Probability Surveys
Vol. 10 (2013) 51–68
ISSN: 1549-5787
DOI: 10.1214/13-PS219

On spectral methods for variance based sensitivity analysis

Alen Alexanderian

The University of Texas at Austin, Institute for Computational Engineering and Sciences, Austin, Texas
e-mail: alen@ices.utexas.edu.

Abstract: Consider a mathematical model with a finite number of random parameters. Variance based sensitivity analysis provides a framework to characterize the contribution of the individual parameters to the total variance of the model response. We consider the spectral methods for variance based sensitivity analysis which utilize representations of square integrable random variables in a generalized polynomial chaos basis. Taking a measure theoretic point of view, we provide a rigorous and at the same time intuitive perspective on the spectral methods for variance based sensitivity analysis. Moreover, we discuss approximation errors incurred by fixing inessential random parameters, when approximating functions with generalized polynomial chaos expansions.

Keywords and phrases: Variance based sensitivity analysis, analysis of variance, spectral methods, generalized polynomial chaos, orthogonal polynomials, conditional expectation.

Received August 2013.

1. Introduction

The aim of this article is to provide a clear and rigorous understanding of spectral methods for variance based sensitivity analysis which employ generalized polynomial chaos expansions. Our discussion concerns square integrable functions of finitely many independent random variables. The original idea of variance based sensitivity analysis goes back to the work of I.M. Sobol in [22]. Other notable subsequent papers in the field include [11, 19, 23]. Unlike *local* sensitivity analysis, which uses derivative information to assess the sensitivity of a model to parameters, variance based sensitivity analysis measures the contribution of each parameter to the total variance. This is why variance based sensitivity analysis is also referred to as *global* sensitivity analysis. Specifically, given a function X of d random inputs, parameterized by random variables $\xi_1, \ldots, \xi_d$, a variance based sensitivity analysis aims to quantify the contribution of each ξ_i (or subcollections of $\xi_1, \ldots, \xi_d$) to the variance of X. In [22], the ANOVA decomposition[1] of functions of independent random variables was used to characterize a number of useful sensitivity indices. The classical numerical recipes

[1] The term ANOVA is short for analysis of variance [5]. The ANOVA decomposition is sometimes referred to as the Sobol decomposition in the literature.

for computation of these indices involve sampling based methods which, in general, require a large number of evaluations of the function X. However, there are many applications where such function evaluations are expensive; for example, evaluating X may require solving a time-dependent partial differential equation. In such cases, sampling based methods tend to become computationally prohibitive.

On the other hand, in the recent years, the spectral methods for quantifying parametric uncertainties, which utilize generalized polynomial chaos expansions, have furnished a host of efficient techniques for analysis of uncertainties in computationally expensive mathematical models; the references [10, 28, 18, 16, 15, 17, 14, 21, 9, 2] represent a small sample of the available literature in this area. The so called generalized polynomial chaos expansions are Fourier expansions in appropriately chosen multivariate orthogonal polynomial bases. The theory of polynomial chaos expansions go back to the seminal work of N. Wiener in [26] and R. Cameron and W. Martin in [7]. The practical applications of polynomial chaos expansions often involve a simple special case of the general theory, where one uses a *finite* number of canonical random variables to parameterize uncertainties in a mathematical model. Once available, these expansions can be used to efficiently characterize the statistical properties of square integrable random variables. In particular, the variance based sensitivity indices can be computed at a negligible computational cost, once such spectral expansions are available. This important point was noted in the papers [8, 25] which describe efficient numerical computation of the sensitivity indices with generalized polynomial chaos expansions.

Most of the recent papers discussing the computation of variance based sensitivity indices via generalized polynomial chaos expansions begin by a discussion of the ANOVA (Sobol) functional decomposition followed by the description of variance based sensitivity indices in terms of this decomposition; subsequently, after discussing the relation of ANOVA decompositions to polynomial chaos, they describe the computation of the indices using the polynomial chaos expansion, often through an informal argument. Finally, after all the dust has settled, one arrives at some simple expressions for the sensitivity indices in terms of polynomial chaos expansions. We choose to take a different path and consider the variance based sensitivity indices, which are defined *independently* of any decomposition, from a measure theoretic point of view. These sensitivity indices are defined in terms of conditional expectations of square integrable random variables. Considering the measure theoretic definition of these indices and noting the probabilistic setup of generalized polynomial chaos expansions reveal a natural mathematical point of view; moreover, this leads to a direct and intuitive way of deriving spectral representations for the conditional expectations involved and subsequently for the variance based sensitivity indices.

The variance based sensitivity indices can be used to identify model parameters that are most responsible for the model variability. Subsequently, the other, *inessential*, parameters may be fixed at some nominal values to reduce the dimension of the parameter space; the latter can lead to significant reductions in the computational overhead for assessing model uncertainties. Moreover, such simplifications are expected to result in negligible approximation errors. The latter point was noted for example in [24] where some useful error estimates, involving variance based sensitivity indices, were derived. We shall consider this important point and study such error estimates in the case of random variables approximated via generalized polynomial chaos expansions.

The structure of this paper is as follows. In Section 2, we list the basic notation and definitions used throughout the paper. In Section 3, we briefly describe the basics of generalized polynomial chaos expansions in the context of spectral methods for uncertainty analysis. In Section 4, which is devoted to variance based sensitivity analysis, we begin by recalling some fundamental ideas regarding conditional expectation and conditional variance and continue by describing spectral representations of the conditional expectations involved; the discussion in the section then proceeds to definitions of the variance based sensitivity indices and their computation via spectral expansions. Section 5 concerns the approximation errors incurred when inessential variables, characterized as such through a variance based sensitivity analysis, are fixed at nominal values. Finally, in Section 6, we provide some concluding remarks.

2. Basic notation and definitions

In what follows $(\Omega, \mathcal{F}, \mu)$ denotes a probability space. The set Ω is a sample space, $\mathcal{F}$ is an appropriate σ-algebra on Ω, and μ is a probability measure. A real-valued random variable U on $(\Omega, \mathcal{F}, \mu)$ is an $\mathcal{F}/\mathcal{B}(\mathbb{R})$-measurable mapping $U : (\Omega, \mathcal{F}, \mu) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, where $\mathcal{B}(\mathbb{R})$ denotes the Borel σ-algebra on $\mathbb{R}$. Given a random variable U on Ω we denote its expectation and variance by

$$\mathrm{E}\left[U\right] := \int_\Omega U(\omega)\, \mu(d\omega), \qquad \mathrm{Var}\left[U\right] := \mathrm{E}\left[U^2\right] - \mathrm{E}\left[U\right]^2.$$

Denote by $L^2(\Omega, \mathcal{F}, \mu)$ the Hilbert space of (equivalence classes of) real-valued square integrable random variables on Ω; this space is equipped with the inner product $(\cdot, \cdot) : L^2(\Omega) \times L^2(\Omega) \to \mathbb{R}$ given by

$$(U, V) = \mathrm{E}\left[UV\right] = \int_\Omega U(\omega)V(\omega)\, \mu(d\omega), \quad U, V \in L^2(\Omega, \mathcal{F}, \mu),$$

and norm $\|U\|_{L^2(\Omega)} = (U, U)^{1/2}$.

Let $\{\xi_i\}_{i\in I}$ be a collection of random variables on Ω, where I in an index set. We denote by $\sigma(\{\xi_i\}_{i\in I})$ the σ-algebra generated by $\{\xi_i\}_{i\in I}$; recall that $\sigma(\{\xi_i\}_{i\in I})$ is the smallest σ-algebra on Ω with respect to which every ξ_i, $i \in I$, is measurable. In the special case where we have a finite collection of random variables, $\{\xi_i\}_{i=1}^d$, we let $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_d)^T$ and use the short-hand notation $\sigma(\boldsymbol{\xi})$ to denote the σ-algebra generated by $\{\xi_i\}_{i=1}^d$.

We use F_ξ to denote the distribution function [13, 27] of a (real-valued) random variable ξ on $(\Omega, \mathcal{F}, \mu)$:

$$F_\xi(x) = \mu(\xi \le x) = \mu\big(\xi^{-1}(-\infty, x]\big), \quad x \in \mathbb{R}.$$

Recall that F_ξ uniquely characterizes the probability law $\mathcal{L}_\xi = \mu \circ \xi^{-1}$ of the random variable ξ. Moreover, for a function $g : \mathbb{R} \to \mathbb{R}$ such that $g(\xi)$ is integrable, we have

$$\int_\Omega g\big(\xi(\omega)\big)\,\mu(d\omega) = \int_{\mathbb{R}} g(x) F_\xi(dx).$$

3. Spectral methods for uncertainty assessment

Here we provide a brief account of generalized polynomial chaos expansions, with their practical applications in mind. In particular, we consider the the finite-dimensional case involving a finite number of random variables which are used to parameterize uncertainties in a mathematical model. The spectral methods for uncertainty assessment generally utilize the spectral representation of random model observables, which are functions of a finite number of independent random variables, in a polynomial chaos basis. The main motivations to use these spectral representations include efficient sampling, efficient computation of statistical properties (e.g., mean, variance), and, more specific to our discussion, immediate access to variance based sensitivity indices. In this section, we briefly describe the generalized polynomial chaos expansions and some of the related probabilistic setup. For an in-depth coverage of the spectral methods for uncertainty assessment and the related numerical algorithms, we refer to the book [14].

3.1. Generalized polynomial chaos expansions

Consider a finite collection $\xi_1, \ldots, \xi_d$ of independent standard normal random variables on $(\Omega, \mathcal{F}, \mu)$. By (a special case of) the Cameron-Martin Theorem [7] we have that every $U \in L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$ admits an expansion of form,

$$U = \sum_{k=0}^{\infty} c_k \Psi_k(\boldsymbol{\xi}), \tag{3.1}$$

where $\boldsymbol{\xi}(\omega) = \big(\xi_1(\omega), \ldots, \xi_d(\omega)\big)^T$, Ψ_k are d-variate Hermite polynomials [1], and the series converges in $L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$. The expansion (3.1) is known as the *polynomial chaos*[2] expansion (or Wiener-Hermite expansion) [26, 7, 12] of U. The polynomials $\{\Psi_k\}_0^\infty$ form a complete orthogonal set in $L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$:

$$\big(\Psi_k(\boldsymbol{\xi}), \Psi_l(\boldsymbol{\xi})\big) = \delta_{kl} \mathrm{E}\left[\Psi_k^2(\boldsymbol{\xi})\right], \tag{3.2}$$

where δ_{kl} is given by,

$$\delta_{kl} = \begin{cases} 1, & k = l, \\ 0, & k \neq l. \end{cases}$$

[2]The term chaos here is unrelated to the concept of chaos from the theory of nonlinear dynamical systems.

TABLE 1
Orthogonal polynomial bases corresponding to the choice of distribution

distribution of ξ	polynomial basis $\{\psi_k(\xi)\}_{k=0}^{\infty}$	support of $\mathcal{L}_\xi$
Standard normal	Hermite Polynomials	$(-\infty, \infty)$
Uniform	Legendre Polynomials	$[-1, 1]$
Gamma	Laguerre Polynomials	$[0, \infty)$
Beta	Jacobi Polynomials	$[a, b]$

The paper [28], which concerns the practical applications of such spectral representations in scientific computing, discusses choosing ξ_i which follow distributions other than standard normal; this is motivated by the need to provide more flexibility in modeling the parametric uncertainties in physical systems. In such cases, when alternate distributions for ξ_i are considered, the orthogonal polynomial basis $\{\Psi_k(\boldsymbol{\xi})\}_{k=0}^{\infty}$ should be chosen accordingly to achieve optimal convergence. The authors of [28] then note that the Wiener-Askey system of orthogonal polynomials can be used as a guide to choose appropriate polynomial bases which are orthogonal with respect to the distribution law of ξ_i. In the latter case, the expansion in (3.1) is commonly referred to as a *generalized* polynomial chaos expansion. An important theoretical gap was subsequently filled in [9], where the authors provided rigorous convergence results for the generalized polynomial chaos expansions. The results in [9] cover the more general case of the generalized polynomial chaos expansions of functions in $L^2(\Omega, \mathcal{V}, \mu)$, where $\mathcal{V}$ is a σ-algebra generated by a countable collection of independent random variables.

We list in Table 1 the commonly used distributions for the random variables ξ_i and the associated orthogonal polynomial bases [28, 14, 9]. We shall refer to a random variable ξ on Ω for which there exists a orthogonal polynomial basis, $\{\psi_k(\xi)\}_{k=1}^{\infty}$ for $L^2(\Omega, \sigma(\xi), \mu)$ as a *basic random variable*. The random variables following distribution listed in Table 1 are examples of basic random variables.

A d-variate orthogonal polynomial basis is constructed as a tensor product of the univariate orthogonal bases in each coordinate ξ_i, $i = 1, \ldots, d$. Note that it is possible to use ξ_i that are independent but *not necessarily identically distributed*, which leads to a mixed generalized polynomial chaos basis. If we denote by $\psi_k(\xi)$ the k^{th} order polynomial basis function in ξ then the d-variate basis functions Ψ_k are given by,

$$\Psi_k(\boldsymbol{\xi}) = \prod_{j=1}^{d} \psi_{\alpha_j^k}(\xi_j), \quad \boldsymbol{\xi} = (\xi_1, \ldots, \xi_d), \tag{3.3}$$

where $\boldsymbol{\alpha}^k$ is the *multi-index* associated with k^{th} basis function Ψ_k. Here α_j^k is a non-negative integer that specifies the order of the univariate basis polynomial in ξ_j, for $j = 1, \ldots, d$. This multi-index notation will be used extensively throughout this paper. We shall provide a concrete example of the tensor product basis construction when discussing truncated generalized polynomial chaos expansions below. In this paper, we shall focus on the case where ξ_i are continuous random variables whose distributions are chosen from among those listed in Table 1. The case of discrete random variables and their associated bases can be found for example in [28, 9].

3.2. The image probability space

Let $\xi_1, \ldots, \xi_d$ be a collection of independent basic random variables on $(\Omega, \mathcal{F}, \mu)$ as above, and let $F_{\boldsymbol{\xi}}$ denote the joint distribution function of the random d-vector $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_d)$; note that for $\boldsymbol{x} \in \mathbb{R}^d$, $F_{\boldsymbol{\xi}}(\boldsymbol{x}) = \prod_1^d F_j(x_j)$ where F_j is the distribution function corresponding to the j^{th} coordinate. For any random variable $U : (\Omega, \sigma(\boldsymbol{\xi}), \mu) \to \big(\mathbb{R}, \mathcal{B}(\mathbb{R})\big)$, we know by Doob-Dynkin Lemma [13], that there exists a Borel function $X : \mathbb{R}^d \to \mathbb{R}$ such that $U(\omega) = X(\boldsymbol{\xi}(\omega))$. We have $\boldsymbol{\xi}$: $(\Omega, \sigma(\boldsymbol{\xi}), \mu) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), F_{\boldsymbol{\xi}}(d\boldsymbol{x}))$ and $X : (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), F_{\boldsymbol{\xi}}(d\boldsymbol{x})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Thus, instead of working in the abstract probability space $(\Omega, \sigma(\boldsymbol{\xi}), \mu)$, it is sometimes more convenient to work in the probability space, $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), F_{\boldsymbol{\xi}}(d\boldsymbol{x}))$. In fact, letting $\Theta \subseteq \mathbb{R}^d$ denote the support of the law of $\boldsymbol{\xi}$, we may work instead in the *image* probability space $(\Theta, \mathcal{B}(\Theta), F_{\boldsymbol{\xi}}(d\boldsymbol{x}))$.

We denote the expectation of a random variable $X : (\Theta, \mathcal{B}(\Theta), F_{\boldsymbol{\xi}}(d\boldsymbol{x})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ by $\langle X \rangle = \int_\Theta X(\boldsymbol{s})\, dF_{\boldsymbol{\xi}}(\boldsymbol{s})$. The space $L^2(\Theta, \mathcal{B}(\Theta), F_{\boldsymbol{\xi}}(d\boldsymbol{x}))$ is endowed with the inner product $(\cdot,\cdot)_\Theta : L^2(\Theta) \times L^2(\Theta) \to \mathbb{R}$ given by

$$(X, Y)_\Theta = \int_\Theta X(\boldsymbol{s}) Y(\boldsymbol{s})\, dF_{\boldsymbol{\xi}}(\boldsymbol{s}) = \langle XY \rangle, \quad X, Y \in L^2(\Theta, \mathcal{B}(\Theta), F_{\boldsymbol{\xi}}(d\boldsymbol{x})),$$

and norm $\|X\|_{L^2(\Theta)} = (X, X)_\Theta^{1/2}$. For random variables $X(\boldsymbol{\xi})$ and $Y(\boldsymbol{\xi})$ in $L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$, it is immediate to note that $\mathrm{E}\,[X(\boldsymbol{\xi})] = \langle X \rangle$, and that $(X(\boldsymbol{\xi}), Y(\boldsymbol{\xi})) = (X, Y)_\Theta$. Therefore, $\|X(\boldsymbol{\xi})\|^2_{L^2(\Omega)} = \|X\|^2_{L^2(\Theta)}$; that is, $X(\boldsymbol{\xi}) \in L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$ if and only if $X \in L^2(\Theta, \mathcal{B}(\Theta), F(d\boldsymbol{x}))$. Moreover, $\{\Psi_k(\boldsymbol{\xi})\}_{k=0}^\infty$ is a complete orthonormal set in $L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$ if and only if $\{\Psi_k\}_{k=0}^\infty$ is a complete orthonormal set in $L^2(\Theta, \mathcal{B}(\Theta), F(d\boldsymbol{x}))$.

3.3. Truncated expansions

Consider the expansion of $U \in L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$ given in (3.1). Since U is $\sigma(\boldsymbol{\xi})$-measurable, as noted above, there exists a Borel function $X : \mathbb{R}^d \to \mathbb{R}$ such that $U(\omega) = X(\boldsymbol{\xi}(\omega))$ and we have, $X(\boldsymbol{\xi}) = \sum_{k=0}^\infty c_k \Psi_k(\boldsymbol{\xi})$. In practical computations, one approximates $X(\boldsymbol{\xi})$ with a truncated series,

$$X(\boldsymbol{\xi}(\omega)) \doteq \sum_{k=0}^{P} c_k \Psi_k(\boldsymbol{\xi}(\omega)), \tag{3.4}$$

where P depends on the truncation strategy used. There are multiple ways of truncating a multivariate orthogonal polynomial basis. A common approach is truncation based on total polynomial degree. That is, given an expansion order p, one uses a truncated basis $\{\Psi_k(\boldsymbol{\xi}) : |\boldsymbol{\alpha}^k| \le p\}$, where $\boldsymbol{\alpha}^k$ are the multi-indices associated with the d-variate basis functions $\Psi_k(\boldsymbol{\xi})$ as described in (3.3), and

Basis functions	Multi-index
$\Psi_0(\xi_1, \xi_2) = 1$	$\boldsymbol{\alpha}^0 = (0, 0)$
$\Psi_1(\xi_1, \xi_2) = \psi_1(\xi_1)$	$\boldsymbol{\alpha}^1 = (1, 0)$
$\Psi_2(\xi_1, \xi_2) = \psi_1(\xi_2)$	$\boldsymbol{\alpha}^2 = (0, 1)$
$\Psi_3(\xi_1, \xi_2) = \psi_2(\xi_1)$	$\boldsymbol{\alpha}^3 = (2, 0)$
$\Psi_4(\xi_1, \xi_2) = \psi_1(\xi_1)\psi_1(\xi_2)$	$\boldsymbol{\alpha}^4 = (1, 1)$
$\Psi_5(\xi_1, \xi_2) = \psi_2(\xi_2)$	$\boldsymbol{\alpha}^5 = (0, 2)$
$\Psi_6(\xi_1, \xi_2) = \psi_3(\xi_1)$	$\boldsymbol{\alpha}^6 = (3, 0)$
$\Psi_7(\xi_1, \xi_2) = \psi_2(\xi_1)\psi_1(\xi_2)$	$\boldsymbol{\alpha}^7 = (2, 1)$
$\Psi_8(\xi_1, \xi_2) = \psi_1(\xi_1)\psi_2(\xi_2)$	$\boldsymbol{\alpha}^8 = (1, 2)$
$\Psi_9(\xi_1, \xi_2) = \psi_3(\xi_2)$	$\boldsymbol{\alpha}^9 = (0, 3)$

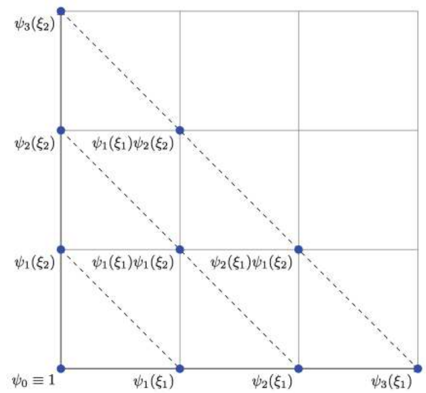

FIG 1. *A bivariate tensor product basis truncated according to total polynomial degree up to three. The table on the left shows the basis functions and the corresponding multi-indices, and the figure on the right provides a visual illustration of the tensor product basis and the truncation. Note that using* (3.5) *with* $p = 3$ *and* $d = 2$ *we have* $1 + P = (3+2)!/(3! \times 2!) = 10$.

$|\boldsymbol{\alpha}^k| = \sum_{j=1}^d \alpha_j^k$. In this case, it is straightforward to show that P in (3.4) is specified by,

$$1 + P = \frac{(d+p)!}{d!p!}. \tag{3.5}$$

See [14] for a construction of an indexing scheme for the multi-indices $\boldsymbol{\alpha}^k$, $k = 0, \dots, P$ which is convenient for computer implementations. To illustrate the tensor-product construction and the truncation strategy described above, we show the construction of a third order ($p = 3$) bivariate ($d = 2$) basis in Figure 1. We point out that in some applications such isotropic truncations may become impractical, and one needs adaptive truncations which exploit the problem structure and choose optimal polynomial orders in different coordinates. However, for simplicity of presentation, in the present work, we consider the total polynomial degree truncation strategy only.

4. Variance based sensitivity analysis

This section is devoted to a detailed study of variance based sensitivity analysis. Since the basic mathematical idea behind variance based sensitivity analysis relies on the concept of conditional expectation, we begin by first recalling some fundamentals regarding conditional expectation and conditional variance in Section 4.1. We then proceed by giving a basic result which enables a spectral approximation of the conditional expectation of a square integrable random variable in Section 4.2. Next, in Section 4.3, starting from the definition of variance based sensitivity indices, we provide their spectral representations in terms of generalized polynomial chaos expansions.

4.1. Conditional expectation and conditional variance

Consider an integrable random variable U on $(\Omega, \mathcal{F}, \mu)$, and consider a sub-$\sigma$-algebra $\mathcal{C}$ of $\mathcal{F}$. The conditional expectation [27] of U with respect to the σ-algebra $\mathcal{C}$, denoted by $\mathrm{E}[U|\mathcal{C}]$ is a $\mathcal{C}$-measurable function such that for every $E \in \mathcal{C}$,

$$\int_E U(\omega)\mu(d\omega) = \int_E \mathrm{E}[U|\mathcal{C}](\omega)\,\mu(d\omega).$$

An intuitive interpretation of the conditional expectation $\mathrm{E}[U|\mathcal{C}]$ is to view it as our best estimate of the random variable U based on the "information content" contained in the σ-algebra $\mathcal{C}$. In the present paper, we consider square integrable random variables, i.e. elements of $L^2(\Omega, \mathcal{F}, \mu)$. In this case, the Hilbert space structure allows defining the conditional expectation as orthogonal projections onto the space $L^2(\Omega, \mathcal{C}, \mu)$. That is, given $U \in L^2(\Omega, \mathcal{F}, \mu)$, the conditional expectation $\mathrm{E}[U|\mathcal{C}]$ is the least-squares best approximation of U in the space $L^2(\Omega, \mathcal{C}, \mu)$. It is also common to talk about conditional expectation with respect to a random variable. In particular, if U and V are random variables on $(\Omega, \mathcal{F}, \mu)$ we write $\mathrm{E}[U|V]$ to mean $\mathrm{E}[U|\sigma(V)]$; recall that $\sigma(V)$ denotes the σ-algebra generated by V.

We also briefly recall the idea of the conditional variance which is defined based on conditional expectation; namely, consider a random variable $U \in L^2(\Omega, \mathcal{F}, \mu)$ and suppose $\mathcal{C} \subseteq \mathcal{F}$ is a sub-σ-algebra, the conditional variance $\mathrm{Var}[U|\mathcal{C}]$ is given by [6]:

$$\mathrm{Var}[U|\mathcal{C}] = \mathrm{E}\left[U^2|\mathcal{C}\right] - \mathrm{E}[U|\mathcal{C}]^2.$$

We also recall the conditional variance formula [6]:

$$\mathrm{Var}[U] = \mathrm{Var}[\mathrm{E}[U|\mathcal{C}]] + \mathrm{E}[\mathrm{Var}[U|\mathcal{C}]]. \tag{4.1}$$

As in the case of conditional expectation, it is common to consider $\mathrm{Var}[U|V]$ where V is random variable on $(\Omega, \mathcal{F}, \mu)$; in this case, $\mathrm{Var}[U|V]$ is understood as $\mathrm{Var}[U|\sigma(V)]$.

4.2. Spectral approximation of conditional expectations

Consider basic random variables $\xi_1, \ldots, \xi_d$ on $(\Omega, \mathcal{F}, \mu)$. We shall, as before, work with random variables $U \in L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$, which as mentioned before can be written as $U = X(\boldsymbol{\xi})$ for a Borel function $X : \mathbb{R}^d \to \mathbb{R}$. To make this dependence on $\boldsymbol{\xi}$ explicit, we refer to elements of $L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$ in the latter form; that is, we say $X(\boldsymbol{\xi}) \in L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$ with the understanding that X is real valued Borel random variable on $\mathbb{R}^d$. Note that $X(\boldsymbol{\xi}) \in L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$ can be expanded in the associated generalized polynomial chaos basis:

$$X(\boldsymbol{\xi}) = \sum_{k=0}^{\infty} c_k \Psi_k(\boldsymbol{\xi}). \tag{4.2}$$

Consider a fixed $i \in \{1, \dots, d\}$ and let y be the conditional expectation $y(\omega) = \mathrm{E}\left[X(\boldsymbol{\xi})|\xi_i\right](\omega)$. As described above y is the orthogonal projection of $X(\boldsymbol{\xi})$ onto $L^2(\Omega, \sigma(\xi_i), \mu)$. Hence, in particular, y is measurable with respect to $\sigma(\xi_i)$ and thus, $y(\omega) = Y(\xi_i(\omega))$ for a real-valued Borel function Y. Moreover, $Y(\xi_i)$ can be expanded in the complete orthogonal basis $\{\psi_k(\xi_i)\}_{\ell=0}^{\infty}$ of $L^2(\Omega, \sigma(\xi_i), \mu)$. That is,

$$Y(\xi_i) = \sum_{\ell=0}^{\infty} d_\ell \psi_\ell(\xi_i). \tag{4.3}$$

Now note that by the tensor product construction of the d-variate basis, the univariate basis $\{\psi_\ell(\xi_i)\}_{\ell=0}^{\infty}$ is a subsequence of the multivariate basis[3] $\{\Psi_k(\boldsymbol{\xi})\}_{k=0}^{\infty}$; that is, $\psi_\ell(\xi_i) = \Psi_{k(\ell)}(\boldsymbol{\xi})$, where $k(\ell) \in \mathbb{Z}^*$ specifies the location of the ℓ^{th} univariate basis function $\psi_\ell(\xi_i)$ in the multivariate basis. Here $\mathbb{Z}^*$ denotes the set of non-negative integers. Next we note that by the definition of orthogonal projection, we have $(Y(\xi_i) - X(\boldsymbol{\xi}), \psi_\ell(\xi_i)) = 0$, for all $\ell \in \mathbb{Z}^*$. Therefore, the expansion coefficients $\{d_\ell\}$ of Y in (4.3) satisfy,

$$d_\ell = \frac{(Y(\xi_i), \psi_\ell(\xi_i))}{(\psi_\ell(\xi_i), \psi_\ell(\xi_i))} = \frac{(X(\boldsymbol{\xi}), \psi_\ell(\xi_i))}{(\psi_\ell(\xi_i), \psi_\ell(\xi_i))} = \frac{\big(X(\boldsymbol{\xi}), \Psi_{k(\ell)}(\boldsymbol{\xi})\big)}{\big(\Psi_{k(\ell)}(\boldsymbol{\xi}), \Psi_{k(\ell)}(\boldsymbol{\xi})\big)} = c_{k(\ell)},$$

for all $\ell \in \mathbb{Z}^*$, where $\{c_k\}$ are the spectral coefficients of $X(\boldsymbol{\xi})$ in (4.2). That is, the coefficients $\{d_\ell\}$ of $Y(\xi_i)$ are a subset of coefficients $\{c_k\}$ of $X(\boldsymbol{\xi})$. Hence, we may write $Y(\xi_i) = \sum_\ell c_{k(\ell)} \Psi_{k(\ell)}(\boldsymbol{\xi})$. Utilizing the tensor product structure of the d-variate basis (3.3) and the multi-index notation, we note that the set $\{k(\ell) : \ell \in \mathbb{Z}^*\}$ which picks the univariate basis functions $\{\psi_\ell(\xi_i)\}$ from $\{\Psi_k(\boldsymbol{\xi})\}$ agrees with the set $\mathcal{E}_i$ defined by

$$\mathcal{E}_i = \{0\} \cup \{k \in \mathbb{N} : \alpha_i^k > 0 \text{ and } \alpha_j^k = 0 \text{ for } j \neq i\}.$$

Thus, it is possible to write the expansion of $Y(\xi_i)$ as,

$$Y(\xi_i) = \sum_{k \in \mathcal{E}_i} c_k \Psi_k(\boldsymbol{\xi}).$$

Moreover, we note that the above developments can be further generalized to consider $\mathrm{E}\left[X(\boldsymbol{\xi})|\{\xi_i\}_{i\in I}\right]$, where I is a subset[4] of $\{1, \dots, d\}$. Repeating an argument similar to the one above we arrive at:

Proposition 4.1. *Suppose, $\xi_1, \dots, \xi_d$ are independent basic random variables on $(\Omega, \mathcal{F}, \mu)$. Let $I \subset \{1, \dots, d\}$ and denote $\boldsymbol{\xi}^{[I]} := \{\xi_i\}_{i\in I}$, and define the index set,*

$$\mathcal{E}_I = \{0\} \cup \{k \in \mathbb{N} : \alpha_i^k > 0 \textit{ for some } i \in I \textit{ and } \alpha_i^k = 0 \textit{ for all } i \notin I\},$$

[3]See also the example in Figure 1 to get an idea of the ordering of the multivariate generalized polynomial chaos basis which is according to increasing total polynomial degree.

[4]To be most precise, we consider $I = \{i_1, \dots, i_r\} \subset \{1, \dots, d\}$ with $i_1 < i_2 < \dots < i_r$.

where $\boldsymbol{\alpha}^k$ are multi-indices associated with the d-variate generalized polynomial chaos basis $\{\Psi_k(\boldsymbol{\xi})\}_{k=0}^\infty$ as described in (3.3). Then, for $X(\boldsymbol{\xi}) \in L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$ the conditional expectation $\mathrm{E}[X(\boldsymbol{\xi})|\boldsymbol{\xi}^{[I]}]$ agrees almost surely with,

$$Y(\boldsymbol{\xi}^{[I]}) = \sum_{k \in \mathcal{E}_I} c_k \Psi_k(\boldsymbol{\xi}^{[I]}), \qquad c_k = \frac{(X(\boldsymbol{\xi}), \Psi_k(\boldsymbol{\xi}))}{(\Psi_k(\boldsymbol{\xi}), \Psi_k(\boldsymbol{\xi}))}.$$

The above result is the main tool used in spectral approximation of variance based sensitivity indices using generalized polynomial chaos expansions.

4.3. Variance based sensitivity indices

This section is devoted to the study of the variance based sensitivity indices [22, 23, 11, 19]. We begin by the definition of the first order, second order, and total sensitivity indices in Section 4.3.1. Then, in Section 4.3.2, we derive the characterization of these indices using generalized polynomial chaos expansions.

4.3.1. The definition of the variance based sensitivity indices

Let $X(\boldsymbol{\xi}) \in L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$. The first order (or main effect) sensitivity indices measure the effect of the i^{th} coordinate ξ_i *alone* on variance of the random variable $X(\boldsymbol{\xi})$. For $i \in \{1, \dots, d\}$, S_i is defined as follows,

$$\mathrm{S}_i = \frac{\mathrm{Var}\left[\mathrm{E}\left[X(\boldsymbol{\xi})|\xi_i\right]\right]}{\mathrm{Var}\left[X(\boldsymbol{\xi})\right]}. \tag{4.4}$$

The second order sensitivity indices describe joint effects. Specifically, for $i, j \in \{1, \dots, d\}$, we define S_{ij} to be the contribution of the *interaction* between ξ_i and ξ_j to the total variance. The mathematical definition of S_{ij} is given by,

$$\mathrm{S}_{ij} = \frac{\mathrm{Var}\left[\mathrm{E}\left[X(\boldsymbol{\xi})|\xi_i, \xi_j\right]\right] - \mathrm{Var}\left[\mathrm{E}\left[X(\boldsymbol{\xi})|\xi_i\right]\right] - \mathrm{Var}\left[\mathrm{E}\left[X(\boldsymbol{\xi})|\xi_j\right]\right]}{\mathrm{Var}\left[X(\boldsymbol{\xi})\right]}. \tag{4.5}$$

Higher order joint sensitivity indices (e.g., S_{ijk}) can be defined also, but are rarely used in applications. Instead, we consider the *total sensitivity index*, which is another useful variance based sensitivity measure. Following [11, 19], for $i \in \{1, \dots, d\}$, we define the total sensitivity index due to ξ_i as,

$$\mathrm{S}_i^{tot} = \frac{\mathrm{E}\left[\mathrm{Var}\left[X(\boldsymbol{\xi})|\boldsymbol{\xi}^{[-i]}\right]\right]}{\mathrm{Var}\left[X(\boldsymbol{\xi})\right]},$$

where $\boldsymbol{\xi}^{[-i]}$ denotes the random vector $\boldsymbol{\xi} = (\xi_1, \dots, \xi_d)$ with ξ_i removed:

$$\boldsymbol{\xi}^{[-i]} := (\xi_1, \dots, \xi_{i-1}, \xi_{i+1}, \dots, \xi_d),$$

in other words, with the notation of Proposition 4.1, $\boldsymbol{\xi}^{[-i]} = \boldsymbol{\xi}^{[I]}$ with $I = \{1, 2, \dots, d\} \setminus \{i\}$. The computation of the total sensitivity indices is facilitated by the following result:

Lemma 4.1. *Let $X(\boldsymbol{\xi}) \in L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$. Then,*

$$\mathrm{S}_i^{tot} = \frac{\mathrm{Var}\left[X(\boldsymbol{\xi})\right] - \mathrm{Var}\left[\mathrm{E}\left[X(\boldsymbol{\xi})|\boldsymbol{\xi}^{[-i]}\right]\right]}{\mathrm{Var}\left[X(\boldsymbol{\xi})\right]}. \tag{4.6}$$

Proof. By the conditional variance formula (4.1), we have,

$$\begin{aligned}\mathrm{Var}\left[X(\boldsymbol{\xi})\right] &= \mathrm{E}\left[\mathrm{Var}\left[X(\boldsymbol{\xi})|\boldsymbol{\xi}^{[-i]}\right]\right] + \mathrm{Var}\left[\mathrm{E}\left[X(\boldsymbol{\xi})|\boldsymbol{\xi}^{[-i]}\right]\right] \\ &= \mathrm{Var}\left[X(\boldsymbol{\xi})\right]\mathrm{S}_i^{tot} + \mathrm{Var}\left[\mathrm{E}\left[X(\boldsymbol{\xi})|\boldsymbol{\xi}^{[-i]}\right]\right]. \end{aligned}$$ □

Remark 4.1. The above result provides an intuitive interpretation of the meaning of a total sensitivity index. Note that the numerator in (4.6) is the total variance minus the variance of the conditional expectation $\mathrm{E}[X(\boldsymbol{\xi})|\boldsymbol{\xi}^{[-i]}]$, which allows quantifying the portion of the variance due to ξ_i. That is, S_i^{tot} is the total contribution of ξ_i, by itself and through its interactions with other coordinates, to the variance.

Remark 4.2. It is also possible to define total sensitivity indices for a subcollection $\{\xi_i\}_{i\in I}$, $I \subseteq \{1, \ldots, d\}$, through,

$$\mathrm{S}_I^{tot} = \frac{\mathrm{Var}\left[X(\boldsymbol{\xi})\right] - \mathrm{Var}\left[\mathrm{E}\left[X(\boldsymbol{\xi})|\boldsymbol{\xi}^{[-I]}\right]\right]}{\mathrm{Var}\left[X(\boldsymbol{\xi})\right]}, \tag{4.7}$$

where $\boldsymbol{\xi}^{[-I]}$ denotes the random vector $\boldsymbol{\xi}$ with coordinates $\{\xi_i\}_{i\in I}$ removed.

4.3.2. Spectral representation of the sensitivity indices

Here we consider spectral approximation of variance based sensitivity indices introduced in the previous section. In practice, given a random variable in $L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$ where $\boldsymbol{\xi}$ is a vector of independent basic random variables, we use its truncated polynomial chaos expansion. Therefore, we state the results of this section for functions $X(\boldsymbol{\xi}) \in \mathcal{V}^p \subseteq L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$, where

$$\mathcal{V}^p = \mathrm{Span}\{\Psi_0(\boldsymbol{\xi}), \Psi_1(\boldsymbol{\xi}), \ldots, \Psi_P(\boldsymbol{\xi})\},$$

which is a assumed to be a sufficiently rich approximation space. Also, we note that Ψ_0 is a constant term, and we use the convention that $\Psi_0(\boldsymbol{\xi}) \equiv 1$. The following result summarizes the rules for computing the variance based sensitivity indices for functions in $\mathcal{V}^p$.

Proposition 4.2. *Let $X(\boldsymbol{\xi}) \in \mathcal{V}^p$. Define the index sets $\mathcal{E}_i$, $\mathcal{J}_{ij}$, and $\mathcal{K}_i$ as follows:*

$$\begin{aligned}&\mathcal{E}_i = \{k \in \{1, \ldots, P\} : \alpha_i^k > 0 \text{ and } \alpha_j^k = 0 \text{ for } j \neq i\}, \\ &\mathcal{J}_{ij} = \{k \in \{1, \ldots, P\} : \alpha_i^k > 0 \text{ and } \alpha_j^k > 0 \text{ and } \alpha_\ell^k = 0 \text{ for } \ell \notin \{i, j\}\}, \\ &\mathcal{K}_i = \{k \in \{1, \ldots, P\} : \alpha_i^k > 0\},\end{aligned}$$

where $\boldsymbol{\alpha}^k$ are multi-indices associated with the d-variate generalized polynomial chaos basis of $\mathcal{V}^p$. Then, S_i, S_{ij} and S_i^{tot} are given by,

$$\mathrm{S}_i = \frac{\sum_{k\in\mathcal{E}_i} c_k^2 \left\|\Psi_k(\boldsymbol{\xi})\right\|^2_{L^2(\Omega)}}{\sum_{k=1}^P c_k^2 \left\|\Psi_k(\boldsymbol{\xi})\right\|^2_{L^2(\Omega)}}, \quad \mathrm{S}_{ij} = \frac{\sum_{k\in\mathcal{J}_{ij}} c_k^2 \left\|\Psi_k(\boldsymbol{\xi})\right\|^2_{L^2(\Omega)}}{\sum_{k=1}^P c_k^2 \left\|\Psi_k(\boldsymbol{\xi})\right\|^2_{L^2(\Omega)}},$$
$$\mathrm{S}_i^{tot} = \frac{\sum_{k\in\mathcal{K}_i} c_k^2 \left\|\Psi_k(\boldsymbol{\xi})\right\|^2_{L^2(\Omega)}}{\sum_{k=1}^P c_k^2 \left\|\Psi_k(\boldsymbol{\xi})\right\|^2_{L^2(\Omega)}}.$$

Proof. First note that for $X(\boldsymbol{\xi}) \in \mathcal{V}^p$, we have, $X(\boldsymbol{\xi}) = \sum_{k=0}^P c_k \Psi_k(\boldsymbol{\xi})$ and thus, by the orthogonality of the basis (and using the convention $\Psi_0 \equiv 1$),

$$\mathrm{Var}\left[X(\boldsymbol{\xi})\right] = \mathrm{E}\left[X(\boldsymbol{\xi})^2\right] - \mathrm{E}\left[X(\boldsymbol{\xi})\right]^2 = \sum_{k=0}^P c_k^2 \mathrm{E}\left[\Psi_k(\boldsymbol{\xi})^2\right] - c_0^2 = \sum_{k=1}^P c_k^2 \left\|\Psi_k(\boldsymbol{\xi})\right\|^2_{L^2(\Omega)}.$$

Now, the proofs of the expressions for S_i and S_{ij} follow immediately from their definition (equations (4.4) and (4.5) respectively) and Proposition 4.1 which gives spectral representations for the conditional expectations involved. As for S_i^{tot}, first note that by Proposition 4.1, we have

$$\mathrm{E}\left[X(\boldsymbol{\xi})|\boldsymbol{\xi}^{[-i]}\right] = \sum_{k\in\{0,\dots,P\}\setminus\mathcal{K}_i} c_k \Psi_k(\boldsymbol{\xi}).$$

Therefore,

$$\begin{aligned} &\mathrm{Var}\left[X(\boldsymbol{\xi})\right] - \mathrm{Var}\left[\mathrm{E}\left[X(\boldsymbol{\xi})|\boldsymbol{\xi}^{[-i]}\right]\right] \\ &\quad= \sum_{k=1}^P c_k^2 \left\|\Psi_k(\boldsymbol{\xi})\right\|^2_{L^2(\Omega)} - \sum_{k\in\{1,\dots,P\}\setminus\mathcal{K}_i} c_k^2 \left\|\Psi_k(\boldsymbol{\xi})\right\|^2_{L^2(\Omega)} \\ &\quad= \sum_{k\in\mathcal{K}_i} c_k^2 \left\|\Psi_k(\boldsymbol{\xi})\right\|^2_{L^2(\Omega)}, \end{aligned}$$

and the expression for S_i^{tot} follows from (4.6). □

Remark 4.3. Note that the index sets $\mathcal{E}_i$, $\mathcal{J}_{ij}$, and $\mathcal{K}_i$ in the above result are determined by the basis of $\mathcal{V}^p$ alone.

Remark 4.4. Note that in view of Proposition 4.1 and Remark 4.2, it is straightforward to derive a spectral representation for S_I^{tot}, where $I \subseteq \{1,\dots,d\}$ specifies a subcollection of the random variables $\xi_1,\dots,\xi_d$.

Remark 4.5. The above result shows that computing sensitivity indices is of trivial computational cost, when a polynomial chaos expansion is available. We point out that in practical applications of spectral methods for uncertainty assessment, the major portion of the computational cost is incurred when computing the expansion coefficients themselves. This issue, which we shall not delve into in the present work, has generated a great amount of research in the recent years. In practice, there exist efficient methods of computing such expansions, albeit in cases of low to moderate parameter dimension. We refer to [14] for a coverage of various strategies for computing polynomial chaos expansions.

In what follows, we also use the notation V_i to denote the total contribution of ξ_i to the variance:

$$\mathrm{V}_i := \mathrm{Var}\left[X\right] - \mathrm{Var}\left[\mathrm{E}\left[X|\boldsymbol{\xi}^{[-i]}\right]\right].$$

Also, to emphasize that the quantities S_i^{tot}, V_i, etc. are computed for a given random variable $X(\boldsymbol{\xi})$, we will denote these quantities by $\mathrm{S}_i^{tot}(X)$, $\mathrm{V}_i(X)$, and so on.

5. Fixing inessential variables and dimension reduction

Consider a function $X(\boldsymbol{\xi}) \in L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$ as before. Suppose a variance based sensitivity analysis is conducted and it is found that one of the variables, say ξ_i, has a very small contribution to the variance of $X(\boldsymbol{\xi})$; that is, S_i^{tot} is "small". It is reasonable to expect that fixing ξ_i at a nominal value will result in a small approximation error. The purpose of this section is to describe estimates of this approximation error. The basic idea behind the developments in this section belongs to [24]. The proofs of the results given below follow in similar lines as the arguments given in [24] (where the authors use Sobol functional decompositions to represent a function of finitely many independent uniformly distributed random variables). The results presented here concern the case of random variables that belong to the space spanned by an appropriate generalized (possibly mixed) polynomial chaos basis; that is we work in $\mathcal{V}^p = \mathrm{Span}\{\Psi_0(\boldsymbol{\xi}), \ldots, \Psi_P(\boldsymbol{\xi})\} \subseteq L^2(\Omega, \sigma(\boldsymbol{\xi}), \mu)$, where as before $\boldsymbol{\xi}$ is a vector of independent basic random variables on Ω.

5.1. A partitioned expansion in the basis of $\mathcal{V}^p$

Consider $X(\boldsymbol{\xi}) \in \mathcal{V}^p$ with a coordinate ξ_i of the random vector $\boldsymbol{\xi}$ fixed at $\xi_i = \vartheta$; we denote this by,

$$X(\boldsymbol{\xi}^{[-i]}; \vartheta) = X(\xi_1, \xi_2, \ldots, \xi_{i-1}, \vartheta, \xi_{i+1}, \ldots, \xi_d).$$

Note that ϑ can be any number in the support of the law of ξ_i. By the tensor-product construction of the basis of $\mathcal{V}^p$, the expansion of $X(\boldsymbol{\xi})$ in this basis,

$$X(\boldsymbol{\xi}) = \sum_{k=0}^{P} c_k \Psi_k(\boldsymbol{\xi}), \tag{5.1}$$

enables the decomposition,

$$X(\boldsymbol{\xi}) = X_0 + \Pi_{\{i\}}\left[X(\boldsymbol{\xi})\right] + \Pi_{\{-i\}}\left[X(\boldsymbol{\xi})\right] + R_X(\boldsymbol{\xi}). \tag{5.2}$$

Here X_0 is the mean of X; moreover, the mapping $\Pi_{\{i\}}[\cdot]$ is the projection of $X(\boldsymbol{\xi})$ in space spanned by $\{\Psi_k(\boldsymbol{\xi})\}_{k\in\mathcal{E}_i}$ with $\mathcal{E}_i$ as defined in Proposition 4.2, $\Pi_{\{-i\}}[\cdot]$ is the projection of $X(\boldsymbol{\xi})$ onto the space spanned by $\{\Psi_k(\boldsymbol{\xi})\}_{k\in I}$ with,

$$I = \{k \in \{1,\dots,P\} : \alpha_i^k = 0\}.$$

and $R_X(\boldsymbol{\xi}) = X(\boldsymbol{\xi}) - X_0 - \Pi_{\{i\}}[X(\boldsymbol{\xi})] - \Pi_{\{-i\}}[X(\boldsymbol{\xi})]$. In other words, $\Pi_{\{i\}}[X(\boldsymbol{\xi})]$ is sum of all the terms in the expansion (5.1) that involve ξ_i only, $\Pi_{\{-i\}}[X(\boldsymbol{\xi})]$ is the sum of all terms in (5.1) that involve $\boldsymbol{\xi}^{[-i]}$ only, and $R_X(\boldsymbol{\xi})$ is the sum of the remaining terms in (5.1).

5.2. Basic estimates

For $i \in \{1,\dots,d\}$, we denote the mean-squared difference between $X(\boldsymbol{\xi}^{[-i]};\vartheta)$ and $X(\boldsymbol{\xi})$ by,

$$\begin{aligned}\delta_X^{(i)}(\vartheta) &= \left\|X(\boldsymbol{\xi}) - X(\boldsymbol{\xi}^{[-i]};\vartheta)\right\|^2_{L^2(\Omega)} \\ &= \int_\Omega \Big(X\big(\boldsymbol{\xi}(\omega)\big) - X\big(\boldsymbol{\xi}^{[-i]}(\omega);\vartheta\big)\Big)^2 \mu(d\omega).\end{aligned} \tag{5.3}$$

Proposition 5.1. *Let $X(\boldsymbol{\xi}) \in \mathcal{V}^p$ and for $i \in \{1,\dots,d\}$ let $\delta_X^{(i)}(\vartheta)$ be defined as in (5.3), with $\vartheta \in \theta_i$, where $\theta_i \subseteq \mathbb{R}$ is the support of the law of ξ_i. Then, the followings hold:*

1. $\delta_X^{(i)}(\vartheta) \ge \mathrm{V}_i(X)$.
2. $\int_{\theta_i} \delta_X^{(i)}(\vartheta)\, dF(\vartheta) = 2\mathrm{V}_i(X)$.

Proof. The decomposition (5.2) of $X(\boldsymbol{\xi})$ gives,

$$X(\boldsymbol{\xi}) = X_0 + U(\xi_i) + W(\boldsymbol{\xi}^{[-i]}) + R_X(\boldsymbol{\xi}),$$

where $U(\xi_i) = \Pi_{\{i\}}[X(\boldsymbol{\xi})]$ and $W(\boldsymbol{\xi}^{[-i]}) = \Pi_{\{-i\}}[X(\boldsymbol{\xi})]$ respectively. Using this decomposition, we have,

$$\mathrm{V}_i(X(\boldsymbol{\xi})) = \mathrm{Var}\left[U(\xi_i)\right] + \mathrm{Var}\left[R_X(\boldsymbol{\xi})\right] \tag{5.4}$$

Also, we can write $X(\boldsymbol{\xi}^{[-i]};\vartheta) = X_0 + U(\vartheta) + W(\boldsymbol{\xi}^{[-i]}) + R_X(\boldsymbol{\xi}^{[-i]};\vartheta)$. We note,

$$\begin{aligned}\delta_X^{(i)}(\vartheta) &= \int_\Omega \Big[X\big(\boldsymbol{\xi}(\omega)\big) - X\big(\boldsymbol{\xi}^{[-i]}(\omega);\vartheta\big)\Big]^2 \mu(d\omega) \\ &= \int_\Omega \Big[U\big(\xi_i(\omega)\big) + R_X\big(\boldsymbol{\xi}(\omega)\big) - U(\vartheta) - R_X\big(\boldsymbol{\xi}^{[-i]}(\omega);\vartheta\big)\Big]^2 \mu(d\omega) \\ &= \int_\Omega \Big[U\big(\xi_i(\omega)\big)^2 + R_X\big(\boldsymbol{\xi}(\omega)\big)^2 + U(\vartheta)^2 + R_X\big(\boldsymbol{\xi}^{[-i]}(\omega);\vartheta\big)^2\Big] \mu(d\omega) \\ &= \mathrm{Var}\left[U(\xi_i)\right] + \mathrm{Var}\left[R_X(\boldsymbol{\xi})\right] + U(\vartheta)^2 + \int_\Omega R_X\big(\boldsymbol{\xi}^{[-i]}(\omega);\vartheta\big)^2 \mu(d\omega) \\ &= \mathrm{V}_i(X) + U(\vartheta)^2 + \int_\Omega R_X\big(\boldsymbol{\xi}^{[-i]}(\omega);\vartheta\big)^2 \mu(d\omega),\end{aligned}$$

where for the third equality, we have used the properties of the basis of $\mathcal{V}^p$, namely orthogonality, tensor product structure, and that $\mathrm{E}[\Psi_k(\boldsymbol{\xi})] = 0$ for $k \geq 1$, and the last equality follows from (5.4). From this, we immediately note that $\delta_X^{(i)}(\vartheta) \geq \mathrm{V}_i(X)$. Next, letting $\Theta^{[-i]}$ denote the support of $F_{\boldsymbol{\xi}^{[-i]}}(d\boldsymbol{x})$, we note that,

$$\int_\Omega R_X\big(\boldsymbol{\xi}^{[-i]}(\omega);\vartheta\big)^2\,\mu(d\omega) = \int_{\Theta^{[-i]}} R_X(\boldsymbol{x};\vartheta)^2\,F_{\boldsymbol{\xi}^{[-i]}}(d\boldsymbol{x}).$$

Thus, letting $\Theta \subseteq \mathbb{R}^d$ denote the support of the joint law $F_{\boldsymbol{\xi}}(d\boldsymbol{s})$, we have,

$$\begin{aligned}\int_{\theta_i} \delta_X^{(i)}(\vartheta)\,dF(\vartheta) &= \mathrm{V}_i(X) + \int_{\theta_i} U(\vartheta)^2 dF(\vartheta) + \int_\Theta R_X(\boldsymbol{s})^2\,F_{\boldsymbol{\xi}}(d\boldsymbol{s})\\ &= \mathrm{V}_i(X) + \int_\Omega U\big(\xi_i(\omega)\big)^2\,\mu(d\omega) + \int_\Omega R_X\big(\boldsymbol{\xi}(\omega)\big)^2\,\mu(d\omega)\\ &= \mathrm{V}_i(X) + \mathrm{Var}\left[U(\xi_i)\right] + \mathrm{Var}\left[R_X(\boldsymbol{\xi})\right] = 2\mathrm{V}_i(X). \qquad \square\end{aligned}$$

Next, as done in [24], we consider the normalized error

$$\Delta_X^{(i)}(\vartheta) = \frac{\delta_X^{(i)}(\vartheta)}{\mathrm{Var}\left[X\right]}.$$

The following result is an immediate consequence of Proposition (5.1):

Corollary 5.1. *For $X(\boldsymbol{\xi}) \in \mathcal{V}^p$, we have $\Delta_X^{(i)}(\vartheta) \geq S_i^{tot}(X)$ and $\int \Delta_X^{(i)}(\vartheta)\,dF(\vartheta) = 2S_i^{tot}(X)$.*

5.3. A probabilistic estimate

According to Corollary 5.1, the the normalized error $\Delta_X^{(i)}(\vartheta)$ is in average in the order of $S_i^{tot}(X)$. As noted in [24], it is possible to take a step further, and use a straightforward argument using Corollary 5.1 and Markov's inequality to derive a probabilistic estimate. In particular, we can state the following result for the case of a square integrable random variable approximated in the space spanned by a generalized polynomial chaos basis.

Proposition 5.2. *Let $X(\boldsymbol{\xi}) \in \mathcal{V}^p$. Then, for any $\beta > 1$, we have*

$$\mathsf{Prob}\Big(\Delta_X^{(i)} \geq \beta S_i^{tot}(X)\Big) \leq 1/(\beta - 1).$$

Proof. Let $\Lambda(\vartheta) = \Delta_X^{(i)}(\vartheta) - \mathrm{S}_i^{tot}(X)$ and note that $\Lambda(\vartheta) \geq 0$. Thus, by Markov's inequality, for the (non-negative) random variable Λ, we have $\mathsf{Prob}(\Lambda \geq \varepsilon) \leq \frac{1}{\varepsilon}\int_{\theta_i} \Lambda(\vartheta)\,dF_{\xi_i}(\vartheta)$ for $\varepsilon > 0$ arbitrary. Moreover, by Corollary 5.1 we have $\int_{\theta_i} \Lambda(\vartheta)\,dF_{\xi_i}(\vartheta) = \mathrm{S}_i^{tot}(X)$. Therefore, we have for any $\varepsilon > 0$,

$$\mathsf{Prob}\Big(\Delta_X^{(i)} - S_i^{tot}(X) \geq \varepsilon\Big) \leq \frac{1}{\varepsilon} S_i^{tot}(X).$$

The statement of the proposition follows by letting $\varepsilon = (\beta - 1)S_i^{tot}(X)$. $\square$

Remark 5.1. Let us note that it is straightforward to generalize Proposition 5.1 and the subsequent results for the case of fixing any subcollection of the coordinates of the random vector $(\xi_1, \ldots, \xi_d)^T$.

6. Concluding remarks

In this paper, we have considered the spectral representation of variance based sensitivity indices from a measure theoretic point of view. This enables a straightforward presentation, which uses spectral representation of conditional expectations as a means of computing variance based sensitivity indices. Moreover, working in the framework of approximations via generalized polynomial chaos expansions, we considered the approximation errors incurred when fixing inessential parameters in a model.

The insight gained from a variance based sensitivity analysis can guide analysis of parametric uncertainties by identifying the parameters most responsible for the variability in a mathematical model. Moreover, such analyses can guide model reduction by fixing inessential parameters. In many physical models, even though the exact value of parameters are difficult to estimate, often there exist widely used values for certain parameters. In such cases, if a variance based sensitivity analysis reveals that such parameters are not influential to model variability, one can consider using the agreed upon nominal values and direct research resources toward more accurate estimation of parameters which are more influential to model variability.

Finally, we mention that practical applications of the idea of variance based sensitivity analysis are abundant in the literature. In addition to the examples given in classical works such as [22, 11, 19, 23, 20], we also point to [25] for an application involving finite element model of a foundation, [4] for an application in ocean circulation modeling under uncertainty, and [3] for a sample application to a biochemical model with random reaction rates. In particular, the application in [4] involves observables which vary over space and time. In that case, it was observed that the balance of sensitivities to different model parameters can change significantly in the space-time domain, as dictated by the physics of the problem.

References

[1] Abramowitz, M. and Stegun, I.A.. *Handbook of Mathematical Functions.* Dover, New York, 9th edition, 1972.

[2] Alexanderian, A., Le Maître, O.P., Najm, H.N., Iskandarani, M., and Knio, O.M. Multiscale stochastic preconditioners in non-intrusive spectral projection. *Journal of Scientific Computing*, 50:306–340, 2012. MR2886330

[3] Alexanderian, A., Rizzi, F., Rathinam, M., Le Maître, O., and Knio, O. Preconditioned Bayesian regression for stochastic chemical kinetics. *Journal of Scientific Computing*, 2013. In press.

[4] Alexanderian, A., Winokur, J., Sraj, I., Srinivasan, A., Iskandarani, M., Thacker, W.C., and Knio, O.M. Global sensitivity analysis in ocean global circulation models: A sparse spectral projection approach. *Computational Geosciences*, 16(3):757–778, 2012.

[5] Archer, G.E.B., Saltelli, A., and Sobol, I.M. Sensitivity measures, ANOVA-like techniques and the use of bootstrap. *Journal of Statistical Computation and Simulation*, 58(2):99–120, 1997.

[6] Athreya, K.B. and Lahiri, S.N. *Measure Theory and Probability Theory*. New York, NY: Springer, 2006. MR2247694

[7] Cameron, R.H. and Martin, W.T. The orthogonal development of non-linear functionals in series of fourier-hermite functionals. *Ann. Math.*, 48:385–392, 1947. MR0020230

[8] Crestaux, T., Le Maitre, O.P., and Martinez, J.-M. Polynomial chaos expansion for sensitivity analysis. *Reliability Engineering & System Safety*, 94(7):1161 – 1172, 2009. Special Issue on Sensitivity Analysis.

[9] Ernst, O.G., Mugler, A., Starkloff, H.-J. and Ullmann, E. On the convergence of generalized polynomial chaos expansions. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46:317–339, 2012. MR2855645

[10] Ghanem, R.G. and Spanos, P.D. *Stochastic Finite Elements: A Spectral Approach*. Dover, 2nd edition, 2002.

[11] Homma, T. and Saltelli, A. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, 52(1):1–17, 1996.

[12] Janson, S. *Gaussian Hilbert Spaces*. Cambridge University Press, 1997. MR1474726

[13] Kallenberg, O. *Foundations of Modern Probability*. Springer, 2nd edition, 2002. MR1876169

[14] Le Maître, O.P. and Knio, O.M. *Spectral Methods for Uncertainty Quantification with Applications to Computational Fluid Dynamics*. Scientific Computation, Springer, 2010. MR2605529

[15] Le Maître, O.P., Mathelin, L., Knio, O.M., and Hussaini, M.Y. Asynchronous time integration for polynomial chaos expansion of uncertain periodic dynamics. *Discrete and Continuous Dynamical Systems*, 28(1):199–226, 2010. MR2629479

[16] Le Maître, O.P., Najm, H.N., Pebay, P.P., Ghanem, R.G., and Knio, O.M. Multi-resolution-analysis scheme for uncertainty quantification in chemical systems. *SIAM Journal on Scientific Computing*, 29(2):864–889, 2007. MR2306272

[17] Najm, H.N. Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics. *Ann. Rev. Fluid Mech.*, 41:35–52, 2009. MR2512381

[18] Reagan, M.T., Najm, H.N., Debusschere, B.J., Le Maître, O.P. Knio, O.M., and Ghanem, R.G. Spectral stochastic uncertainty quantification in chemical systems. *Combustion Theory and Modelling*, 8:607–632, 2004.

[19] Saltelli, A. Sensitivity analysis for importance assessment. *Risk Analysis*, 22(3):579–590, 2002.

[20] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. *Global Sensitivity Analysis: The Primer*. Wiley.com, 2008. MR2382923

[21] Sargsyan, K., Debusschere, B., Najm, H.N., and Le Maître, O.P. Spectral representation and reduced order modeling of the dynamics of stochastic reaction networks via adaptive data partitioning. *SIAM J. Sci. Comput.*, 31:4395–4421, 2010. MR2594987

[22] Sobol, I.M. Estimation of the sensitivity of nonlinear mathematical models. *Matematicheskoe Modelirovanie*, 2(1):112–118, 1990. MR1052836

[23] Sobol, I.M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55(1–3):271–280, 2001. The Second IMACS Seminar on Monte Carlo Methods. MR1823119

[24] Sobol, I.M., Tarantola, S., Gatelli, D., Kucherenko, S.S., and Mauntz, W. Estimating the approximation error when fixing unessential factors in global sensitivity analysis. *Reliability Engineering & System Safety*, 92(7):957–960, 2007.

[25] Sudret, B. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety*, 93(7):964–979, 2008.

[26] Wiener, N. The Homogeneous chaos. *Amer. J. Math.*, 60:897–936, 1938. MR1507356

[27] Williams, D. *Probability with Martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991. MR1155402

[28] Xiu, D.B. and Karniadakis, G.E. The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24:619–644, 2002. MR1951058