j}}(\mathbf{y}^{(l)})}{\prod_{i=1}^d \alpha_0 |\sum_{\mathbf{y}}| \sum_{k \in \mathcal{K}(y_i^{(l)})} |\sum_{\mathbf{y}}^k / \sum_{y_i}^k| \sum_{\mathbf{j} \in \mathcal{Q}_i} \alpha_{\mathbf{j}} \psi_{\mathbf{j}}(y_i^{(l)})}. \tag{9}$$

In Section 6.1, we test the convergence of this approach for a variably correlated multi-variate Gaussian in two dimensions. Finally, we remark that the marginals could also be obtained from multi-wavelet regression of frequency plots from the MCMC parameter traces, possibly combined with smoothing through convolutions with the Gaussian density.

### 3. Multi-resolution stochastic expansion

In this section, we propose a generalized construction of Alpert multi-wavelets using arbitrary probability measures defined on a hypercube partition [21,16]. We start with the construction of a one-dimensional basis in the next section and discuss the multi-dimensional case in Section 3.2.

#### 3.1. Alpert's multi-wavelet in one dimension with arbitrary measure

Consider a random input  $y: \Omega \rightarrow (0, 1)$ , with distribution  $\rho(y)$ . The measure  $\rho(y)$  is assumed known at  $n_q$  quadrature locations  $\{\rho(y^{(i)}), i = 1, \dots, n_q\}$ . We propose a modification of the approach introduced in [21] to construct multi-wavelet basis with orthogonality with respect to the measure  $\rho(y)$ , instead of the originally assumed uniform measure. To do so, we first construct the multi-scaling functions  $\{\phi_i(y), i = 0, \dots, m - 1\}$  as the set of continuous polynomials on  $(0, 1)$ , orthogonal with respect to  $\rho(y)$ . Specifically, the orthogonality is defined as

$$\langle \phi_i, \phi_j \rangle_{L^2_{\rho}} = \int_{(0,1)} \phi_i(y) \phi_j(y) \rho(y) dy = 0, \forall i \neq j. \tag{10}$$

Legendre multi-scaling polynomials are obtained for an underlying uniform distribution  $\rho(y) = 1$  and the self-similarity of this measure over partitions of  $(0, 1)$  ensures that the same basis is generated across different elements, unlike the general case with non-uniform  $\rho(y)$  defined below.

The basis  $\{\phi_i, i = 0, \dots, m - 1\}$  satisfies the three term recurrence for orthogonal polynomials. If we define

$$\phi_{-1} \stackrel{\text{def}}{=} 0, \text{ and } \phi_0 \stackrel{\text{def}}{=} 1, \quad (11)$$

the generic multi-scaling function  $\phi_{i+1}$  obeys the relationship

$$\phi_{i+1}(y) = (y - \alpha_i)\phi_i(y) - \beta_i\phi_{i-1}(y) \text{ for } i = 0, \dots, m - 1, \quad (12)$$

with coefficients  $\alpha_i$  and  $\beta_i$  obtained through (see, e.g., [23])

$$\alpha_i = \frac{\langle \phi_i, y\phi_i \rangle_{L^2_\rho}}{\langle \phi_i, \phi_i \rangle_{L^2_\rho}} \text{ and } \beta_i = \frac{\langle \phi_i, \phi_i \rangle_{L^2_\rho}}{\langle \phi_{i-1}, \phi_{i-1} \rangle_{L^2_\rho}}, \quad (13)$$

respectively. In (13), the inner products are evaluated using a numerical quadrature rule as follows:

$$\langle f, g \rangle_{L^2_\rho} \approx \sum_{j=1}^{n_q} f(y^{(j)})g(y^{(j)})\rho(y^{(j)})w_j, \quad (14)$$

where  $w_j$  denotes the weights for the selected quadrature rule. Here we employ a *double* Clenshaw–Curtis quadrature rule [24] of order 30, using two Clenshaw–Curtis quadrature rules on  $(0, 1/2)$  and  $(1/2, 1)$ , respectively. Since one-dimensional multi-wavelets are piecewise polynomials defined over  $(0, 1/2)$  and  $(1/2, 1)$ , this choice leads to a greater accuracy than using a grid spanning the whole  $(0, 1)$  interval.

The basis  $\{\phi_i, i = 0, \dots, m - 1\}$  spans the space  $\mathbf{V}_0^m$  of continuous polynomials in  $L^2_\rho(0, 1)$  with order less than  $m$ . The subscript 0 here refers to the fact that these functions are defined over the full unit interval. Bases for  $\mathbf{W}_0^m$  are obtained, at this point, in a similar manner as discussed in [21] such that  $\mathbf{V}_0^m \perp \mathbf{W}_0^m$  and  $\mathbf{V}_0^m \oplus \mathbf{W}_0^m = \mathbf{V}_1^m$  with  $\mathbf{V}_1^m$  defined as the space of polynomials with order less than  $m$ , and continuous on  $(0, 1/2)$  and  $(1/2, 1)$ , respectively. The construction of bases for  $\mathbf{W}_0^m$  starts by selecting  $2m$  functions spanning the space of polynomials of degree less than  $m$  on the interval  $(-1, 0)$  and on  $(0, 1)$ , then orthogonalizing

$m$  of them, first to the functions  $1, y, y^2, \dots, y^{m-1}$ , then to the functions  $y^m, y^{m+1}, \dots, y^{2m-1}$ . Finally, we enforce mutual orthogonality, followed by rescaling on  $(0, 1)$  and normalization. These steps are described in more detail in the next sections, where the symbols  $\phi, \psi$  and  $f$  are used to denote multi-scaling, multi-wavelet and auxiliary functions generated at intermediate construction steps, respectively.

**3.1.1. Orthogonalization to monomials**—As a first step,  $m$  functions  $f_0^0, f_1^0, \dots, f_{m-1}^0$  are defined as

$$f_j^0(y) = \begin{cases} y^j, & y \in (0, 1) \\ -y^j, & y \in (-1, 0) \\ 0, & \text{otherwise} \end{cases} \quad (15)$$

which together with the monomials  $\{1, y, \dots, y^{m-1}\}$  span the space of polynomials of degree less than  $m$  on  $(0, 1)$  and  $(-1, 0)$ . We aim at creating a new family of functions  $\{f_j^1, j=0, \dots, m\}$ , each with  $m$  vanishing moments, i.e.:

$$\langle y^i, f_j^1 \rangle_{L^2_\rho} = \int_{(-1,1)} y^i f_j^1 \rho(y) dy = 0 \quad \forall 0 \leq i, j < m. \quad (16)$$

To do so, we write

$$f_i^1 = f_i^0 + \sum_{j=0}^{m-1} a_{i,j} y^j, \quad i=0, \dots, m-1, \quad (17)$$

and enforce the condition (16) by computing the coefficients  $a_{k,j}$  as the solution of a linear system whose  $k$ th equation is expressed as:

$$\sum_{j=0}^{m-1} \langle y^i, y^j \rangle_{L^2_\rho} a_{k,j} = \langle y^i, f_k^0 \rangle_{L^2_\rho}, \quad i=0, \dots, m-1 \quad (18)$$

where all products  $\langle \cdot, \cdot \rangle_{L^2_\rho}$  in (18) are evaluated using numerical quadrature. As suggested in [21], this orthogonality is preserved by the following steps, which only produce linear combinations of the  $f_j^1$ . The coefficients  $a_{k,j}$  in (18) are stored in  $\mathbf{A} \in \mathbb{R}^{m \times m}$ , such that  $(\mathbf{A})_{k,j} = a_{k,j}$ . Consider a generic vector  $\mathbf{f}_i^0(\mathbf{y}^{(j)}) = (f_i^0(y_1^{(j)}), \dots, f_i^0(y_{n_q}^{(j)}))$  containing realizations of the function  $f_i^0$  at a number  $n_q$  of locations, stored in the vector  $\mathbf{y}^{(j)} \in \mathbb{R}^{n_q}$ . Also consider the matrix  $\mathbf{F}^0$  with columns  $\mathbf{f}_i^0, i=0, \dots, m-1$ , and  $\mathbf{Y}$  containing columns  $(\mathbf{y}^{(j)})^i, i=0, \dots, m$



$- 1, j = 1 \dots, n_q$ . Using (17), the set of functions  $f_i^1$  at  $\mathbf{y}^{(j)}$  are evaluated using the following expression:

$$\mathbf{F}^1 = \mathbf{F}^0 + \mathbf{A} \mathbf{Y}, \quad (19)$$

where the matrix  $\mathbf{F}^1$  has columns  $\mathbf{f}_i^1, i = 0 \dots, m - 1$ .

**3.1.2. Additional orthogonalizations to higher order monomials**—The next sequence of steps involves constructing  $m - 1$  functions orthogonal to  $y^m$ , of which  $m - 2$  functions are orthogonal to  $y^{m+1}$ , and so forth, down to 1 function which is orthogonal to  $y^{2m-1}$  [21]. To this aim, we first reorder  $f_j^1$  such that the first function is not orthogonal to  $y^m$ , i.e.,  $\langle f_0^1, y^m \rangle \neq 0$ . We then define  $f_j^2 = f_j^1 - b_j^2 f_0^1$  where  $b_j^2$  is chosen so that  $\langle f_j^2, y^m \rangle_{L_p^2} = 0, j = 1, \dots, m - 1, j = 1, \dots, m - 1$ , achieving the desired orthogonality. The  $m - 1$  coefficients  $b_j^2$  are thus obtained as the solution of the equation:

$$b_j^2 = \frac{\langle f_j^1, y^m \rangle_{L_p^2}}{\langle f_0^1, y^m \rangle_{L_p^2}}, j = 1, \dots, m - 1. \quad (20)$$

Similarly, we write  $f_j^3 = f_j^2 - b_j^3 f_1^2, j = 2, \dots, m - 1$  and determine the  $b_j^3$  using the condition  $\langle f_j^3, y^{m+1} \rangle_{L_p^2} = 0, j = 2, \dots, m - 1$ , leading to

$$b_j^3 = \frac{\langle f_j^2, y^{m+1} \rangle_{L_p^2}}{\langle f_1^2, y^{m+1} \rangle_{L_p^2}} = \frac{\langle f_j^1, y^{m+1} \rangle_{L_p^2}}{\langle f_j^2, y^{m+1} \rangle_{L_p^2}} - a_j \frac{\langle f_0^1, y^{m+1} \rangle_{L_p^2}}{\langle f_1^2, y^{m+1} \rangle_{L_p^2}}, j = 2, \dots, m - 1. \quad (21)$$

The procedure is repeated with orthogonalizations with respect to  $y^{m+2}, \dots, y^{2m-1}$ , to obtain the set  $\{f_0^1, f_1^2, f_2^3, \dots, f_{m-1}^m\}$  satisfying the conditions

$$\langle f_j^{j+1}, y^i \rangle_{L_p^2} = 0, \forall i \leq j + m - 1, \forall i \quad j + m - 1.$$

For a more compact notation, we consider the matrix  $\mathbf{B} \in \mathbb{R}^{(m-1) \times (m-1)}$  with components

$$(\mathbf{B})_{i,j} = \begin{cases} b_{i+j}^{i+1} = \frac{\langle f_{i+j}^i, y^{m+j-1} \rangle_{L^2_\rho}}{\langle f_{i-1}^i, y^{m+j-1} \rangle_{L^2_\rho}}, & j < m - i, \\ 0, & \text{otherwise,} \end{cases} \quad (22)$$

where reordering assures that the denominator  $\langle f_{i-1}^i, y^{m+j-1} \rangle_{L^2_\rho}$  does not vanish. We then recursively apply

$$f_j^i = f_j^{i-1} - \sum_{k=1}^{m-1} (\mathbf{B})_{i,k} f_{i+k}^{i-1}. \quad (23)$$

**3.1.3. Mutual orthogonality, normalization and rescaling**—Orthogonalization of the functions  $f_i^{i+1}$ ,  $i = 0, \dots, m - 1$ , is accomplished, at this point, by using the Gram–Schmidt procedure. The first function is assigned as  $\varphi_0 = f_{m-1}^m$ , while the successive ones are computed through

$$\varphi_i = f_{m-i-1}^{m-i} + \sum_{j=0}^{i-1} h_{i,j} \varphi_j, \quad (24)$$

where the coefficients  $h_{i,j}$  are determined from the condition  $\langle \varphi_i, \varphi_k \rangle_{L^2_\rho} = 0$ , resulting in the system

$$\sum_{j=0}^{i-1} \langle \varphi_j, \varphi_k \rangle_{L^2_\rho} h_{i,j} = - \langle f_{m-i-1}^{m-i}, \varphi_k \rangle_{L^2_\rho}, \quad \forall k < i. \quad (25)$$

The solution of (25) is stored in the matrix  $(\mathbf{H})_{i,j} = h_{i,j}$  and used every time  $\varphi_i(y^{(i)})$  needs to be evaluated. We finally normalize and rescale from  $(-1, 1)$  to  $(0, 1)$  to obtain the functions  $\{\varphi_i, i = 0, \dots, m - 1\}$  that satisfy  $\langle \varphi_i, \varphi_j \rangle_{L^2_\rho(0,1)} = \delta_{i,j}$ , and  $\delta_{i,j}$  is the Kronecker delta.

**Remark 1 (Re-evaluating Multi-wavelet Basis).** In order to reduce the computational burden of evaluating multi-wavelet functions, the systems (18) and (25) are solved only once and the matrices  $\mathbf{A}$ ,  $\mathbf{B}$  and  $\mathbf{H}$  stored. Every successive evaluation of the bases uses (17), (23) and (24).

**3.1.4. Properties of the basis**—From the construction carried out in the previous sections, it is readily seen that the functions  $\{\varphi_i(y) : (0, 1) \rightarrow \mathbb{R}, i = 0, \dots, m\}$ , have the following properties [21]:

1. The restriction of  $\varphi_i$  to the interval  $(1/2, 1)$  is a polynomial of degree  $m - 1$ . This property is extremely important when constructing multi-wavelet basis in multiple dimensions as it prevents sparse tensorizations from being used as discussed in more detail in Section 3.2.
2. The function  $\varphi_i$  is extended to the interval  $(0, 1/2)$  as an even or odd function according to the parity of  $i+m - 1$  (see Fig. 1).
3. The family  $\{\varphi_i, i = 0, \dots, m - 1\}$  satisfies, by construction, the orthogonality conditions

$$\int_{(0,1)} \varphi_i(y) \varphi_j(y) \rho(y) dy = \delta_{i,j}, \quad i, j = 0, \dots, m - 1. \quad (26)$$

1. The function  $\varphi_j$  has the following vanishing moments

$$\int_{(0,1)} \varphi_j(y) y^i \rho(y) dy = 0, \quad i = 0, 1, \dots, j+m - 1. \quad (27)$$

As discussed in [21], properties 1 and 2 imply that there are  $(m - 1)^2$  polynomial coefficients that determine the functions  $\{\varphi_0, \varphi_1, \dots, \varphi_{m-1}\}$ , while properties 3 and 4 provide  $(m - 1)^2$  constraints. Therefore, the multi-wavelet basis functions are unique (up to sign). Examples of the proposed multi-resolution basis for an underlying uniform and truncated Gaussian measure are shown in Fig. 1.

**Remark 2 (Basis at Finer Resolutions).** After constructing the multi-scaling and multi-wavelet bases spanning  $\mathbf{V}_0^m$  and  $\mathbf{W}_0^m$ , respectively, the construction of a basis for  $\mathbf{W}_1^m$  through scaling and shifting operations is not at all trivial if  $\rho(y)$  is not constant. Specifically, it appears that the *self-similarity* of the underlying measure on the hypercube partitions is a key property to build a basis for higher resolution in the arbitrary measure case. In this work, we focus only on building a basis for  $\mathbf{W}_0^m$  and use the multi-wavelet coefficients to guide the refinement of current elements. Once refined, a new set of bases for  $\mathbf{V}_0^m$  and  $\mathbf{W}_0^m$  is re-computed independently of the basis used on the parent element.

### 3.2. Multi-dimensional multi-wavelets for binary refinements

For  $d > 1$ , we consider the vector  $\mathbf{m} = (m_1, \dots, m_d) \in \mathbb{N}^d$  and introduce

$\mathbf{V}_0^{\mathbf{m}} = \mathbf{V}_0^{m_1} \otimes \dots \otimes \mathbf{V}_0^{m_d}$  as the product space spanned by the tensorizations of univariate multi-scaling functions. In this setting, the multi-index  $\mathbf{i} \in \mathcal{Q} = \{(i_1, \dots, i_d) : 0 \leq i_l < m_l, l = 1, \dots, d\}$ , is used to define the  $d$ -dimensional tensor product multi-scaling basis

$$\phi_{\mathbf{i}}(\mathbf{y}) = \phi_{i_1}(y_1) \dots \phi_{i_d}(y_d).$$

To unify the notation, we define

$$\psi_i^0 = \phi_i, \quad \psi_i^1 = \varphi_i, \quad \psi_{\mathbf{i}}^{\mathbf{q}}(\mathbf{y}) = \psi_{i_1}^{q_1}(y_1) \dots \psi_{i_d}^{q_d}(y_d),$$

where  $\mathbf{q}$  is a multi-index set in  $Q = \{(q_1, \dots, q_d) : q_l \in \{0, 1\}, l = 1, \dots, d\}$ . In other words,  $\mathbf{q}$  defines the products of univariate multi-scaling and multi-wavelet functions  $\phi_i$  and  $\varphi_i$  in  $\psi_{\mathbf{i}}^{\mathbf{q}}(\mathbf{y})$ , where  $\mathbf{q} = (0, \dots, 0)$  identifies the set of multi-scaling basis.

We now focus on multi-indices  $\mathbf{q}$  that are compatible with *binary refinements* of an element, i.e., subdivisions along a single dimension. As the presence of a multi-wavelet basis along multiple dimensions would imply a corresponding ability to perform subdivisions, *only a single multi-wavelet function per dimension* is needed in this case. This also reduces the number of possible multi-indices  $\mathbf{q}$  from  $2^d$  to  $d$ , and therefore the complexity of forming a multi-resolution basis under increasing dimensionality. Under these conditions, a stochastic response  $u(\mathbf{y}) \in L^2_{\rho}(0, 1)^d$  can be expanded as

$$u(\mathbf{y}) \approx \sum_{\mathbf{i} \in I} \tilde{\alpha}_{\mathbf{i}} \psi_{\mathbf{i}}^0(\mathbf{y}) + \sum_{\mathbf{q} \in \mathcal{Q} \setminus \{(0, \dots, 0)\}} \sum_{\mathbf{i} \in I} \tilde{\beta}_{\mathbf{i}} \psi_{\mathbf{i}}^{\mathbf{q}}(\mathbf{y}). \quad (28)$$

The cardinality  $P$  of the proposed multi-variate representation is also significantly reduced compared to the approach proposed in [16], due to the freedom in selecting the multi-index  $\mathbf{i}$ . It should be noted, in this regard, that multi-wavelet functions are polynomials of degree  $m - 1$  on  $(0, 1/2)$  and  $(1/2, 1)$ , respectively. Hence, partial ten-sorization of multi-wavelets in different dimensions leads to an incomplete representation and a full tensor product must be employed. For binary refinements instead, a maximum of one component in each multi-index is associated to a multi-wavelet function. This enables one to employ sparse multi-variate tensorizations of one-dimensional multi-wavelets, for example, with multi-indices of total order less than  $m$ , or less expensive alternatives (see, e.g., hyperbolic cross multi-indices [25]). Finally, we simplify the notation by rewriting (28) in the form

$$u(\mathbf{y}) \approx \sum_{\mathbf{i} \in \mathcal{I}} \alpha_{\mathbf{i}} \psi_{\mathbf{i}}(\mathbf{y}). \quad (29)$$

Examples of times and storage requirements to build the multi-wavelet matrix ( $\Psi$ ) with components  $(\Psi)_{i,j} = \psi_{\mathbf{j}}(\mathbf{y}^{(i)})$  (note that here we assumed  $\mathbf{j}$  to be the  $j$ th multi-index in  $\mathcal{I}$ ) are reported in Table 1.

#### 4. Relevance vector machine (RVM) regression

After generating a basis with the desired properties, the expansion coefficients  $\alpha_{\mathbf{i}}, i \in \mathcal{I}$ , in (29) need to be evaluated. In this study, we use Relevance Vector Machines (RVM), a sparsity promoting Bayesian regression framework introduced in [26,27]. Note that, in this study, we consider a Gaussian hyperprior on the coefficients  $\alpha_{\mathbf{i}}$ , while a Gamma hyperprior

was employed in [26]. Consider, in this context, a generic  $d$ -dimensional random vector  $\mathbf{y} \in \mathbb{R}^d$  and denote a set of  $M$  realizations as  $\mathbf{y}^{(i)}$ ,  $i = 1, \dots, M$ . Also assume the stochastic response samples  $\mathbf{u} = [u(\mathbf{y}^{(1)}), u(\mathbf{y}^{(2)}), \dots, u(\mathbf{y}^{(M)})]$  to be related with realizations of  $\mathbf{y}$  through an expression of the form:

$$\mathbf{u} = \Psi \boldsymbol{\alpha} + \boldsymbol{\varepsilon} \quad (30)$$

where  $\Psi \in \mathbb{R}^{M \times P}$  contains the realizations of our multi-resolution basis in matrix form,  $\boldsymbol{\alpha} \in \mathbb{R}^P$  is the expansion coefficient vector and  $\boldsymbol{\varepsilon} \in \mathbb{R}^M$  is a Gaussian error vector with zero-mean components and diagonal covariance  $\sigma^2 \mathbf{I}_M$ . Before proceeding, we stress that, in most applications, the coefficients are expected to be sparse, i.e.,  $\|\boldsymbol{\alpha}\|_0 = \#\{\alpha_i : \alpha_i \neq 0, i \in \mathcal{Q}\} \ll P$ , due to the increasing smoothness of the stochastic response for progressive element refinements. This means that only a few  $\psi_i(\mathbf{y})$  are expected to be characterized by a non-zero expansion coefficient  $\alpha_i$  in (29) or, as will be alternatively referred to, are included in the *model*  $\boldsymbol{\varepsilon} = \{\mathbf{i} : \alpha_i \neq 0\}$ , i.e., the set of multi-indices associated to non-zero expansion coefficients. This motivates the adoption of a framework for sparse regression. The likelihood induced by (30) is thus expressed as:

$$P(\mathbf{u}|\boldsymbol{\alpha}, \sigma^2) = \frac{1}{(2\pi)^{M/2} \sigma^M} \exp \left[ -\frac{1}{2} \frac{(\mathbf{u} - \Psi \boldsymbol{\alpha})^T (\mathbf{u} - \Psi \boldsymbol{\alpha})}{\sigma^2} \right]. \quad (31)$$

Prior knowledge on  $\boldsymbol{\alpha}$  is also introduced using independent zero-mean Gaussian hyperpriors with parameters  $\beta_i$ ,  $i = 0, \dots, P-1$ . Note that  $\beta_i$  are *precisions*, i.e., the inverse variances of the coefficients  $\alpha_i$ , and that a zero mean Gaussian hyperprior provides a natural mechanism to promote sparse representations. We express these hyperpriors as

$$P(\boldsymbol{\alpha}|\boldsymbol{\beta}) = \frac{1}{(2\pi)^{P/2} \beta^P} \exp \left[ -\frac{\boldsymbol{\alpha}^T \mathbf{B} \boldsymbol{\alpha}}{2} \right], \quad (32)$$

where the diagonal matrix  $\mathbf{B}$  is such that  $(\mathbf{B})_{i,i} = \beta_i$ . Due to the adoption of a conjugate prior, the posterior density of the coefficients  $\boldsymbol{\alpha}$  is also Gaussian, with mean and covariance, respectively, equal to (see [26]):

$$\boldsymbol{\alpha} = \sigma^{-2} \sum \Psi^T \mathbf{u}, \quad \Sigma = (\mathbf{B} + \sigma^{-2} \Psi^T \Psi)^{-1}. \quad (33)$$

The estimates (33) are iteratively updated by computing the hyperparameters  $\boldsymbol{\beta}$  from the maximization of the marginal log-likelihood:

$$\mathcal{M}(\boldsymbol{\beta}) = \log p(\mathbf{u}|\boldsymbol{\beta}, \sigma^2) = \log \int_{-\infty}^{+\infty} p(\mathbf{u}|\tilde{\boldsymbol{\alpha}}, \sigma^2) p(\tilde{\boldsymbol{\alpha}}|\boldsymbol{\beta}) d\tilde{\boldsymbol{\alpha}} = \mathcal{M}(\boldsymbol{\beta}_{-r}) + \Delta M(\beta_r), \quad (34)$$

where  $\tilde{\boldsymbol{\alpha}}$  denotes an iterate of the expansion coefficients. The increment  $\Delta M(\beta_r)$  produced by assigning a non zero  $\alpha_r$  to the  $r$ th basis function can be explicitly computed as

$$\Delta \mathcal{M}(\beta_r) \propto \log \beta_r - \log(\beta_r + s_r) + \frac{q_r^2}{\beta_r + s_r}, \quad (35)$$

where  $q_r$  and  $s_r$  are the quality and sparsity factors, respectively (the interested reader is referred to [26,27] for a derivation of the intermediate steps). The quality factor  $q_r$  provides a measure of the correlation between the basis vector  $\Psi_r$  and the residual  $\mathbf{r} = \mathbf{u} - \Psi \tilde{\boldsymbol{\alpha}}$ . The sparsity factor instead quantifies how a given basis vector is similar to another basis already in  $\mathcal{E}$ . Due to the importance of these two factors, the sign of  $\theta_r = q_r^2 - s_r$  determines if a basis function needs to be *added* or *removed* from  $\mathcal{E}$ , or if the associated coefficient needs to be *re-estimated*.

In a previous study [16], greedy heuristics for sparse regression – e.g., Orthogonal Matching Pursuit (OMP) and Tree-based Orthogonal Matching Pursuit (TOMP) – were proposed in conjunction with a monolithic multi-resolution expansion approach, i.e., an approach where a single multi-resolution basis was adopted to approximate the stochastic response  $u(\mathbf{y})$ . We found that RVM offered improvements over OMP and TOMP in three different areas.

The first is flexibility in adding and removing basis functions from the model  $\mathcal{E}$ . During the OMP iterations, an arbitrary basis function  $\Psi_i$  can only be added to  $\mathcal{E}$ , based on its correlation with the residual vector at the current iteration. In RVM, the increment in the marginal likelihood (35) suggests adding a basis function to  $\mathcal{E}$ , removing it, but also to re-estimating the associated coefficient  $\alpha_i$ , if it already belongs to  $\mathcal{E}$ . Second, both OMP/TOMP consider a fixed noise variance. Estimation of the noise level producing the minimum regression residual can be performed at a significant increase in the computational cost, for example through cross validation techniques. RMVs naturally deal with this problem by alternating estimations of the coefficient vector  $\boldsymbol{\alpha}$  and its covariance  $\Sigma$  with re-estimation of the noise intensity. Finally, RMV has the ability of directly estimating the full covariance matrix  $\Sigma$  for the coefficient vector  $\boldsymbol{\alpha}$ , unlike OMP/TOMP. This is useful to assess the confidence in the estimate of  $\boldsymbol{\alpha}$ .

Before discussing the application of RVM regression in the context of uncertainty propagation, we note that this approach promotes a sparse representation and is effective in approximating the stochastic response in progressively small parameter space elements. Unlike *greedy heuristics* for sparse regression, where the non zero coefficients are computed for the basis mostly correlated with the regression residual, or in *basis pursuit* where a relaxed  $l_1$  convex regularizer is employed, here sparsity in the resulting representation is promoted by specifying a zero-mean hyperprior for every expansion coefficient, consistent

with the Bayesian character of the approach [26,27]. Specifically, it can be seen from (33), that a larger value of the  $i$ th precision hyperparameter  $\beta_i$  will cause the associated expansion coefficient  $\alpha_i$  to be concentrated at zero thus *switching off* its contribution to the model [26]. It is also evident how approximations on progressively smaller partitions are characterized by an increasing smoothness in the local stochastic response. This means that elements not containing stochastic discontinuities will likely be accurately represented using only a small number of expansion coefficients.

## 5. Adaptive multi-resolution uncertainty propagation

This section presents the main features of an uncertainty propagation approach that includes the multi-resolution stochastic expansion and sparse regression methods discussed in Sections 3 and 4. In particular, we describe the features that differentiate it from previous studies in multi-resolution uncertainty propagation (see, e.g., [16]).

We believe the generality of the proposed approach to be well suited for applications in cardiovascular blood flow simulation. This includes situations where some of the parameters (e.g., patient specific boundary conditions) are estimated from model output observations through the solution of an inverse problem, whereas other, independent, parameters may be added to study the sensitivity to changes in the anatomical model geometry, material properties or other model parameters. In such a case, both the ability to support random sampling from arbitrary underlying measures and the built-in adaptivity allowing for approximations of stochastic responses possibly characterized by steep gradients are key to an efficient uncertainty propagation method.

### 5.1. Multi-element approach

Regression of a stochastic response on  $(0, 1)^d$  using a single multi-wavelet expansion (monolithic approach) has several drawbacks. First, the cardinality of the basis dictionary will exponentially increase, and the regression problem rapidly becomes intractable even for a moderate number of random inputs. Second, there is no easy way to distribute the computations across processors when running on parallel architectures, as the regression needs to be computed on the whole domain even if large parts of this domain are characterized by a zero multi-wavelet coefficient for a number of successive resolutions.

These difficulties can be overcome by computing multiple regressions on the separate elements. Likely, this will lead to the stochastic response to be progressively smoother on elements of decreasing size, providing a stronger justification for the use of sparse regression. Moreover, this needs to be performed only on elements where the approximation has not yet converged, thus better allocating the available computational resources.

### 5.2. Built-in adaptivity via analysis of variance

This section shows how the information encoded in the multi-resolution coefficients  $\alpha$  can be leveraged to assess whether the current element needs to be subdivided. Consider a Sobol' decomposition [28] of the stochastic response  $u(\mathbf{y})$  with independent inputs  $\mathbf{y}$ , in terms of zero mean contributions of increasing dimensionality

$$u(y_1, \dots, y_d) = u_0 + \sum_{j=1}^d u_j(y_j) + \sum_{j=1}^{d-1} \sum_{k=j+1}^d u_{j,k}(y_j, y_k) + \dots + u_{1,\dots,d}(y_1, \dots, y_d), \text{ where} \\ \int_{(0,1)} u_{j_1,\dots,j_l}(y_{j_1}, \dots, y_{j_l}) \rho_{j_k}(y_{j_k}) dy_{j_k} = 0, \forall 1 \leq k \leq l, \forall \{j_1, \dots, j_l\} \subset \{1, \dots, d\}. \quad (36)$$

Under these assumptions,  $u_0$  represents the mean  $\tilde{v}_1$  of  $u$ , while its variance becomes

$$v_2 = \int_{(0,1)^d} u(\mathbf{y})^2 \rho(\mathbf{y}) d\mathbf{y} - u_0^2 = \sum_{j=1}^d \sigma_j^2 + \sum_{j=1}^{d-1} \sum_{k=j+1}^d \sigma_{j,k}^2 + \dots + \sigma_{1,\dots,d}^2, \quad (37)$$

as products of terms in (36) are all zero, and where

$$\sigma_{j_1,\dots,j_l}^2 = \int_{(0,1)^l} u(y_{j_1}, \dots, y_{j_l})^2 \rho_{j_1}(y_{j_1}) \dots \rho_{j_l}(y_{j_l}) dy_{j_1} \dots dy_{j_l}. \quad (38)$$

This leads to a decomposition of the total variance of  $u$  in single parameter effects, combined effects of two parameters, three parameters and so on. The combined global effect of the parameter set  $\{y_{j_1}, \dots, y_{j_l}\}$  can therefore be quantified using the following *direct sensitivity index*

$$S_{y_{j_1}, \dots, y_{j_l}} = \frac{\sigma_{j_1, \dots, j_l}^2}{\sigma^2}, \text{ where, } \sum_{j=1}^d S_j + \sum_{j=1}^{d-1} \sum_{k=j+1}^d S_{j,k} + \dots + S_{1,\dots,d} = 1. \quad (39)$$

At this point, we note that the orthonormal multi-resolution expansion in (29) represents a Sobol' decomposition of  $u$ . In this context, the following relationship between multi-resolution expansion coefficients and contributions to the total variance can be expressed as

$$\alpha_i^2 = \sigma_i^2. \quad (40)$$

It is therefore possible to separate the contributions to the total variance associated with the multi-scaling and multi-wavelet family as follows

$$S_{MS} = \frac{\sum_{i \in I_{MS}} \alpha_i^2}{\sum_{j \in I} \alpha_j^2} = \frac{\sum_{i \in I_{MS}} \alpha_i^2}{\sigma^2}, \quad (41)$$



$$S_{MW} = \frac{\sum_{\mathbf{i} \in \mathcal{I}_{MW}} \alpha_{\mathbf{i}}^2}{\sigma^2}, \quad (42)$$

$$\gamma = S_{MW} \cdot \sigma^2, \quad (43)$$

where the sets  $\mathcal{I}_{MS}$ , ( $\mathcal{I}_{MW}$  contain basis function multi-indices associated with members of the multi-scaling and multi-wavelet families, respectively. It is easy to see that  $S_{MW}$  quantifies how much of the total variance of the stochastic response  $u(\mathbf{y})$  in the current partition is explained by differences between levels of refinements (contribution of multi-wavelet basis). In the examples presented in Section 6, we use  $\gamma$  as a *refinement metric*, subdividing an element if  $\gamma > \gamma_t$  where  $\gamma_t$  is a selected threshold in the range  $[10^{-8}, 10^{-6}]$ .

Moreover, as previously discussed, elements are subdivided only along a single dimension. Thus, the algorithm needs to compute the refinement direction when  $\gamma > \gamma_t$ . This can be achieved by looking at the multi-index  $\mathbf{i} = (i_1, i_2, \dots, i_d)$  associated with  $\alpha_{\mathbf{i}}$ , where each index  $i_j, j = 1, \dots, d$ , denotes the order of a univariate multi-scaling or multi-wavelet basis along the  $j$ th dimension. Consider the set  $\mathcal{I}_j$  containing multi-indices that are non zero only along dimension  $j$ . We define the following variances

$$\sigma_{MW,j}^2 = \sum_{\mathbf{i} \in \mathcal{I}_j} \alpha_{\mathbf{i}}^2, \text{ and } j_{ref} = \arg \max_j \sigma_{MW,j}^2, \quad (44)$$

where  $j_{ref}$  denotes the dimension along which the refinement is performed.

For the sake of clarity, we remark that, in this section, an analysis of the variance of  $u(\mathbf{y})$  was leveraged to decide whether an element should be further subdivided. In particular, we showed how the total multi-wavelet variance can be obtained from the multi-resolution expansion coefficients and also how to identify which dimension needs refinement. Similar refinement metrics are discussed in [9,15].

### 5.3. Binary tree refinement

The iterative partitions of the unitary hypercube are stored using a binary tree data structure. The *leaves* of this representation (i.e., elements in the current partition) are visited at every iteration, and a multi-resolution approximation of the stochastic response in these elements is computed if none of the convergence criteria has been met.

After processing all elements in the partition, single element statistics need to be assembled into statistics of the overall stochastic response. If we define  $|\sum_{\mathbf{y}}^k| = \int \sum_{\mathbf{y}}^k \rho^k(\mathbf{y}) d\mathbf{y}$  as the

*volume* of the  $k$ th element, the overall stochastic response average  $\tilde{\nu}_1$  is assembled from the single element averages  $\tilde{\nu}_1^k, k=1, \dots, K$ , as follows

$$\tilde{\nu}_1 = \frac{1}{|\sum_{\mathbf{y}}|} \sum_{k=1}^K \tilde{\nu}_1^k |\sum_{\mathbf{y}}^k|, \quad (45)$$

while the overall centered variance  $\nu_2$  is evaluated using the expression

$$\nu_2 = \frac{1}{|\sum_{\mathbf{y}}|} \sum_{k=1}^K \left[ (v_2^k)^2 + (\tilde{\nu}_1^k - \tilde{\nu}_1)^2 \right] |\sum_{\mathbf{y}}^k|. \quad (46)$$

#### 5.4. Multi-resolution regression and random sampling

For a smooth stochastic response, collocation on quadrature grids of increasing order will likely lead to an exact computation of the expansion coefficients, thus enhancing the effectiveness of the adaptivity indicators encoded in the multi-wavelet basis. Conversely, for randomly distributed parameter realizations  $\mathbf{y}^{(i)}, i = 1, \dots, M$ , regression errors in  $\{a_i, i \in \mathcal{Q}\}$  may negatively affect the estimated adaptivity metrics. This is especially true when the number of samples contained in a single element is progressively reduced as a result of refinement. This may happen, for example, when a fixed number of solution realizations are available and no more can be generated.

Quantification of how well smooth (e.g., constant or linear) responses can be captured from their random samples is therefore useful to quantify the amount of oversampling in more general scenarios. Specifically, we look at the ability of the proposed multi-resolution approach to approximate two multi-variate functions  $u_1$  and  $u_2$  of the form

$$u_1(\mathbf{y}) = 0.3, \quad u_2(\mathbf{y}) = \sum_{i=1}^d y_i. \quad (47)$$

These functions can be represented exactly only in terms of multi-scaling basis coefficients, and therefore with identically zero multi-wavelet variance. Note that the minimum number of samples needed to uniquely identify  $u_1$  and  $u_2$  is 1 and  $d+1$ , respectively. The decay of the multi-wavelet variance shown in Fig. 2 is informative on the amount of oversampling required by the proposed approach under a variable number of dimensions  $d$ .

As shown in Fig. 2, a constant function is always correctly represented with very few samples. This means that sub-domains containing only one sample will be correctly approximated with a constant function by MW regression and the variance of both the multi-scaling and multi-wavelet components will be exactly zero, indicating no further refinement to be needed, as expected.

A more challenging situation results from approximating the linear function  $u_2$ , where the number of samples leading to a zero multi-wavelet variance is greater than  $d + 1$  and increases with the dimensionality.

## 6. Numerical benchmarks

The performance of the proposed approach is assessed using various benchmarks. We begin by showing how the integration constant of a variably correlated multi-variate Gaussian can be determined. Identity maps are then used to verify the advantages of constructing representations that are orthogonal to the marginal PDFs of the random inputs over other approaches, e.g., transformation through projection on the cumulative distribution function. The accuracy of approximating two analytic functions with aligned discontinuity and not aligned slope discontinuity is assessed in Sections 6.3 and 6.4, respectively. Next we consider the one-, two- and three-dimensional Kraichnan–Orszag problem, a challenging benchmark often used to assess the convergence of adaptive uncertainty propagation algorithms. Finally, we consider an application in multi-scale cardiovascular simulation where 13 random inputs are first sampled using MCMC and then propagated forward with the proposed approach.

Several criteria are used to assess convergence at the element level in the above benchmarks, related to geometry, number of samples and total multi-wavelet variance. Geometrical criteria include minimum element size and aspect ratio, while various thresholds on the minimum number of samples are applied, based on the discussion in Section 5.4. A converged multi-wavelet variance is also obtained by comparing  $\gamma$  with  $\gamma_b$  as discussed in 5.2. Special cases are also addressed, for example, a zero RVM approximation residual or overall element variance. The algorithm requires other free parameters and thresholds listed in Table 2.

In the benchmarks reported in the following sections, the errors in the estimated statistical moments are evaluated as

$$\varepsilon_1 = \left| \frac{\tilde{\nu}_1^* - \tilde{\nu}_1}{\tilde{\nu}_1} \right| \text{ and } \varepsilon_2 = \left| \frac{\sqrt{\nu_2^*} - \sqrt{\nu_2}}{\sqrt{\nu_2}} \right|, \quad (48)$$

where  $\tilde{\nu}_1^*$  and  $\nu_2^*$  represent the estimates computed through the proposed multi-resolution approach, while  $\tilde{\nu}_1$  and  $\nu_2$  the corresponding exact value. The latter has been computed exactly for stochastic responses with known analytical expression (Sections 6.2–6.4) and through Monte Carlo Simulation (MCS) using  $2 \times 10^6$  samples (Sections 6.6–6.8). In cases where the denominator in (48) is zero, only the numerator is considered.

In the convergence plots shown in the following sections, we consider full tensor grid orders 2, 5 and 8 as well as random sampling with a minimum of 2, 10 and 20 samples in each element. These are denoted “O2”, “O5”, “O8” and “O2”, “O10” and “O20”, respectively. Errors in the Monte Carlo statistics are also reported as “MCS”. Also, a common feature

observed in all convergence plots in the upcoming sections, is that random sampling is typically more oscillatory than collocation on regular grids. This is expected and depends on two facts. First, for random sampling, element refinement and adaptive sampling are performed at two successive iterations, and therefore element refinement at a fixed number of samples may lead to a drop in accuracy for the statistics of interest. This relates to the implementation details of the algorithm. In other words, a multi-wavelet approximation is re-computed after element refinement, without generating additional samples. Second, samples drawn uniformly on bigger elements, may not be sufficiently well spaced on successive refinements, thus reducing the accuracy of the multi-wavelet approximant.

To count the number of samples for collocation grid arrangements, new grids were constructed at each element refinement, and samples from all the parent elements were accounted for when computing the total number of samples.

### 6.1. Regression of variably correlated multi-variate Gaussians

The ability of a multi-resolution approximation with an underlying uniform measure to determine the integration constant of a variably correlated Gaussian is illustrated in this section. Two random variables  $(y_1, y_2) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}_\tau)$  are selected where

$$\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ and } \boldsymbol{\Sigma}_\tau = \begin{bmatrix} 1 & \tau \\ \tau & 1 \end{bmatrix},$$

and the parameter  $\tau$  denotes the correlation coefficient of  $y_1$  and  $y_2$ . While in this simple example we can use a standard transformation of  $y_1$  and  $y_2$  to two independent Gaussians, we use the approach of Section 2.2 for illustration of the idea.

We mimicked a set of MCMC samples by randomly generating realizations from  $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}_\tau)$  and by scaling the associated PDF by a known factor. We then tried to recover this factor using multi-resolution regression. Fig. 3 shows that a few regression iterations are sufficient to determine the integration constant up to an error of a few percent, and that the algorithm is robust for variably correlated  $y_1$  and  $y_2$ . Due to the effect of the correlation and the mapping between an arbitrary support and  $(0, 1)^d$ , it is likely for large areas of the regression domain to be characterized by a small number of samples (as apparent in Fig. 3(a)). To force the regressor to progressively vanish for rare realizations, we enforced zero PDF samples at  $\pm 5$  marginal standard deviations from the mean.

**Remark 3** (*Decorrelation of MCMC Samples*). In regard to the use of regression as in (6) to determine the integration constant from arbitrary correlated samples, it should be noted that, as suggested in [22], it is possible to apply a decorrelation transformation (e.g., through a finite dimensional Karhunen-Loève expansion, see, e.g., [29]) thus improving the regularity of the distribution to approximate. However, we observe how a transformation could prevent the ability to recover sensitivity indices in terms of the original variables. In other words, the sensitivities computed using the transformed variables may not have a straightforward interpretation.

## 6.2. Simple identity maps of arbitrary distributed random variables

In this benchmark, the first and second order moments of the identity  $u(\mathbf{y}) = \mathbf{y}$  are computed for a random input  $\mathbf{y} \sim \mathcal{N}(0, 1)$ . Moments are computed using two different methods, i.e., by using the proposed approach and by retaining an underlying uniform measure and using projections on the Gaussian cumulative distribution. As shown in Figs. 4 and 5, the proposed approach converges significantly faster. We note that distributions with infinite support need to be truncated to fit the unitary hypercube where the proposed multi-resolution basis is constructed. The results in this section have been computed by truncating the distribution at  $\pm 5$  standard deviations from the mean.

## 6.3. Discontinuous function in two dimensions

We consider the discontinuous function proposed in [30]:

$$u(y_1, y_2) = \begin{cases} \sin(\pi y_1) \sin(\pi y_2) & \text{if } y_1 \leq 0.5, y_2 \leq 0.5 \\ 0 & \text{otherwise} \end{cases}, \quad (49)$$

where the random variables  $y_1, y_2$  are independent and uniformly distributed on  $(0, 1)^2$ . This function has two line discontinuities along  $y_1 = 1/2$  and  $y_2 = 1/2$ , respectively, and thus methods based on global polynomial regression are likely to result in poor approximation. Convergence plots for the mean and standard deviation are shown in Fig. 7. The 2nd-order sparse grid is unable to provide an accurate approximation of  $v_2$ . While areas where  $u(y_1, y_2) = 0$  are correctly approximated, computation of expansion coefficients involving the product of a quadratic basis with a sinusoidal response is not accurate.

## 6.4. Non aligned slope discontinuity in two dimensions

Convergence of the proposed formulation is demonstrated, in this section, for an analytical function with slope discontinuity not aligned with the Cartesian axis. This example is proposed in [10] to show the convergence properties of an adaptive hierarchical sparse grid collocation algorithm. Consider the function on  $(0, 1)^2$  defined as:

$$u(y_1, y_2) = \frac{1}{|0.3 - y_1^2 - y_2^2| + \delta}, \quad (50)$$

where  $\delta = 10^{-1}$ . Convergence plots are shown in Fig. 8. The advantages of random sampling (with a minimum number of 10 and 20 samples per element) start to appear both in the computation of the average value and standard deviation. For responses with features not aligned with the adopted Cartesian integration grid, improved approximation results from a combination of random sampling and RVM regression.

For this and the previous case (a visualization of their stochastic response and parameter space refinement is shown in Fig. 6), we look at both the convergence in terms of the first and second order statistical moments, and the  $\ell_1$ ,  $\ell_2$  and  $\ell_\infty$  error norms evaluated at  $n_{MC} =$

$10^3$  random points uniformly distributed on  $\Sigma_{\mathbf{y}} = (0, 1)^2$ . Specifically, the following expressions are used:

$$\begin{aligned} e_{\infty} &= \max_{i=1, \dots, n_{MC}} \left| u^{(i),*}(y_1, y_2) - u^{(i)}(y_1, y_2) \right|, \\ e_1 &= \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \left| u^{(i),*}(y_1, y_2) - u^{(i)}(y_1, y_2) \right|, \\ e_2 &= \left\{ \frac{1}{n_{MC}} \sum_{i=1}^{n_{MC}} \left[ u^{(i),*}(y_1, y_2) - u^{(i)}(y_1, y_2) \right]^2 \right\}^{1/2}, \end{aligned} \quad (51)$$

where  $u^{(i),*}$  represents the approximation computed through the proposed multi-resolution regression at the  $i$ th check point and the associated exact solution  $u^{(i)}$ . Monotonic convergence is observed for both quadrature grids and random samples (see Fig. 9), where the former require significant additional samples. This plot further highlights the benefits of the proposed approach with respect to grid collocation.

### 6.5. The Kraichnan–Orszag problem

The Kraichnan–Orszag (KO) problem has become a standard benchmark for adaptive uncertainty propagation algorithms (see, e.g., [10,9]). It consists of a system of three coupled non-linear ODEs derived from the simplified inviscid Navier–Stokes equations [31]. When uncertainty is injected in the initial conditions of the three state variables, the stochastic response is characterized by sharp gradients that rapidly evolve with time. In this section, we adopt a formulation of the KO problem discussed in [9], expressed through the following system:

$$\frac{du_1}{dt} = u_1 u_3, \quad \frac{du_2}{dt} = u_2 u_3, \quad \frac{du_3}{dt} = u_1^2 + u_2^2. \quad (52)$$

Note that when the set of initial conditions is selected such that the planes  $u_1 = 0$  and  $u_2 = 0$  are consistently crossed, the accuracy of the global polynomial approximations (at the stochastic level) deteriorates rapidly in time [9].

### 6.6. One dimensional Kraichnan–Orszag problem

We assume initial conditions for (52) to be random and specified as

$$u_1(t=0)=1, \quad u_2(t=0)=0.2y - 0.1, \quad u_3(t=0)=0, \quad (53)$$

where  $y$  is uniformly distributed on  $(0, 1)$ . Fig. 12(a) illustrates the adaptive reconstruction of  $u_1(y)$  for an increasing number of elements with samples generated from a double Clenshaw–Curtis integration rule, while Fig. 10 shows the convergence of the first two estimated statistical moments of  $u_1(y)$ . Convergence plots are shown in Fig. 10. While lower grid orders or number of samples provide inaccurate estimates of the standard deviation,

significant improvements over Monte Carlo Sampling are obtained by adding more samples. We note that the stochastic response is characterized by sharp local features (Fig. 12(a)). Under these conditions, the efficiency of collocation grids is maximized, resulting in similar sample requirements as random sampling.

### 6.7. Two dimensional Kraichnan–Orszag problem

The initial conditions of the Kraichnan–Orszag problem are assumed to be uncertain and functions of two random variables

$$u_1(t=0)=1, \quad u_2(t=0)=0.2y_1 - 0.1, \quad u_3(t=0)=2y_2 - 1, \quad (54)$$

where  $y_1$  and  $y_2$  are independent and uniformly distributed on  $(0, 1)$ . A two dimensional multi-wavelet approximation of  $u_1$  at  $t = 10$  s is generated with  $m = 2$ . Fig. 11(a) shows the convergence of average value and standard deviation of  $u_1$  using MCS, and a multi-wavelet approximation evaluated on sparse grid and random adaptive sampling. Finally, Fig. 12(b) and (c) show the obtained partition of  $(0, 1)^2$  using tensor grids at refinement iteration 17. Convergence plots are shown in Fig. 11. The proposed approach with random sampling outperforms both grid collocation and MCS.

### 6.8. Three dimensional Kraichnan–Orszag problem

The initial conditions of the three-dimensional Kraichnan–Orszag problem are

$$u_1(t=0)=0.2y_1 - 1.0, \quad u_2(t=0)=0.2y_2 - 0.1, \quad u_3(t=0)=0.2y_3 - 1.0, \quad (55)$$

where we assume that  $y_1$ ,  $y_2$  and  $y_3$  are independent and uniformly distributed on  $(0, 1)$ . The multi-wavelet approximation at  $t = 10$  s is constructed for  $u_1$  using  $m = 2$ . The convergence plots for the standard deviation are shown in Fig. 13. Unlike the prior results, random sampling outperforms quadrature grids but shows a convergence similar to Monte Carlo Sampling, due to the complex features of the stochastic response [10]. While we do not expect, in practice, stochastic responses with features as complicated as this benchmark problem, this case illustrates that the algorithm can compute converging statistics in such cases.

## 7. Application: Forward propagation of local hemodynamic statistics

### 7.1. Motivation

We discuss the application of the proposed framework to a problem in uncertainty propagation for cardiovascular blood flow simulation. Specifically, we are interested in determining both mean values and variability of local hemodynamic quantities in the coronary arteries, in patients following coronary artery bypass graft (CABG) surgery.

To do so, we employ a multi-scale modeling strategy that includes a three-dimensional model of the patient's aorta, aortic branches and coronary arteries, receiving pressure/velocity boundary conditions from a *lumped parameter* description of the peripheral circulation. This lumped parameter or *zero-dimensional* (0D) boundary model is analogous to an electrical circuit and parameterized through a set of resistances (viscous resistance of the blood), capacitances (vessels compliance), inductances (blood inertia) and other parameters regulating the heart function. An three-dimensional anatomical model was constructed using the Simvascular software package [32], starting from a set of patient-specific CT images. A schematic representation of a typical multi-scale model for coronary artery disease is depicted in Fig. 14.

It is evident how a multiplicity of pressure/velocity boundary conditions (or alternatively of 0D model parameters) are able to generate multi-scale model outputs compatible with non-invasively acquired clinical data, especially if these data are affected by uncertainty. We are therefore interested in first determining the *distributions* of boundary conditions by performing data assimilation, i.e., by estimating the 0D model parameters using adaptive Markov chain Monte Carlo (MCMC). Due to the number of model evaluations typically required to produce independent samples from a stationary posterior distribution in MCMC, a condensed resistive representation of the three-dimensional model is also computed, leading to a dramatic reduction in the overall computational cost (from several hours to a fraction of a second, see [33], for a single model solution). The proposed multi-resolution approach to uncertainty propagation is employed to propagate the distributions of assimilated boundary conditions through a full multi-scale model, therefore quantifying the variability in local hemodynamic indicators of interest.

Finally, we address the reasons why the proposed approach should be preferred to Monte Carlo Simulation for cardiovascular applications. The primary factors in the choice of method are the moderate dimensionality, the non-smoothness of the stochastic response (likely limited to a small subsets of dimensions), and the high computational cost needed to produce each multi-scale model evaluation. These conditions arise particularly when studying the effect of combined uncertainty sources, e.g., clinical data, model geometry and vessel wall material properties. Specifically, changes in the model geometry for configurations characterized by competing flow (e.g., total cavopulmonary connection in Fontan completion surgery [34]) are likely to produce a non-smooth variation in the outputs. Under these conditions and from the results of the numerical experiments in Section 6, the proposed approach is likely to result in computational savings with respect to MCS due to the smaller number of model evaluations needed to capture the statistics of interest.

## 7.2. Clinical targets selection and parameter identifiability study

In this section we provide details on the preliminary steps needed to produce random input samples for the uncertainty propagation task. Specifically, we discuss clinical target selection for data assimilation and preliminary identifiability analysis. Please refer to [33] for further details. Targets were acquired using routine clinical measurements, population average values, echocardiography data, and complemented with literature data on coronary flow. Diastolic, systolic and mean aortic pressures were estimated from the patient's cuff



pressures during echocardiogram. Other patient-specific measurements from echocardiography included stroke volume, ejection fraction, ratio of early to late flows into the left ventricle, valve opening times relative to the heart cycle duration, systolic pressure difference between the right ventricle and right atrium measured from tricuspid valve regurgitation, and the mean right atrial pressure estimated from the diameter and degree of inspiratory collapse of the inferior vena cava [33].

Additional data from literature was used to improve parameter identifiability. In this regard, a mean pulmonary pressure of 14.0 mmHg was considered together with 4% of the cardiac output entering the coronaries. Representative left and right coronary waveforms from Doppler flow wire measurements [35] were also used to quantify the diastolic to systolic coronary peak flow ratio, the diastolic to systolic coronary blood volume, and the ratio of the flow in the coronary arteries occurring in the first 1/3 and 1/2 of the heart cycle.

Standard deviations for these targets were determined from the literature and personal communication with our clinical collaborators. Measurements derived from tracking the left ventricular wall during echocardiography are characterized by limited uncertainty, whereas higher uncertainties occur in flow measurements from doppler echocardiography.

Identification of both influential and unimportant parameters was performed using a combination of prior experience in manual tuning, analysis of structural identifiability through the Fisher Information Matrix (FIM), and the marginal parameter distributions resulting from Bayesian estimation (where a flat marginal posterior is generally indicative of unimportant parameters). Previous experience with manual tuning allowed us to immediately identify parameters significantly affecting the systemic blood pressure, stroke volume, systemic to coronary flow split, and qualitative shape of the coronary flow waveforms. The left ventricular elastance  $E_{lv}$  affects the stroke volume and mean systemic blood pressure, while the aortic compliance  $C_{ao}$  affects the systemic diastolic to systolic pressure range. The coronary downstream resistance scaling factor  $R_{am,cor}$  has a strong influence on the flow splits, while the coronary capacitances and  $dP/dt_r$  (i.e., the intramyocardial pressure time derivative) significantly affect the shape of the coronary waveforms. The downstream systemic resistance scaling factor  $R_{scr}$  affects both the aortic diastolic to systolic pressure range and the stroke volume.

The Fisher Information Matrix (FIM) rank is a local measure of identifiability, i.e., the ability to uniquely characterize parameter combinations from the available observations [36]. A singular FIM reveals the presence of non-identifiable parameter combinations and an analysis of the eigenvectors associated with the zero eigenvalues (so-called *null eigenvectors*) can be useful to identify these parameters. In this regard, a null eigenvector with a dominant component along a certain parameter typically suggests this parameter to be unimportant. Thus, we used the FIM to detect all the unimportant parameters until no more trivial null eigenvectors were found.

We then examined the estimated posterior marginal variance as a measure of global identifiability. In particular, we quantified the ability to learn the parameter  $y_i$  by comparing the posterior and prior marginal variances using the coefficient  $\theta_i$  as

$$\theta_{i=1} = \sqrt{\frac{\nu_2[y_i|\mathbf{d}]}{\nu_2[\mathbf{d}]}} \quad (56)$$

where  $\nu_2[y_i|\mathbf{d}]$  and  $\nu_2[\mathbf{d}]$  are the posterior and prior marginal variances, respectively. Well-learned parameters will have a much smaller variance after being conditioned to the clinical observations ( $\theta \rightarrow 1$ ), while poorly learned parameters, on the other hand, will have resulting variances that are close to their prior variance ( $\theta \rightarrow 0$ ). Intuitively, parameters associated with large marginal variances will likely have a limited impact on the model results selected for identification, since these parameters can take on a wide range of values, but produce results with similar posterior.

Prior tuning experience, FIM analysis, Bayesian learning metrics, and qualitative knowledge of expected physiologic coronary waveforms, allowed us to determine a reduced set of 13 parameters (see Table 3). This subset consistently generated excellent agreement between model outputs and clinical targets, and produced flow and pressure waveforms consistent with qualitative expectations [33].

An adaptive Markov chain Monte Carlo strategy was used, at this point, to sample from the posterior distribution of these 13 model parameters, constructed using the clinical data and associated uncertainty discussed above.

### 7.3. Uncertainty propagation results

The availability of samples from the joint posterior distribution of the random inputs, is a prerequisite to performing uncertainty propagation using the proposed approach, with the goal of statistically quantifying the outputs of interest. As a first step, the marginals were determined for the random inputs  $\mathbf{y}$  (Fig. 15) as well as the coefficients  $\beta^{(k)}$ ,  $k = 1, \dots, M$ . We then used 64 *sub-samples* from the MCMC parameter traces (after removing the realizations produced during *burn-in*, see [37]) and computed solutions of the full multi-scale model at these realizations using a stabilized finite element solver [38]. Note how the parameter realizations provided by MCMC are roughly 1.5M and it would be impossible to compute a stabilized finite element solution for all the MCMC parameter realizations. Statistics were computed through Monte Carlo averaging and by forward propagation using the proposed multi-resolution approach, using the same number of model solutions (64) for various multi-wavelet orders. Results were very similar to those provided by MCS, as shown in Figs. 16 and 17. Fig. 16 reports the average value and standard deviation for the pressure distribution in one of the seven patients analyzed in [33]. We see that the standard deviation is not negligible in this case, and the proposed approach with a multi-wavelet order of 1 produces the same results as MCS. Fig. 17 represents a typical case where, instead, the variability of wall shear stress-derived quantities is negligible with respect to the associated mean value. We observe that, even in these cases, there is a good agreement between the MCS results and those produced by the proposed approach.

This confirms the accuracy of the approach when random inputs are not independent, but available through MCMC samples. Moreover, we have shown in the previous sections that

the proposed approach outperforms MCS, requiring in most cases a significantly smaller number of model evaluations to accurately compute the output statistics. This suggests that multi-resolution uncertainty propagation has enough flexibility to handle cases where estimated distributions of boundary conditions are combined with assumed geometrical and/or material parameter uncertainties. Future studies will investigate the performance of the proposed approach under these more general settings.

## 8. Conclusion

We propose a generalized multi-resolution expansion with random inputs characterized by arbitrary probability distributions and random sample locations. The framework has built-in adaptivity metrics based on the natural decomposition of the total regressor variance in multi-scaling and multi-wavelet contributions. A set of *ad-hoc* multi-resolution bases is constructed orthonormal with respect to an arbitrary distribution function defined on a subset of the input space. This approach, particularized to uniform underlying measures with Legendre multi-scaling, can be used to evaluate the integration constant and the marginal distributions from Markov chain Monte Carlo samples, allowing to perform expectations of arbitrary correlated random inputs. Performance and convergence of the method was demonstrated on several model problems and on a large scale computational model in cardiovascular hemodynamics. For problems in stochastic cardiovascular modeling, especially under a multiplicity of uncertainty sources, the proposed approach is expected to provide a computationally more efficient alternative to Monte Carlo Sampling, leading to accurate statistics with a smaller number of required multi-scale model evaluations.

## Acknowledgments

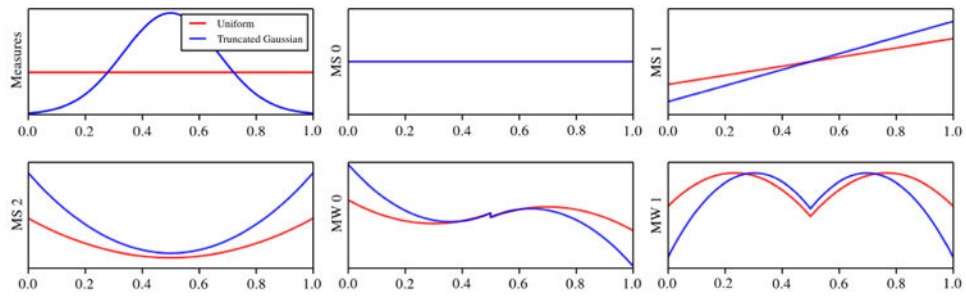
The authors would like to thank the two anonymous Reviewers for their generous comments and feedback that greatly contributed to improve the consistency and quality of the present contribution. This work was supported by American Heart Association Grant #15POST23010012 (Daniele Schiavazzi) a Burroughs Wellcome Fund Career award at the Scientific Interface, NSF CAREER OCI-1150184, NIH R01HL123689, NIH R01 PA16285 (Alison Marsden) and used computational resources from the Extreme Science and Engineering Discovery Environment (XSEDE), supported by National Science Foundation grant number ACI-1053575. This material is based upon work supported by the U.S. Department of Energy Office of Science, Office of Advanced Scientific Computing Research, under Award Number DE-SC0006402 and NSF grant CMMI-145460. We also acknowledge the open source SimVascular project at [www.simvascular.org](http://www.simvascular.org).

## References

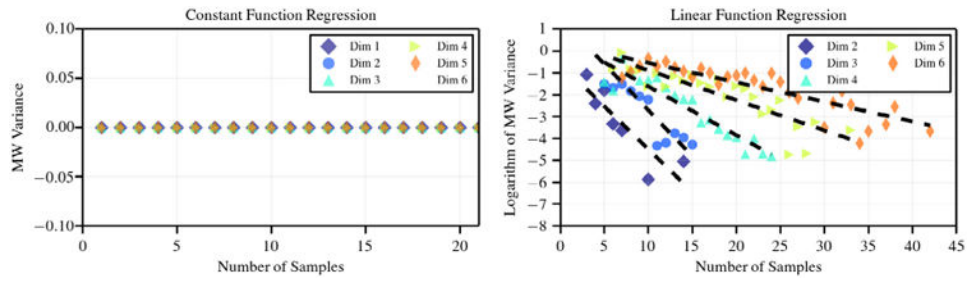
1. Wiener N. The homogeneous chaos. *Amer J Math.* 1938; 60:897–936.
2. Cameron R, Martin W. The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals. *Ann of Math.* 1947:385–392.
3. Ghanem, R., Spanos, P. *Stochastic Finite Elements: A Spectral Approach.* Courier Corporation; 2003.
4. Xiu D, Karniadakis G. The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM J Sci Comput.* 2002; 24:619–644.
5. Ernst O, Mugler A, Starkloff H, Ullmann E. On the convergence of generalized polynomial chaos expansions. *ESAIM Math Model Numer Anal.* 2012; 46:317–339.
6. Babuka I, Nobile F, Tempone R. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM J Numer Anal.* 2007; 45:1005–1034.
7. Nobile F, Tempone R, Webster C. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J Numer Anal.* 2008; 46:2309–2345.

8. Nobile F, Tempone R, Webster C. An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J Numer Anal.* 2008; 46:2411–2442.
9. Wan X, Karniadakis G. An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. *J Comput Phys.* 2005; 209:617–642.
10. Ma X, Zabarar N. An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations. *J Comput Phys.* 2009; 228:3084–3113.
11. Witteveen J, Iaccarino G. Simplex stochastic collocation with random sampling and extrapolation for nonhypercube probability spaces. *SIAM J Sci Comput.* 2012; 34:A814–A838.
12. Doostan A, Owhadi H. A non-adapted sparse approximation of PDEs with stochastic inputs. *J Comput Phys.* 2011; 230:3015–3034.
13. Peng J, Hampton J, Doostan A. A weighted  $\ell_1$ -minimization approach for sparse polynomial chaos expansions. *J Comput Phys.* 2014; 267:92–111.
14. Bilonis I, Zabarar N. Multidimensional adaptive relevance vector machines for uncertainty quantification. *SIAM J Sci Comput.* 2012; 34:B881–B908.
15. Le Maître O, Najm H, Ghanem R, Knio O. Multi-resolution analysis of Wiener-type uncertainty propagation schemes. *J Comput Phys.* 2004; 197:502–531.
16. Schiavazzi D, Doostan A, Iaccarino G. Sparse multiresolution regression for uncertainty propagation. *Int J Uncertain Quantif.* 2014; 4
17. Sankaran S, Marsden A. A stochastic collocation method for uncertainty quantification and propagation in cardiovascular simulations. *J Biomech Eng.* 2011; 133:031001. [PubMed: 21303177]
18. Sankaran S, Marsden A. The impact of uncertainty on shape optimization of idealized bypass graft models in unsteady flow. *Phys Fluids.* 2010; 22:121902. 1994-present.
19. Xiu D, Sherwin S. Parametric uncertainty analysis of pulse wave propagation in a model of a human arterial network. *J Comput Phys.* 2007; 226:1385–1407.
20. Schiavazzi D, Arbia G, Baker C, Hlavacek A, Hsia T, Marsden A, Vignon-Clementel I, et al. Uncertainty quantification in virtual surgery hemodynamics predictions for single ventricle palliation. *Int J Numer Methods Biomed Eng.* 2015
21. Alpert B. A class of bases in  $L^2$  for the sparse representation of integral operators. *SIAM J Math Anal.* 1993; 24:246–262.
22. Soize C, Ghanem R. Physical systems with random uncertainties: chaos representations with arbitrary probability measure. *SIAM J Sci Comput.* 2004; 26:395–410.
23. Golub G, Welsch J. Calculation of Gauss quadrature rules. *Math Comp.* 1969; 23:221–230.
24. Clenshaw C, Curtis A. A method for numerical integration on an automatic computer. *Numer Math.* 1960; 2:197–205.
25. Blatman G, Sudret B. Adaptive sparse polynomial chaos expansion based on least angle regression. *J Comput Phys.* 2011; 230:2345–2367.
26. Tipping M. Sparse Bayesian learning and the relevance vector machine. *J Mach Learn Res.* 2001; 1:211–244.
27. Tipping M, Faul A, et al. Fast marginal likelihood maximisation for sparse Bayesian models. *Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics.* 2003; 1
28. Sobol I. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math Comput Simulation.* 2001; 55:271–280.
29. Loeve M. *Probability Theory II (Graduate Texts in Mathematics).* 1994
30. Agarwal N, Aluru N. A domain adaptive stochastic collocation approach for analysis of MEMS under uncertainties. *J Comput Phys.* 2009; 228:7662–7688.
31. Kraichnan R. Direct-interaction approximation for a system of several interacting simple shear waves. *Phys Fluids.* 1963; 6:1603–1609. 1958-1988.
32. Wilson, N., Wang, K., Dutton, R., Taylor, C. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2001.* Springer; 2001. A software framework for creating patient specific geometric models from medical imaging data for simulation based medical planning of vascular surgery; p. 449-456.

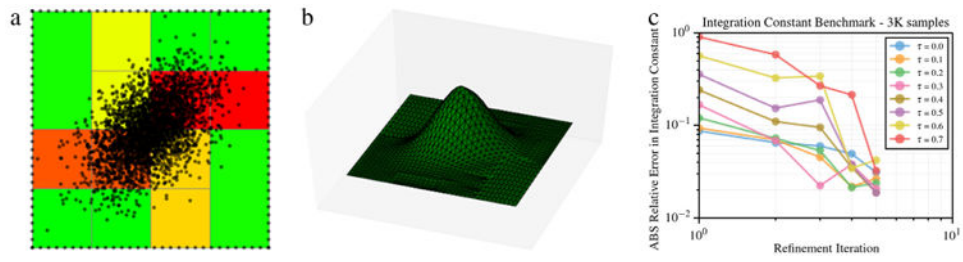
33. Tran J, Schiavazzi D, Bangalore Ramachandra A, Kahn A, Marsden A. Automated tuning for parameter identification and uncertainty quantification in multi-scale coronary simulations. *Comput Fluids*. 2016 in press.
34. Fontan F, Baudet E. Surgical repair of tricuspid atresia. *Thorax*. 1971; 26:240–248. [PubMed: 5089489]
35. Ofli E, Kern M, Vrain J, Donohue T, Bach R, Al-Joundi B, Aguirre F, Castello R, Labovitz A. Differential characterization of blood flow, velocity, and vascular resistance between proximal and distal normal epicardial human coronary arteries: analysis by intracoronary Doppler spectral flow velocity. *Am Heart J*. 1995; 130:37–46. [PubMed: 7611121]
36. Rothenberg T. Identification in parametric models. *Econometrica*. 1971:577–591.
37. Geyer C. Practical Markov chain Monte Carlo. *Statist Sci*. 1992:473–483.
38. Whiting C, Jansen K. A stabilized finite element method for the incompressible Navier–Stokes equations using a hierarchical basis. *Int J Numer Methods Fluids*. 2001; 35:93–116.



**Fig. 1.** Examples of multi-scaling and multi-wavelet basis orthogonal to a uniform and truncated Gaussian distribution, respectively. MS denotes multi-scaling functions while MW multi-wavelet functions.



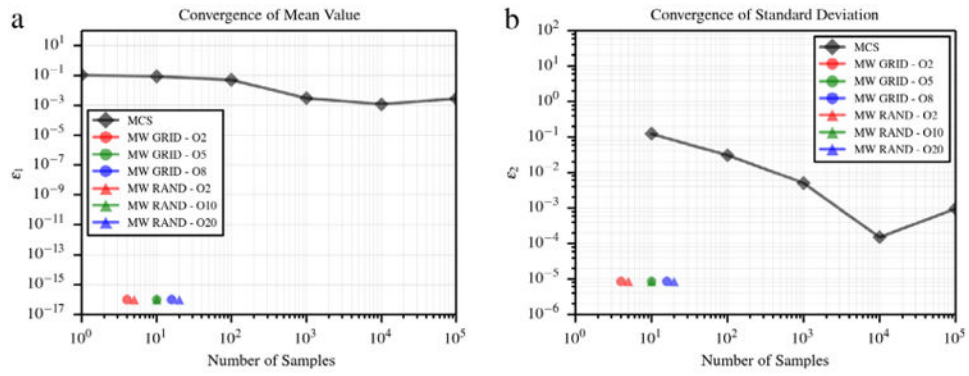
**Fig. 2.** Decay of multi-wavelet variance for an increasing number of samples for constant (left) and quadratic (right) functions with exact representation.



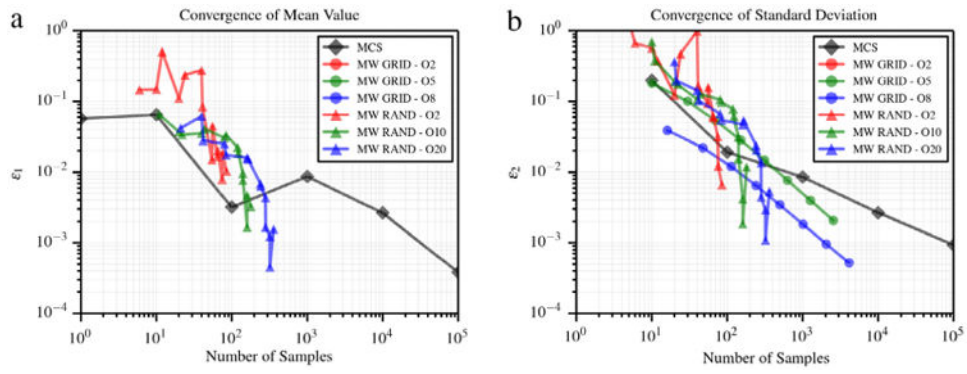
**Fig. 3.**

Computing the integration constant of a variably correlated multi-variate Gaussian with multi-resolution regression. The sample distribution and refined hypercube partitions are shown for  $\tau = 1/2$  (a), together with the reconstructed distribution (b) and convergence plots in terms of the absolute (ABS) relative error between the estimate and the true integration constant (c).

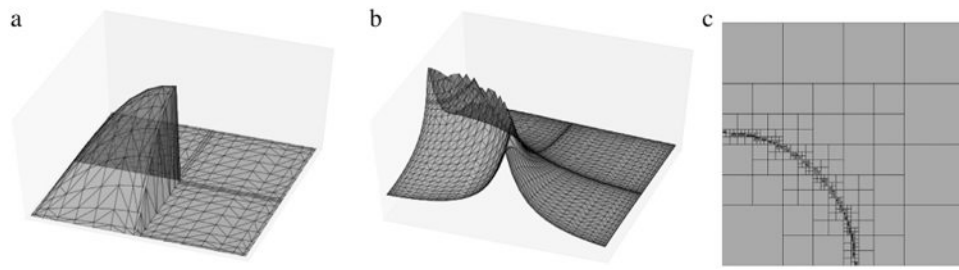




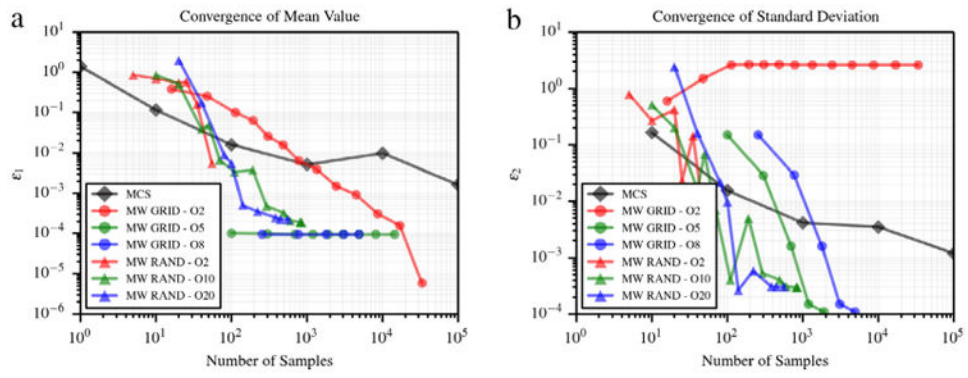
**Fig. 4.** Error in average value (a) and standard deviation (b) for identity map of Gaussian random variable computed using a multi-wavelet expansion with orthogonality with respect to the truncated Gaussian distribution.



**Fig. 5.** Error in average value (a) and standard deviation (b) for identity map of Gaussian random variable computed using a multi-wavelet expansion with orthogonality with respect to  $\mathcal{U}(0, 1)$  and transformation through projections on the cumulative distribution. Note that in (a) the statistics computed using quadrature grids of increasing order are not visible, due to the very small associated error.



**Fig. 6.** Representation of the two non-smooth benchmark functions. The first (a) is expressed through the relationship (49) and has been proposed in [30], while the second (b) is expressed using (50) and proposed in [10]. For this latter function, the hypercube partition at refinement iteration 19 is also shown (c).



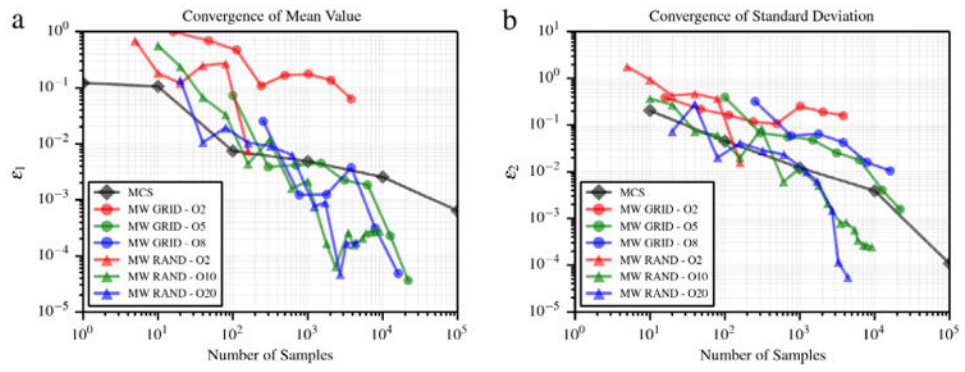
**Fig. 7.** Error in average value (a) and standard deviation (b) for the discontinuous function proposed in [30]

Author Manuscript

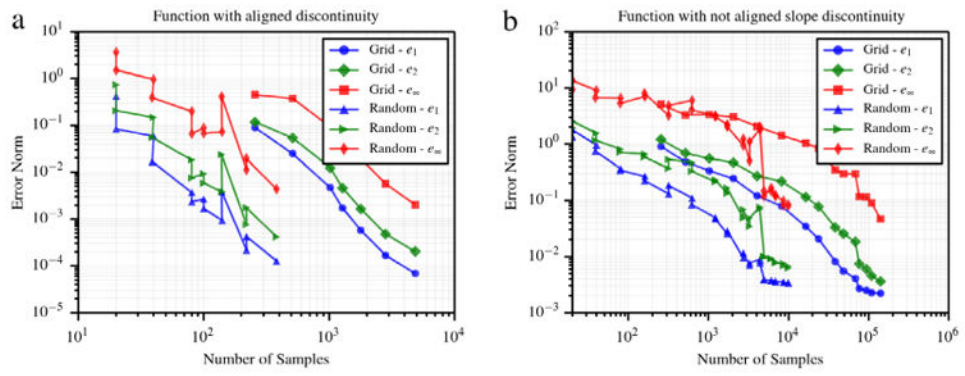
Author Manuscript

Author Manuscript

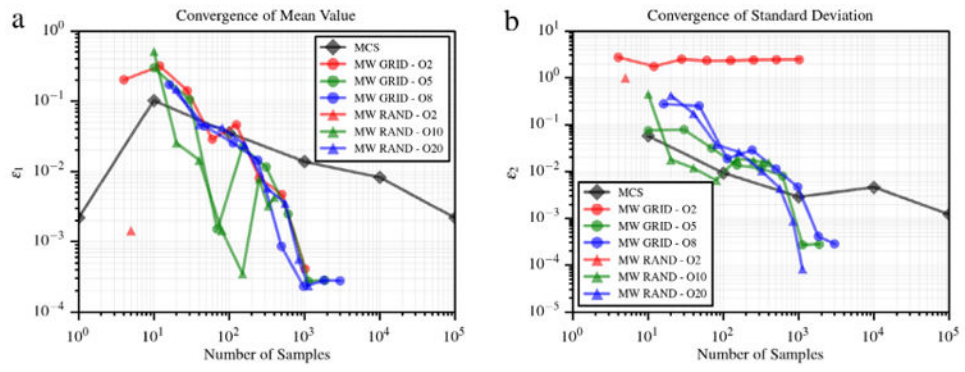
Author Manuscript



**Fig. 8.** Error in average value (a) and standard deviation (b) for the function with slope discontinuity proposed in [10].



**Fig. 9.** Convergence of error norms  $e_1$ ,  $e_2$  and  $e_\infty$ , plotted using 1000 random check points. (a) Convergence for (49), (b) convergence for (50).



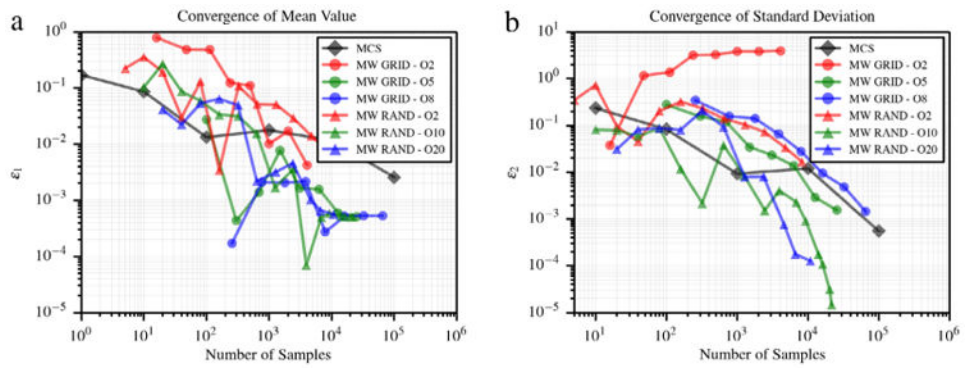
**Fig. 10.** Convergence in average value (a) and standard deviation (b) for the 1D KO problem.

Author Manuscript

Author Manuscript

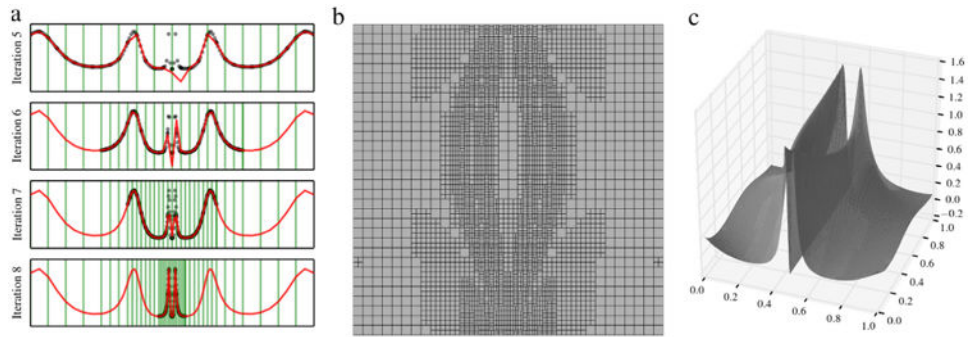
Author Manuscript

Author Manuscript

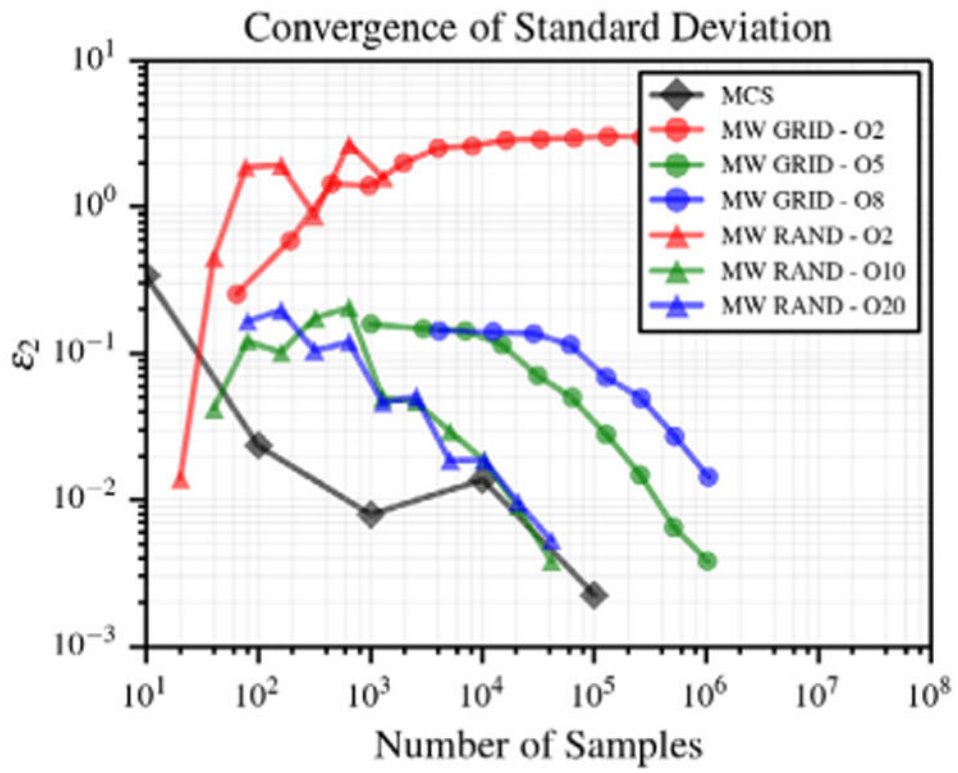


**Fig. 11.** Convergence in average value (a) and standard deviation (b) for the 2D KO problem.

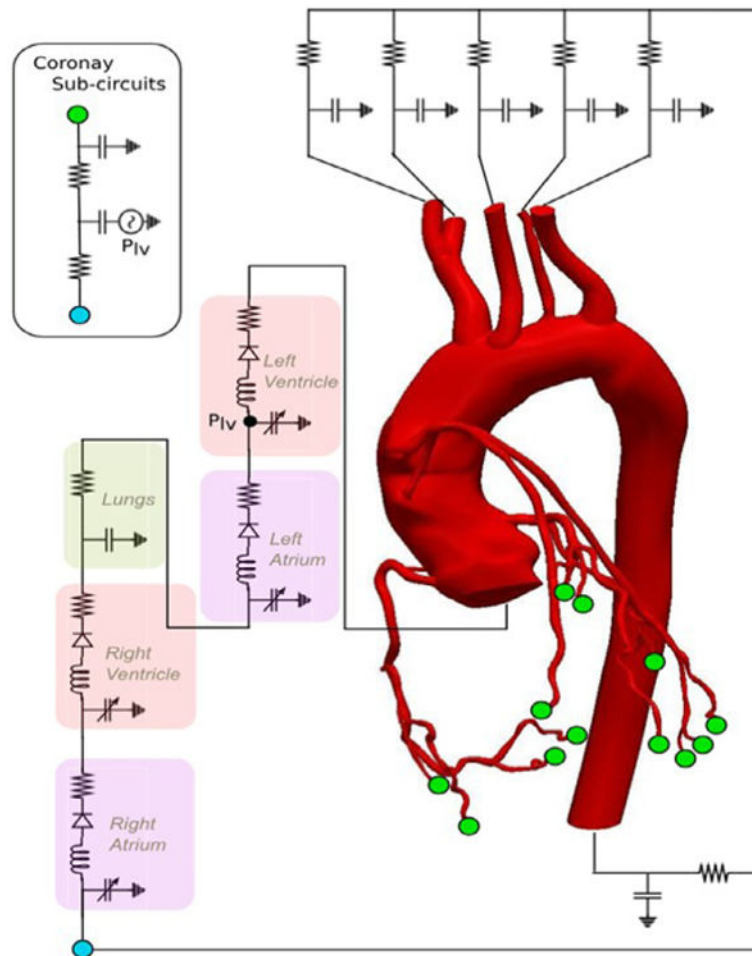




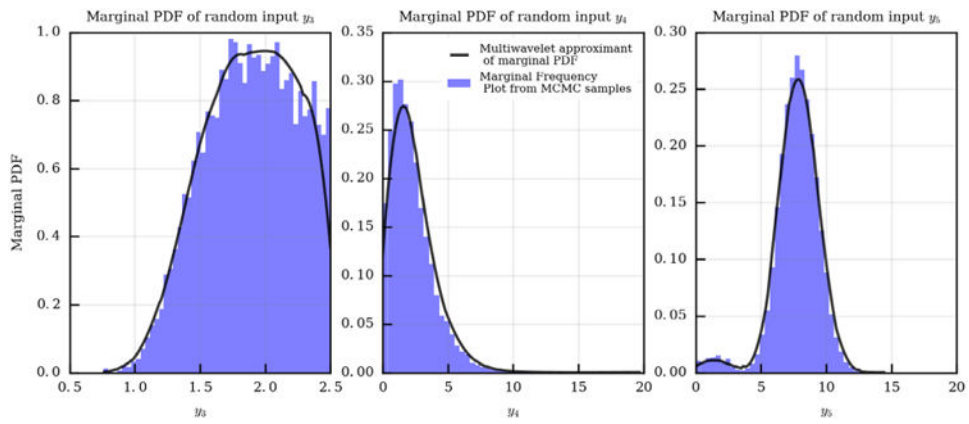
**Fig. 12.** Iterative refinement of  $(0, 1)$  and progressive approximation of the stochastic response  $u_1(y, t)$  at  $t = 30$  s for the one-dimensional KO problem (a). Refinement of  $(0, 1)^2$  (b) and reconstructed response (c) for the two-dimensional KO problem.



**Fig. 13.** Convergence in standard deviation for the 3D KO problem.



**Fig. 14.** Schematic representation of a multi-scale model for coronary artery bypass surgery from [33]. The anatomical model contains the aortic arch, thoracic aortic branches, coronary arteries and bypass grafts. The boundary conditions are provided by a lumped parameter model that mimics the heart and peripheral circulation.



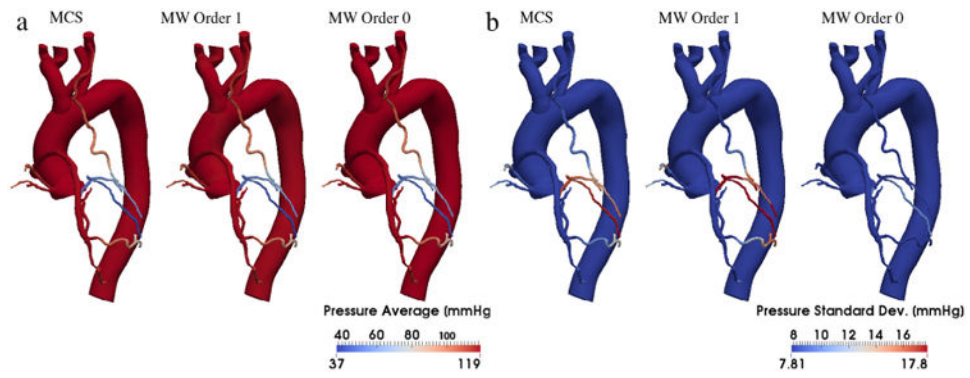
**Fig. 15.** Normalized frequency plots and multiwavelet approximants for the marginal PDF of random inputs  $y_3$ ,  $y_4$  and  $y_5$ .

Author Manuscript

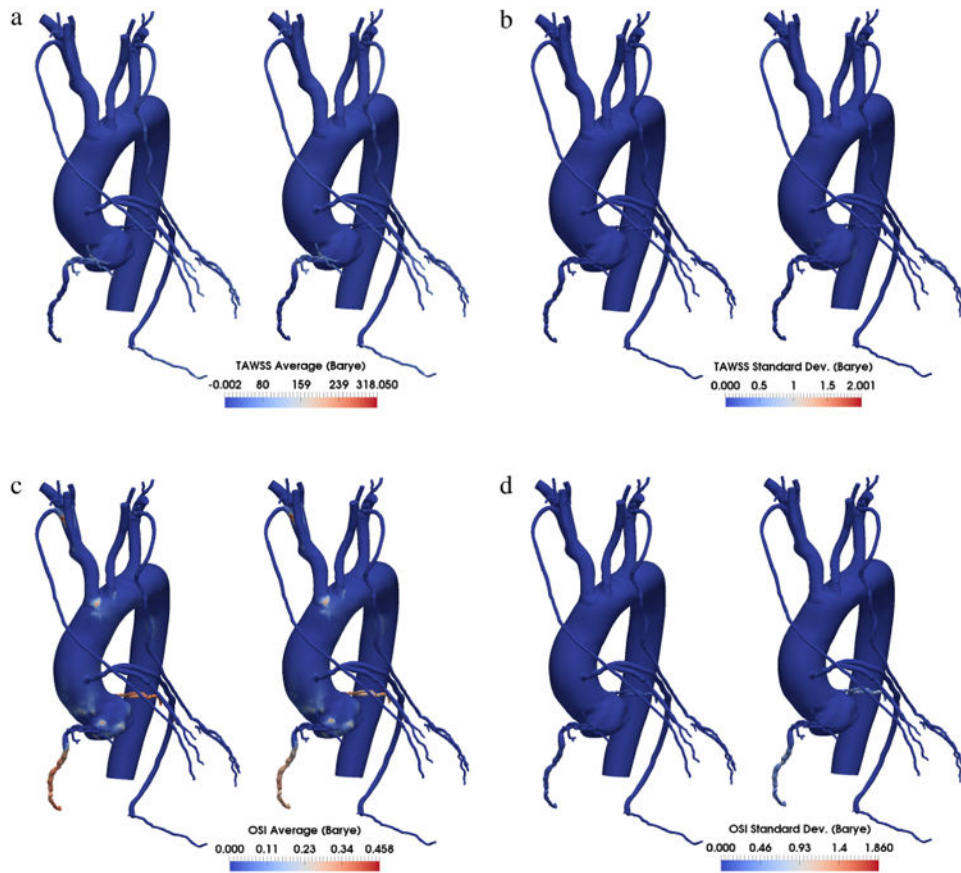
Author Manuscript

Author Manuscript

Author Manuscript



**Fig. 16.** Contour plot representing the estimates of the average pressure (a) and its standard deviation (b) on a three-dimensional multi-scale model of the aorta and coronary circulation. The statistics of the proposed multi-resolution approach are compared to those obtained from Monte Carlo sampling, computed over 64 realizations.



**Fig. 17.** Comparison between average and standard deviation of time averaged wall shear stress (TAWSS) and oscillatory shear index (OSI) computed using Monte Carlo simulation and the proposed multi-resolution approach. The two approaches result in practically the same contours for the quantities of interests.

Assembly time and storage requirements for multi-wavelet matrix evaluated at 100 realizations of the random inputs. Sparse tensor product multi-indices with total order less than  $m$  are employed. The times reported were obtained on an Intel® Core™ i7-4770 CPU, 3.40 GHz with a cache size of 8192 KB.

**Table 1**

Test No.	Dimensions	MW order	MW Quad. order	Time [s]	Mat. rows	Mat. columns	Mat. size [MB]
1	5	1	10	0.0048	100	36	0.027
2	5	2	10	0.0049	100	126	0.096
3	5	3	10	0.0136	100	336	0.256
4	20	1	10	0.0405	100	441	0.336
5	20	2	10	0.4598	100	4851	3.701
6	20	3	10	3.8565	100	37191	28.374
7	50	1	10	0.559620	100	2601	1.984
8	50	2	10	14.948165	100	67626	51.595

**Table 2**

List and description of the free parameters and threshold quantities used in the proposed uncertainty propagation approach.

Free parameter	Value for test cases in Section 6	Description
Number of initial samples (random sampling) or Sampling grid order (grid collocation).	2, 10, 20 for “O2”, “O10” and “O20” test cases, respectively.	This parameter quantifies the initial number of model solutions that will be used to evaluate a Multi-wavelet expansion on the full hypercube. Alternatively, for collocation on quadrature grids, it quantifies the grid order in every element.
Maximum multi-wavelet order.	$m = 2$ .	This is the maximum order of the multi-wavelet basis. Generally orders 1 and 2 are preferred. We found that a piecewise quadratic approximation usually suffices to obtain a satisfactory multi-wavelet approximant.
Quadrature order for multi-wavelet basis construction.	$n_q = 30$ .	Sets the order of the univariate quadrature rule when computing integrals during the construction of the multi-wavelet basis, as discussed in Section 3.1.
Minimum element size.	$1.0 \times 10^{-3}$ (1/1000 of the minimum size of $\mathcal{S}_y$ ).	This is the minimum element size allowed for a single dimension. Elements having size smaller than this threshold are left unrefined.
Maximum partition aspect ratio.	Default set to 1.0 (promotes <i>squared</i> elements).	This parameter governs element aspect ratios.
Minimum number of samples in partition (random sampling only).	2, 10, 20 for “O2”, “O10” and “O20” test cases, respectively.	For random sampling, elements with less than this number of samples are left unrefined.
Multi-wavelet variance ratio $\gamma_t$ .	$\gamma_t \in [10^{-8}, 10^{-6}]$ .	This parameter (discussed in Section 5.2) sets a limit in the multi-wavelet variance for each element. Elements characterized by a multi-wavelet variance smaller than $\gamma_t$ are regarded as converged.



**Table 3**

Thirteen identifiable 0D model parameters included as stochastic inputs.

Parameter	Description	Parameter	Description
$E_{r\sigma}$	Right ventricular elastance.	$C_{am,l}$	Distal capacitance at left coronary outlet.
$E_{l\sigma}$	Left ventricular elastance.	$C_{a,l}$	Proximal capacitance at left coronary outlet.
$E_{l\sigma p}$	Scaling factor for left ventricular elastance rate.	$C_{am,r}$	Distal capacitance at right coronary outlet.
$C_{ao}$	Aortic compliance.	$C_{a,r}$	Proximal capacitance at right coronary outlet.
$K_{sp,ra}$	Passive right atrial curve scaling factor.	$R_{rcr}$	Downstream systemic resistance scaling factor.
$K_{sp,la}$	Passive left atrial curve scaling factor.	$dP/dt_r$	Intramyocardial pressure time derivative.
$R_{am}$	Coronary downstream resistance scaling factor.		