Decomposing high-order statistics for sensitivity analysis

By G. Geraci, P.M. Congedo† AND G. Iaccarino

1. Motivation and objectives

Sensitivity analysis in the presence of uncertainties in operating conditions, material properties, and manufacturing tolerances poses a tremendous challenge to the scientific computing community. In particular, in realistic situations, the presence of a large number of uncertain inputs complicates the task of propagation and assessment of output uncertainties; many of the popular techniques, such as stochastic collocation or polynomial chaos, lead to exponentially increasing costs, thus making these methodologies unfeasible (Foo & Karniadakis 2010). Handling uncertain parameters becomes even more challenging when robust design optimization is of interest (Kim *et al.* 2006; Eldred 2009). One of the alternative solutions for reducing the cost of the Uncertainty Quantification (UQ) methods is based on approaches attempting to identify the relative importance of the input uncertainties. In the literature, global sensitivity analysis (GSA) aims at quantifying how uncertainties in the input parameters of a model contribute to the uncertainties in its output (Borgonovo *et al.* 2003). Traditionally, GSA is performed using methods based on the decomposition of the output variance (Sobol 2001), i.e., ANalysis Of VAriance, ANOVA. The ANOVA approach involves splitting a multi-dimensional function into its contributions from different groups of dependent variables. The ANOVA-based analysis creates a hierarchy of dominant input parameters, for a given output, when variations are computed in terms of variance. A limitation of this approach is the fact that it is based on the variance since it might not be a sufficient indicator of the overall output variations.

The main idea of this work is that the hierarchy of important parameters based on second-order statistical moment (as in ANOVA analysis) is not the same if a different statistic is considered (a first attempt in this direction can be found in Abgrall *et al.* 2012). Depending on the problem, the decomposition of the nth-order moment might be more insightful. Our goal is to illustrate a systematic way of investigating the effect of high-order interactions between variables to understand if they are dominant or not. We introduce a general method to compute the decomposition of high-order statistics, then formulate an approach similar to ANOVA but for skewness and kurtosis. This is a fundamental step in order to also formulate innovative optimization methods for obtaining robust designs that account for a complete description of the output statistics. For instance, by knowing the relative importance of each variable (or subset of variables) over the design spaces, reduced UQ propagation problems can be solved adaptively by choosing only the influent variables. A similar approach (Congedo *et al.* 2013) using variance-based sensitivity indices has been demonstrated to be effective in the overall reduction of the numerical cost associated with a design optimization of a turbine blade for Organic Rankine Cycle (ORC) application with a large number of uncertain inputs.

† Inria Bordeaux–Sud-Ouest, France

The methodology proposed in this work illustrates how third- and fourth-order statistic moments can be decomposed (in a way which mimics what has been done for the variance). It is shown that this decomposition is correlated to a polynomial chaos (PC) expansion, enabling us to compute each term and propose new sensitivity indices. The new decomposition technique is illustrated by considering several test functions. In particular, a functional decomposition based on variance, skewness, and kurtosis is computed, displaying how sensitivity indices vary according to the order of the statistical moment. Moreover, the decomposition of high-order statistics is used to drive the model reduction of the metamodel. The effect of the high-order decomposition is also evaluated, for several test cases, in terms of its impact on the probability density functions.

2. High-order statistics definition

Let us consider a real function $f = f(\boldsymbol{\xi})$ with $\boldsymbol{\xi}$ a vector of independent and identically distributed random inputs $\boldsymbol{\xi} \in \Xi^d = \Xi_1 \times \cdots \times \Xi_n$ $(\Xi \subset \mathbb{R}^d)$ and $\boldsymbol{\xi} \in \Xi^d \longmapsto f(\boldsymbol{\xi}) \in L^4(\Xi^d, p(\boldsymbol{\xi}))$, where $p(\boldsymbol{\xi}) = \prod_{i=1}^d p(\xi_i)$ is the probability density function of $\boldsymbol{\xi}$.

The central moments of order n can be defined as

$$\mu^n(f) = \int_{\Xi^d} (f(\boldsymbol{\xi}) - E(f))^n p(\boldsymbol{\xi}) \mathrm{d}\xi, \quad \text{where } E(f) = \int_{\Xi^d} f(\boldsymbol{\xi}) p(\boldsymbol{\xi}) \mathrm{d}\xi. \tag{2.1}$$

In the following, we indicate with $\sigma^2 = \mu^2(f)$, $s = \mu^3(f)$, and $k = \mu^4(f)$ the variance (second-order moment), the skewness (third-order), and the kurtosis (fourth-order), respectively. We note here that according to standard definitions of the skewness and kurtosis, we should include a normalization factor, namely the third power of the standard deviation and the square of the variance, respectively. However, in this context, interest is only in the relative contribution of each term of the decomposition; thus, distorting the nomenclature somewhat, we refer to skewness and kurtosis following the definitions as in Eq. (2.1).

3. Functional ANOVA decomposition

Let us apply the definition of the Sobol functional decomposition (Sobol 2001) to the function f as

$$f(\boldsymbol{\xi}) = \sum_{i=0}^{N} f_{\boldsymbol{m_i}}(\boldsymbol{\xi} \cdot \boldsymbol{m_i}), \tag{3.1}$$

where the multi-index $\boldsymbol{m}$, of cardinality $card(\boldsymbol{m}) = d$, can contain only elements equal to 0 or 1. The total number of admissible multi-indices $\boldsymbol{m_i}$ is $N + 1 = 2^d$; this number represents the total number of contributes up to the dth-order of the stochastic variables $\boldsymbol{\xi}$. The scalar product between the stochastic vector $\boldsymbol{\xi}$ and $\boldsymbol{m_i}$ is employed to identify the functional dependences of $f_{\boldsymbol{m_i}}$. In this framework, the multi-index $\boldsymbol{m}_0 = (0, \ldots, 0)$, is associated with the mean term $f_{\boldsymbol{m}_0} = \int_{\Xi^d} f(\boldsymbol{\xi}) p(\boldsymbol{\xi}) \mathrm{d}\xi$. As a consequence, $f_{\boldsymbol{m}_0}$ is equal to the expectancy of f, i.e., $E(f)$. In the following, we assume the first d indices as the multi-indices associated to the single variables, while the second-order interaction terms follow, and so on.

The decomposition Eq. (3.1) is of ANOVA-type in the sense of Sobol (Sobol 2001) if all the members in Eq. (3.1) are orthogonal, i.e., as

$$\int_{\Xi^d} f_{\boldsymbol{m_i}}(\boldsymbol{\xi}\cdot\boldsymbol{m_i}) f_{\boldsymbol{m}_j}(\boldsymbol{\xi}\cdot\boldsymbol{m}_j) p(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi} = 0 \quad \text{with} \quad \boldsymbol{m_i} \neq \boldsymbol{m}_j, \tag{3.2}$$

and for all the terms $f_{\boldsymbol{m_i}}$, except f_0, it holds

$$\int_{\Xi^d} f_{\boldsymbol{m_i}}(\boldsymbol{\xi}\cdot\boldsymbol{m_i}) p(\xi_j)\mathrm{d}\xi_j = 0 \quad \text{with} \quad \xi_j \in (\boldsymbol{\xi}\cdot\boldsymbol{m_i}). \tag{3.3}$$

Each term $f_{\boldsymbol{m_i}}$ of (3.1) can be expressed as

$$f_{\boldsymbol{m_i}}(\boldsymbol{\xi}\cdot\boldsymbol{m_i}) = \int_{\Xi^{d-\mathrm{card}(\hat{\boldsymbol{m}}_{\boldsymbol{i}})}} f_{\boldsymbol{m_i}}(\boldsymbol{\xi}\cdot\boldsymbol{m_i}) p(\bar{\boldsymbol{\xi}}_i)\mathrm{d}\bar{\boldsymbol{\xi}}_i \quad - \sum_{\substack{\boldsymbol{m}_j \neq \boldsymbol{m_i} \\ \mathrm{card}(\hat{\boldsymbol{m}}_j) < \mathrm{card}(\boldsymbol{m_i})}} f_{\boldsymbol{m}_j}(\boldsymbol{\xi}\cdot\boldsymbol{m}_j), \tag{3.4}$$

where the symbol $\hat{\boldsymbol{m}}_{\boldsymbol{i}}$ indicates a vector of cardinality equal to the number of non-null elements in $\boldsymbol{m_i}$, i.e., $\mathrm{card}(\boldsymbol{m_i})$ indicates the number of variables involved in $\boldsymbol{m}_i$, while $\bar{\boldsymbol{\xi}}_i$ contains all the variables that satisfy $(\boldsymbol{\xi}\cdot\boldsymbol{m_i}) \cup \bar{\boldsymbol{\xi}}_i = \boldsymbol{\xi}$.

Hereafter, we refer, for brevity, to the probability measure: $\mathrm{d}\mu_i = p(\boldsymbol{\xi}\cdot\boldsymbol{m_i})\mathrm{d}(\boldsymbol{\xi}\cdot\boldsymbol{m_i})$. Variance can be expressed as the summation of all the conditional contributions

$$\sigma^2 = \sum_{i=1}^{N} \sigma^2_{\boldsymbol{m_i}}, \quad \text{where} \quad \sigma^2_{\boldsymbol{m_i}} = \int_{\hat{\Xi}_i} f^2_{\boldsymbol{m_i}}(\boldsymbol{\xi}\cdot\boldsymbol{m_i})\mathrm{d}\mu_i. \tag{3.5}$$

The notation is made more compact by means of $\hat{\Xi}_i = \Xi^{\mathrm{card}(\hat{\boldsymbol{m}}_i)}$. Because of the properties of the ANOVA terms, all the mixed contributions are zero due to orthogonality. Analogously, the skewness, first by taking the third power of $f(\boldsymbol{\xi}) - f_0$ and by neglecting the orthogonal contributions, is equal to

$$\begin{aligned} s = \int_{\Xi} (f(\boldsymbol{\xi}) - f_0)^3 \mathrm{d}\mu = & \sum_{p=1}^{N} \int_{\hat{\Xi}_p} f^3_{\boldsymbol{m_p}} \mathrm{d}\mu_p + 3 \sum_{\boldsymbol{m_p}} \sum_{\boldsymbol{m_q} \subset \boldsymbol{m_p}} \int_{\hat{\Xi}_{pq}} f^2_{\boldsymbol{m_p}} f_{\boldsymbol{m_q}} \mathrm{d}\mu_{pq} \\ & + 6 \sum_{p=1}^{N} \sum_{q=p+1}^{N} \sum_{\substack{r=q+1 \\ \boldsymbol{m_{pq}} = \boldsymbol{m_r}}}^{N} \int_{\hat{\Xi}_{pq}} f_{\boldsymbol{m_p}} f_{\boldsymbol{m_q}} f_{\boldsymbol{m_r}} \mathrm{d}\mu_{pq}. \end{aligned} \tag{3.6}$$

In the previous expression, the multi-index $\boldsymbol{m_{pq}}$ represents the union between $\boldsymbol{m_p}$ and $\boldsymbol{m_q}$, also indicated as $\boldsymbol{m_{pq}} = \boldsymbol{m_p} \boxplus \boldsymbol{m_q}$. After some manipulations, it is possible to demonstrate the following (additive) form: $s = \sum_{i=1}^{N} s_{\boldsymbol{m_i}}$. In particular, by considering each multi-index $\boldsymbol{m_i}$ associated with a set of $2^{|\boldsymbol{m_i}|} - 1$ sub-interactions and by denoting this set as $\mathcal{P}_i$ ($\mathcal{P}_{i,\neq}$ is shorthand for $\mathcal{P}_i - \{\boldsymbol{m}_i\}$), each contribution can be expressed as

$$s_{\boldsymbol{m_i}} = \int_{\hat{\Xi}_i} f^3_{\boldsymbol{m_i}} \mathrm{d}\mu_i + 3 \int_{\hat{\Xi}_i} f^2_{\boldsymbol{m_i}} \sum_{\boldsymbol{m_q} \in \mathcal{P}_{i,\neq}} f_{\boldsymbol{m_q}} \mathrm{d}\mu_i + 6 \sum_{\boldsymbol{m_p} \in \mathcal{P}_{i,\neq}} \sum_{\substack{\boldsymbol{m_p} \neq \boldsymbol{m_q} \in \mathcal{P}_{i,\neq} \\ \boldsymbol{m_{pq}} = \boldsymbol{m_i}}} \int_{\hat{\Xi}_i} f_{\boldsymbol{m_i}} f_{\boldsymbol{m_p}} f_{\boldsymbol{m_q}} \mathrm{d}\mu_i. \tag{3.7}$$

Similar considerations lead to the additive form of the kurtosis, where each conditional term is

$$\begin{aligned} k_{\boldsymbol{m_i}} = & \int_{\hat{\Xi}_i} f^4_{\boldsymbol{m_i}} \mathrm{d}\mu_i + 4 \int_{\hat{\Xi}_i} f^3_{\boldsymbol{m_i}} \sum_{\boldsymbol{m_q} \in \mathcal{P}_{i,\neq}} f_{\boldsymbol{m_q}} \mathrm{d}\mu_i + 6 \sum_{\boldsymbol{m_p} \in \mathcal{P}_i} \sum_{\substack{\boldsymbol{m_p} \neq \boldsymbol{m_q} \in \mathcal{P}_i \\ \boldsymbol{m_{pq}} = \boldsymbol{m_i}}} \int_{\hat{\Xi}_i} f^2_{\boldsymbol{m_p}} f^2_{\boldsymbol{m_q}} \mathrm{d}\mu_i \\ & + 12 \sum_{\boldsymbol{m_p}} \sum_{\boldsymbol{m_p} \neq \boldsymbol{m_q} \in \mathcal{P}_i} \sum_{\substack{\boldsymbol{m_r} \in \mathcal{P}_i, r>q \\ \boldsymbol{m_p} \boxplus \cap_{qr} = \boldsymbol{m_i}}} \int_{\hat{\Xi}_i} f^2_{\boldsymbol{m_p}} f_{\boldsymbol{m_q}} f_{\boldsymbol{m_r}} \mathrm{d}\mu_i \\ & + 24 \sum_{\boldsymbol{m_p} \in \mathcal{P}_i} \sum_{\boldsymbol{m_q} \in \mathcal{P}_i, q>p} \sum_{\boldsymbol{m_r} \in \mathcal{P}_i, r>q} \sum_{\substack{t>r, \boldsymbol{m_r} \in \mathcal{P}_i \\ \boldsymbol{m_i} \subseteq \boldsymbol{m_{pq}} \boxplus \cap_{rt} \\ \boldsymbol{m_i} \subseteq \boldsymbol{m_{rt}} \boxplus \cap_{pq}}} \int_{\hat{\Xi}_i} f_{\boldsymbol{m_p}} f_{\boldsymbol{m_q}} f_{\boldsymbol{m_r}} f_{\boldsymbol{m_t}} \mathrm{d}\mu_i. \end{aligned} \tag{3.8}$$

Hereafter, the symbol $\cap_{pq}$ indicates the set of variables contained in both $\boldsymbol{m_p}$ and $\boldsymbol{m_q}$.

4. Correlation with polynomial chaos framework

Variance, skewness, and kurtosis from the functional decomposition are correlated with the terms contained within a polynomial chaos expansion. This correlation establishes a rigorous numerical approach to compute the terms present in the functional (additive) decomposition of the central moments. As usual in the PC framework, a truncated series of $P + 1 = (n_0 + d)!/(n_0!d!)$ terms of total degree n_0 can be obtained as

$$f(\boldsymbol{\xi}) \approx \tilde{f}(\boldsymbol{\xi}) = \sum_{k=0}^{P} \beta_k \Psi_k(\boldsymbol{\xi}), \quad \text{with} \quad \Psi_k(\boldsymbol{\xi} \cdot \boldsymbol{m}^{\star,k}) = \prod_{i=1}^{d} \psi_{\alpha_i^k}(\xi_i). \tag{4.1}$$

Each polynomial $\Psi_k(\boldsymbol{\xi})$ is a multivariate polynomial form which involves tensorization of 1D polynomials by using a multi-index $\boldsymbol{\alpha}^k \in \mathbb{N}^d$, with $\sum_{i=1}^d \boldsymbol{\alpha}_i^k \leq n_0$. The multi-index $\boldsymbol{m}^{\star,k} = \boldsymbol{m}^{\star,k}(\boldsymbol{\alpha}^k) \in \mathbb{N}^d$ is a function of $\boldsymbol{\alpha}^k$: $\boldsymbol{m}^{\star,k} = (m_1^{\star,k}, \ldots, m_d^{\star,k})$, with $m_i^{\star,k} = \alpha_i^k / \left|\left|\alpha_i^k\right|\right|_{\neq 0}$, where the function $\left|\left|\cdot\right|\right|_{\neq 0}$ is defined as

$$\left|\left|\alpha\right|\right|_{\neq 0} = \begin{cases} |\alpha| & \text{if} \quad \alpha \neq 0 \\ 1 & \text{if} \quad \alpha = 0. \end{cases} \tag{4.2}$$

The polynomial basis is chosen according to the Wiener-Askey scheme in order to select orthogonal polynomial terms with respect to the probability density function $p(\boldsymbol{\xi})$ of the input. Thanks to the orthogonality, the coefficients of the expansion in a non-intrusive spectral projection framework are numerically evaluated as

$$\beta_k = \frac{\langle f(\boldsymbol{\xi}), \Psi_k(\boldsymbol{\xi}) \rangle}{\langle \Psi_k(\boldsymbol{\xi}), \Psi_k(\boldsymbol{\xi}) \rangle} \quad \forall k, \tag{4.3}$$

where $\langle \cdot \rangle$ denotes the $L^2(\Xi^d, p(\boldsymbol{\xi}))$ inner product.

For variance decomposition, each conditional term can be computed as

$$\sigma^2_{\boldsymbol{m_i}} = \sum_{k \in K_{\boldsymbol{m_i}}} \beta_k^2 \langle \Psi_k^2(\boldsymbol{\xi}) \rangle, \tag{4.4}$$

where $K_{\boldsymbol{m_i}}$ represents the set of indices associated with the variables $(\boldsymbol{\xi} \cdot \boldsymbol{m_i})$

$$K_{\boldsymbol{m_i}} = \left\{ k \in \{1, \ldots, P\} \,|\, \boldsymbol{m}^{\star,k} = \boldsymbol{m}^{\star,k}(\boldsymbol{\alpha}^k) = \boldsymbol{m_i} \right\}. \tag{4.5}$$

Similar correlations, within PC framework, can also be found for skewness and kurtosis. The final form for skewness is equal to

$$\begin{aligned} s = &\sum_{p=1}^{P} \beta_p^3 \langle \Psi_p^3(\boldsymbol{\xi}) \rangle + 3 \sum_{p=1}^{P} \beta_p^2 \sum_{\substack{q=1 \\ q \neq p}}^{P} \beta_q \langle \Psi_p^2(\boldsymbol{\xi}), \Psi_q(\boldsymbol{\xi}) \rangle \Delta_q^p \\ &+ 6 \sum_{p=1}^{P} \sum_{q=p+1}^{P} \sum_{r=q+1}^{P} \beta_p \beta_q \beta_r \langle \Psi_p(\boldsymbol{\xi}), \Psi_q(\boldsymbol{\xi}) \Psi_r(\boldsymbol{\xi}) \rangle \Delta_{pqr}, \end{aligned} \tag{4.6}$$

where two functions are introduced for the selection, namely

$$\Delta_q^p = \begin{cases} 0 & \text{if} \quad \alpha_j^p = 0 \quad \text{and} \quad \boldsymbol{m_q}_j = 1 \\ 1 & \text{otherwise} \end{cases} \qquad \Delta_{pqr} = \begin{cases} 0 & \text{if} \quad \boldsymbol{m_p}_j + \boldsymbol{m_q}_j + \boldsymbol{m_r}_j = 1, 2 \\ 1 & \text{otherwise.} \end{cases} \tag{4.7}$$

The previous expression reduces, for a fixed $\boldsymbol{m_i}$, to

$$\begin{aligned} s_{\boldsymbol{m_i}} = &\sum_{p \in K_{\boldsymbol{m_i}}} \beta_p^3 \langle \Psi_p^3(\boldsymbol{\xi}) \rangle + 3 \sum_{p \in K_{\boldsymbol{m_p}}} \beta_p^2 \sum_{\substack{q \in K_{\boldsymbol{m_q}} \\ \boldsymbol{m_{pq}} = \boldsymbol{m_i}}} \beta_q \langle \Psi_p^2(\boldsymbol{\xi}), \Psi_q(\boldsymbol{\xi}) \rangle \Delta_q^p \\ &+ 6 \sum_{p \in K_{\boldsymbol{m_p}}} \sum_{\substack{q \in K_{\boldsymbol{m_q}} \\ q \geq p+1}} \sum_{\substack{r \in K_{\boldsymbol{m_r}} \\ \boldsymbol{m_{pqr}} = \boldsymbol{m_i}}} \beta_p \beta_q \beta_r \langle \Psi_p(\boldsymbol{\xi}), \Psi_q(\boldsymbol{\xi}) \Psi_r(\boldsymbol{\xi}) \rangle \Delta_{pqr}. \end{aligned} \tag{4.8}$$

In the case of kurtosis, the final expression reads

$$\begin{aligned} k = &\sum_{p=1}^{P} \beta_p^4 \langle \Psi_p^4(\boldsymbol{\xi}) \rangle + 4 \sum_{p=1}^{P} \beta_p^3 \sum_{\substack{q=1 \\ q \neq p}}^{P} \beta_q \langle \Psi_p^3, \Psi_q \rangle \Delta_q^p \\ &+ 6 \sum_{p=1}^{P} \beta_p^2 \sum_{q=p+1}^{P} \beta_q^2 \langle \Psi_p^2, \Psi_q^2 \rangle + 12 \sum_{p=1}^{P} \beta_p^2 \sum_{\substack{q=1 \\ q \neq p}}^{P} \beta_q \sum_{\substack{r=q+1 \\ r \neq p}}^{P} \beta_r \langle \Psi_p^2, \Psi_q \Psi_r \rangle \Delta_{qr}^p \\ &+ 24 \sum_{p=1}^{P} \sum_{q=p+1}^{P} \sum_{r=q+1}^{P} \sum_{t=r+1}^{P} \beta_p \beta_q \beta_r \beta_t \langle \Psi_p \Psi_q, \Psi_r \Psi_t \rangle \Delta_{pqrt}, \end{aligned} \tag{4.9}$$

where the function Δ_q^p is already introduced in Eq. (4.7), while

$$\Delta_{qr}^p = \begin{cases} 0 & \text{if} \;\; \alpha_j^p = 0 \;\; \text{and} \;\; \boldsymbol{m_q}_j + \boldsymbol{m_r}_j = 1, 2 \\ 1 & \text{otherwise} \end{cases} \tag{4.10}$$

$$\Delta_{pqrt} = \begin{cases} 0 & \text{if} \;\; \boldsymbol{m_p}_j + \boldsymbol{m_q}_j + \boldsymbol{m_r}_j + \boldsymbol{m_t}_j = 1, 2 \\ 1 & \text{otherwise.} \end{cases} \tag{4.11}$$

The conditional contribution associated with $\boldsymbol{m_i}$ is

$$
\begin{aligned}
k_{\boldsymbol{m_i}} = {} & \sum_{k \in K_{\boldsymbol{m_i}}} \beta_k^4 \langle \Psi_k^4(\boldsymbol{\xi}) \rangle + 4 \sum_{p \in K_{\boldsymbol{m_p}}} \beta_p^3 \sum_{\substack{q \in K_{\boldsymbol{m_q}} - \{p\} \\ \boldsymbol{m_p} \boxplus \boldsymbol{m_q} = \boldsymbol{m_i}}} \beta_q \langle \Psi_p^3, \Psi_q \rangle \Delta_q^p \\
& + 6 \sum_{p \in K_{\boldsymbol{m_p}}} \beta_p^2 \sum_{\substack{q \in K_{\boldsymbol{m_q}} - \{p\} \\ \boldsymbol{m_p} \boxplus \boldsymbol{m_q} = \boldsymbol{m_i}}} \beta_q^2 \langle \Psi_p^2, \Psi_q^2 \rangle \\
& + 12 \sum_{p \in K_{\boldsymbol{m_p}}} \beta_p^2 \sum_{q \in K_{\boldsymbol{m_q}} - \{p\}} \beta_q \sum_{\substack{r \in K_{\boldsymbol{m_r}} \\ r \geq q+1 \\ \boldsymbol{m_{pqr}} = \boldsymbol{m_i}}} \beta_r \langle \Psi_p^2, \Psi_q \Psi_r \rangle \Delta_{qr}^p \\
& + 24 \sum_{p \in K_{\boldsymbol{m_p}}} \sum_{\substack{q \in K_{\boldsymbol{m_q}} \\ q \geq p+1}} \sum_{\substack{r \in K_{\boldsymbol{m_r}} \\ r \geq q+1}} \sum_{\substack{t \in K_{\boldsymbol{m_t}} \\ t \geq r+1 \\ \boldsymbol{m_{pqrt}} = \boldsymbol{m_i}}} \beta_p \beta_q \beta_r \beta_t \langle \Psi_p \Psi_q, \Psi_r \Psi_t \rangle \Delta_{pqrt}.
\end{aligned}
\tag{4.12}
$$

As a final remark, with respect to the computational cost of this procedure, it is evident that the number of terms (multidimensional integrals) to compute grows with both the stochastic dimensions and the total polynomial degree. However, the procedure does not require additional model evaluations and can be performed in the post-processing phase just by using the coefficients associated with the polynomial expansion. Moreover, parallel strategies could very easily be implemented, considering that each integral can be computed independently.

4.1. *Additional sensitivity indices and model reduction*

Sensitivity indices (SI) for variance and for skewness and kurtosis can be defined as

$$
\sigma_{\boldsymbol{m_i}}^{2,\mathrm{SI}} = \frac{\sigma_{\boldsymbol{m_i}}^2}{\sigma^2}, \quad s_{\boldsymbol{m_i}}^{\mathrm{SI}} = \frac{s_{\boldsymbol{m_i}}}{s} \quad k_{\boldsymbol{m_i}}^{\mathrm{SI}} = \frac{k_{\boldsymbol{m_i}}}{k}. \tag{4.13}
$$

The first d sensitivity indices are commonly referred as first-order indices because they are associated with the single variables. On the contrary, the remaining terms are associated with the high-order interactions between variables. The sensitivity indices can be used not only to understand the role played by each single variable or groups of variables, but also to guide the choice of a surrogate representation of the function $f(\boldsymbol{\xi})$. For instance, referring to the functional approximation Eq. (4.1), it is possible to obtain a polynomial series which includes only terms of degree up to a fixed order of interaction, namely t

$$
f(\boldsymbol{\xi}) = \sum_{\boldsymbol{m_i}} f_{\boldsymbol{m_i}} \simeq \sum_{\mathrm{card}(\hat{\boldsymbol{m}}_{\boldsymbol{i}}) \leq t} f_{\boldsymbol{m_i}} = \hat{f}(\boldsymbol{\xi}). \tag{4.14}
$$

Henceforth, we refer to the polynomial approximation $\hat{f}(\boldsymbol{\xi})$ as the metamodel of $f(\boldsymbol{\xi})$ with order t. Note that $\mathrm{card}(\hat{\boldsymbol{m}}_{\boldsymbol{i}})$ defines the number of non-null elements in $\boldsymbol{m_i}$, i.e., it measures the order of interaction between variables in the ANOVA sense. Obtaining a metamodel which neglects the interaction terms of order greater than t is strictly related to the idea of effective dimension in the superposition sense introduced in Caflisch *et al.* (1997).

Another measure of sensitivity is the so-called Total Sensitivity Index (TSI) associated with each variable. This measure of sensitivity is computed by including the overall influence of a single variable

$$\mathrm{TSI}_j = \sum_{\xi_j \in (\boldsymbol{\xi}\cdot \boldsymbol{m}_i)} \sigma^{2,\mathrm{SI}}_{\boldsymbol{m}_i}, \qquad \mathrm{TSI}^s_j = \sum_{\xi_j \in (\boldsymbol{\xi}\cdot \boldsymbol{m}_i)} s^{\mathrm{SI}}_{\boldsymbol{m}_i}, \qquad \mathrm{TSI}^k_j = \sum_{\xi_j \in (\boldsymbol{\xi}\cdot \boldsymbol{m}_i)} k^{\mathrm{SI}}_{\boldsymbol{m}_i}. \tag{4.15}$$

The information associated with the TSI can also be exploited for reducing a model. If a threshold is fixed, the dimensionality of the surrogate model can be reduced by neglecting the variables whose TSI is lower than the threshold. In this case, the metamodel choice relies on the definition of effective dimension in the truncation sense (Caflisch *et al.* 1997). In the following section, for brevity, we do not present results related to the reduction of the model in the truncation sense, but rather we present several results on the truncation in the superposition sense. The interested reader can refer to (Gao & Hesthaven 2011; Congedo *et al.* 2013) for further discussions on the importance of the TSI measure for model reduction in the uncertainty quantification and robust optimization settings, respectively.

5. Numerical results

The numerical test cases are chosen to highlight high-order conditional contributions including multiple sources of uncertainties.

5.1. *Computing conditional statistics by means of PC*

In this section, the problem of the computation of high-order conditional terms is analyzed by means of the PC expansion series. Consider the following function

$$f(\boldsymbol{\xi}) = \prod_{i=1}^{d} \sin(\pi \xi_i), \tag{5.1}$$

where each variable $\xi_i \sim \mathcal{U}(0,1)$ with dimension d up to three. Sensitivity indices' (relative) errors are systematically computed with respect to the analytical solution. Conditional statistics can be computed using a PC approach using Eqs. (4.6) and (4.9). In Figure 5.1, we consider the case $d = 2$, and we report the errors in first-order statistics v_1, s_1, and k_1 (where for symmetry $\sigma_1^2 = \sigma_2^2$, $s_1 = s_2$, and $k_1 = k_2$) and interaction terms (v_{12}, s_{12}, k_{12}) computed with respect to the analytical solution. These statistics are well converged at $N = (n_0 + 1)^2 = 121$. In Figures 3, 4, and 5, we consider the case $d = 3$, and percentage errors for conditional statistics are reported. Convergence for conditional statistics is attained at nearly $N = 1500$. The results show that, in this case, all the moments converge in a reasonably consistent way; it is also clear that the dimensionality of the problem has a strong impact on the convergence, especially when high-order contributions are non-negligible.

In the following section, the high-order conditional statistics are employed to show the importance of the high-order interactions between uncertain parameters for the reduction of a numerical model.

5.2. *On the advantages of high-order indices for global sensitivity analysis*

The importance of including the computation of high-order conditional terms in the statistical analysis is demonstrated in this section by means of several test functions.

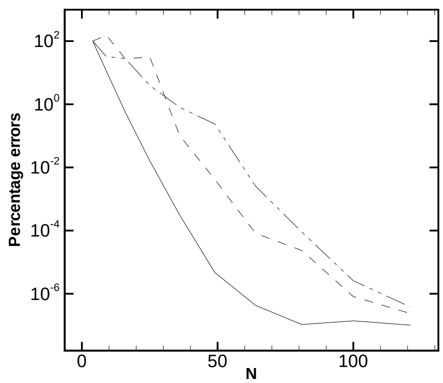

FIGURE 1. First order interactions for Eq. (5.1) with $d = 2$. Variance, continuous line; skewness, dashed line; and kurtosis, dash-dot line.

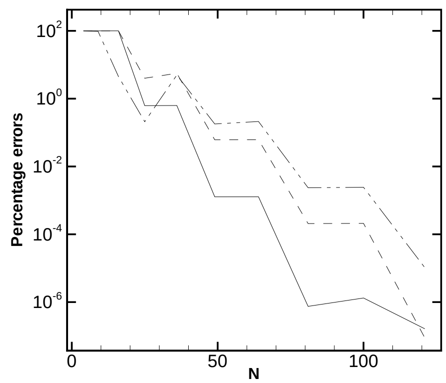

FIGURE 2. Second order interactions for Eq. (5.1) with $d = 2$. Variance, continuous line; skewness, dashed line; and kurtosis, dash-dot line.

Let us consider the classical Sobol function (with four stochastic dimensions)

$$f(\boldsymbol{\xi}) = \prod_{i=1}^{4} \frac{|4\xi_i - 2| + a_i}{1 + a_i}, \tag{5.2}$$

where $\xi_i \sim \mathcal{U}(0,1)$. Two possible choices of the coefficients are considered here: (i) $a_i = (i-1)/2$ the so called linear g-function f_{glin}, or (ii) $a_i = i^2$ the so called quadratic g-function f_{gquad}.

In Figure 6, sensitivity indices for the linear g-function f_{glin} are reported. Several differences can be noticed between the sensitivity indices computed using the variance or other high-order moments. The variance-based ranking illustrates that the first-order sensitivity indices are higher than the second-order one, while these last are higher than those of the third- and fourth-order. This is not the case for skewness and kurtosis, where the second-order contributions are higher than the first-order and third-order ones. This behavior indicates that a variance analysis is able to represent the absolute ranking of the variables in terms of first-order contributions, but the importance associated with higher-order interactions between the parameters is lost. From a practical point of view, underestimating the importance of high-order interactions between variables can lead to wrong decisions in a dimension reduction strategy. The variance based only on first-order contributions exceeds 0.8, whereas skewness and kurtosis do not attain 0.1. This can be demonstrated to be very influential if the probability distribution for reduced models is considered. In Table 1, the total sensitivity indices for the four variables are reported.

The same functional form can lead to slightly different results if the quadratic function coefficients are considered. In Figure 7, the sensitivity indices for the g-function with a quadratic dependence of the coefficients are reported. In this case, the difference between the first order contributions and high-order terms is even more evident. For the variance, first-order contributions exceed 0.98, while a value larger than 0.5 is computed for high-order interactions, when considering skewness and kurtosis. For both skewness and kurtosis, attaining a level equal to 0.8 is only possible including the second order interaction between the first and second variable. In Table 2, total sensitivity indices are

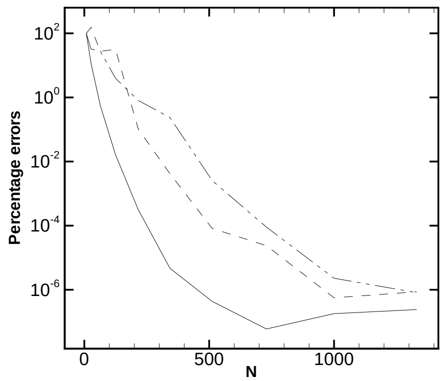

FIGURE 3. First order interactions for Eq. (5.1) with $d = 3$. Variance, continuous line; skewness, dashed line; and kurtosis, dash-dot line.

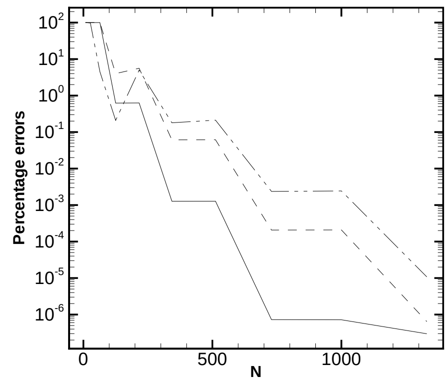

FIGURE 4. Second order interactions for Eq. (5.1) with $d = 3$. Variance, continuous line; skewness, dashed line; and kurtosis, dash-dot line.

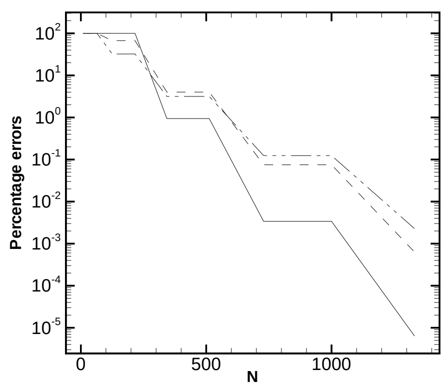

FIGURE 5. Third order interactions for Eq. (5.1) with $d = 3$. Variance, continuous line; skewness, dashed line; and kurtosis, dash-dot line.

Variable	TSI	TSI^s	TSI^k
ξ_1	0.57	0.79	0.86
ξ_2	0.29	0.56	0.64
ξ_3	0.17	0.36	0.44
ξ_4	0.11	0.24	0.31

TABLE 1. Total sensitivity indices for the linear g-function Eq. (5.2) based on a PC series with total degree $n_0 = 5$.

reported for the four variables. In this case, variance contributions for both the third and

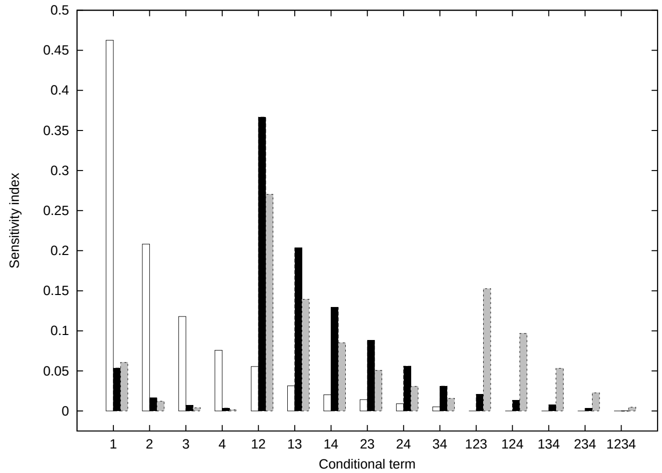

FIGURE 6. Sensitivity indices for the linear g-function f_{glin} Eq. (5.2) obtained with a PC series with total degree $n_0 = 5$. Variance terms are indicated in white, skewness in black and kurtosis in gray.

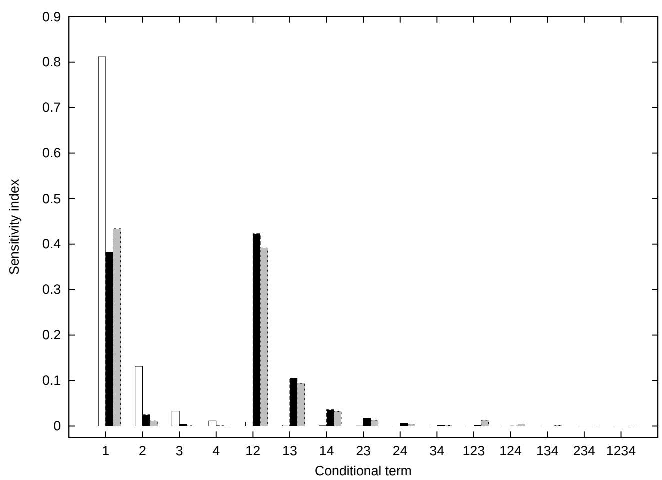

FIGURE 7. Sensitivity indices for the quadratic g-function f_{gquad} Eq. (5.2) obtained with a PC series with total degree $n_0 = 5$. Variance terms are indicated in white, skewness in black, and kurtosis in gray.

fourth variables are below 0.05, while for both skewness and kurtosis, only the fourth variable contribution takes a TSI value of 0.04, which can be considered non significant. A low level of TSI for the variables ξ_3 and ξ_4 could suggest truncating the dimensionality of the model to the first two variables or neglecting the contributions related to an order higher than one.

Let us now consider the following functions

$$f_1 = \xi_1 e^{\frac{\xi_2}{\xi_3^2+1}} + \xi_1 \xi_2 \quad \text{and} \quad f_2 = \prod_{i=1}^{3} \frac{2\xi_i + 1}{2}, \tag{5.3}$$

where the parameters are $\xi_i \sim \mathcal{U}(0,1)$.

Variable	TSI	TSI^s	TSI^k
ξ_1	0.82	0.95	0.97
ξ_2	0.14	0.47	0.44
ξ_3	0.04	0.13	0.12
ξ_4	0.01	0.04	0.04

TABLE 2. Total sensitivity indices for the quadratic g-function f_{gquad} Eq. (5.2) based on a PC series with total degree $n_0 = 5$.

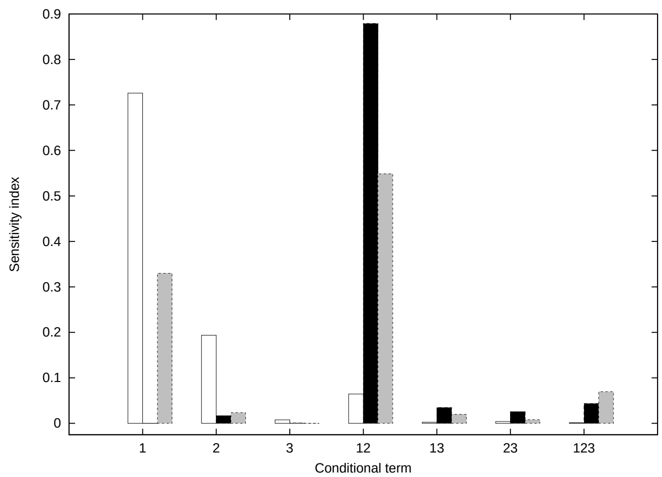

FIGURE 8. Sensitivity indices for the first function f_1 Eq. (5.3) obtained with a PC series with total degree $n_0 = 7$. Variance terms are indicated in white, skewness in black, and kurtosis in gray.

Sensitivity indices associated to the first function f_1 are reported in Figure 8. For the function f_1, the most important variable is ξ_1. For the variance, the first-order sensitivity index relative to ξ_1 is also the most important SI. By contrast, for both skewness and kurtosis, the highest SI is associated to the second-order interaction between the first and the second variable. In this case, inspection of the total sensitivity indices, reported in Table 3, suggests that the third variable ξ_3 does not contribute to the variance. However, if this information is used together with the high-order total sensitivity indices' information, the choice of ignoring the third variable should be considered more carefully. This reflects the importance of ξ_3 in the actual form of the probability density function of f_1, even if its variance is not heavily influenced by it.

The last example, i.e., the function f_2 Eq. (5.3), includes an equal contribution of three variables. However, looking at Figure 9, note that the variance is concentrated only on first-order contributions of the single variables and their sum exceeds 0.9. The skewness and kurtosis contributions, on the other hand, are concentrated in second-order interactions. For kurtosis, the third-order interaction is the highest contribution. Note that even if the sum of the first-order variance contribution exceeds 0.9, a reduction of the model that neglects the high orders of interaction could lead to wrong conclusions. In this case, the skewness associated to the first-order metamodel does not include any

Variable	TSI	TSI^s	TSI^k
ξ_1	0.79	0.96	0.97
ξ_2	0.26	0.96	0.67
ξ_3	0.02	0.10	0.10

TABLE 3. Total sensitivity indices for the first function f_1 Eq. (5.3) based on a PC series with total degree $n_0 = 7$.

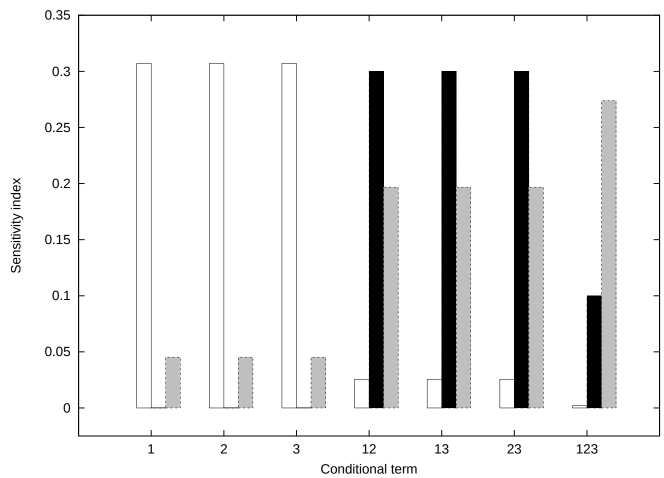

FIGURE 9. Sensitivity indices for the function f_2 Eq. (5.3) obtained with a PC series with total degree $n_0 = 7$. Variance terms are indicated in white, skewness in black, and kurtosis in gray.

Variable	TSI	TSI^s	TSI^k
ξ_1	0.36	0.70	0.71
ξ_2	0.36	0.70	0.71
ξ_3	0.36	0.70	0.71

TABLE 4. Total sensitivity indices for the first function f_2 Eq. (5.3) based on a PC series with total degree $n_0 = 7$.

information associated to the non-symmetric behavior of the probability distribution of the function f_2.

Values for the total sensitivity indices for this case are reported in Table 4 . Note that the sum of the total sensitivity indices over the three variables is much higher for skewness and kurtosis with respect to the variance. Then, both of them refer, correctly, to an intrinsically high-order (of interaction) function (see Eq. (5.3) for the definition of f_2).

From a practical point of view, the information related to the high-order interactions can be exploited for driving a model reduction. With regard to the function f_{glin} (see Figure 6 and Table 1), the partial contribution on the variance of the first-order interaction, since it exceeds 0.8, could lead to the decision to build a metamodel including

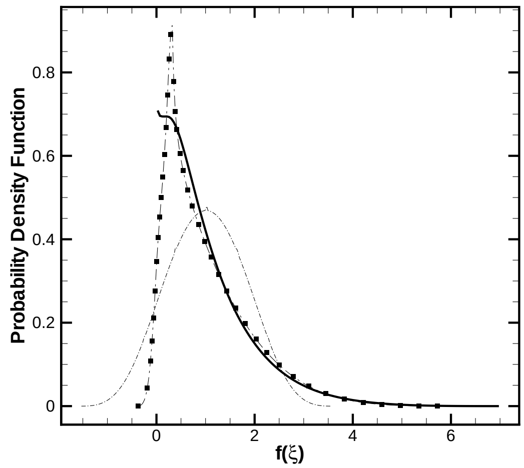

FIGURE 10. PDFs for the complete linear g-function f_{glin} (bold line) and the reduced models: first (dash-dot) and second order (dashed with symbols).

only the first order, i.e., four ANOVA terms. However, the high-order SI indicate the importance of, at least, second-order interactions. The comparison, in term of probability density function of f_{glin}, is reported in Figure 10. It is evident that the first-order model has zero skewness. The situation greatly improves by including contributions up to second order.

A similar behavior can be observed by applying a model reduction to the function f_{gquad}. Even if the variance is almost fully captured by the first-order contributions of the ANOVA expansions (see Figure 7), the information relative to the skewness is mostly related to the second-order contributions. This is evident in terms of the probability distributions, as is shown in Figure 11.

Moreover, the function f_1 features a different behavior in terms of sensitivity indices. The skewness is associated to the second-order interaction between the first and the second variable. The effect of neglecting the second-order terms is evident in Figure 12 where the probability distributions are shown. The first-order metamodel completely misrepresents the tails of the distribution, whereas a great improvement is associated to the the second order terms. The right tail is well captured, but the model still fails to capture correctly the left tail.

In Figure 13, the pdf for the complete model and the first and second orders are reported. Even if more than 90% of the variance is included in the first-order model, its pdf contains no information about the skewness, and the tails appear to be totally lost. However, if the second-order interactions between variables are included, the quality of the pdf improves consistently.

Numerical test cases presented in this section illustrate how information relative to variance-based sensitivity indices seems to be incomplete if the true dependence of a model from its variables is to be understood. In particular, variance gives more weight to low-order interactions with respect to the sensitivity indices associated to skewness and kurtosis. This factor could be even more important if the aim is to reduce the dimensionality of the problem and to build an accurate metamodel.

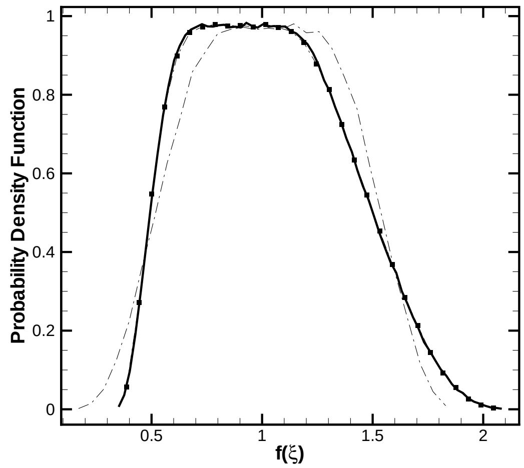

FIGURE 11. PDFs for the complete quadratic g-function f_{gquad} (bold line) and the reduced models: first (dash-dot) and second order (dashed with symbols).

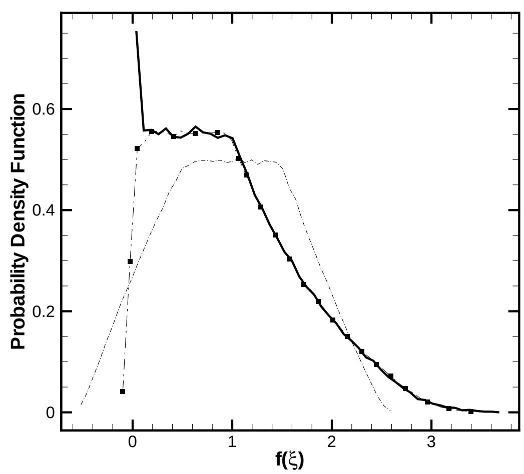

FIGURE 12. PDFs for the complete f_1 (bold line) and the reduced models: first (dash-dot) and second order (dashed with symbols).

6. Conclusions and future perspectives

This study is focused on analysis of the decomposition of high-order statistical moments of multi-variate stochastic functions, obtained as results of problems subject to uncertainty in the inputs. A correlation was found between functional decomposition, as depicted by Sobol, and polynomial chaos expansion. This allows each term in the decomposition to be clearly defined and also provides a practical way to compute all these terms. This procedure is assessed on several analytic test cases computing the conver-

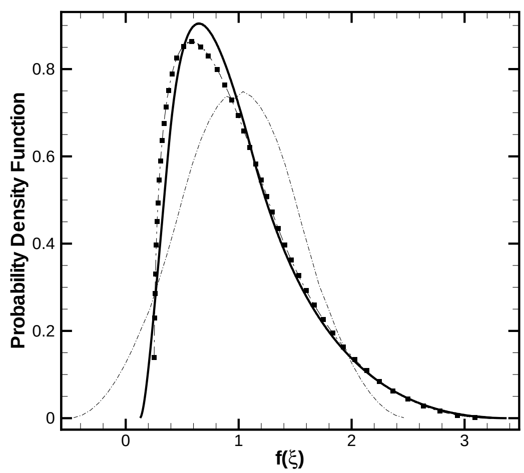

FIGURE 13. PDFs for the complete f_2 (bold line) and the reduced models: first (dash-dot) and second order (dashed with symbols).

gence curves obtained by using PC. Furthermore, sensitivity indices based on skewness and kurtosis decomposition are introduced. The importance of ranking the predominant uncertainties in terms not only of the variance but also of higher-order moments (then extending the ANOVA analysis also to higher-order statistic moments) was demonstrated.

Future plans will be directed towards the application of high-order sensitivity indices' information, as already done for the variance, for adaptive dimensional reduction in the context of robust design optimization. Moreover, efforts to estimate, a priori, the quality of a polynomial metamodel, in which the high-order interactions cannot be accurately assessed, are currently under investigation.

Acknowledgments

The authors acknowledge Mike Emory for insightful comments on a preliminary draft of this work. The associated team AQUARIUS (Inria-UQ Lab Stanford) is also acknowledged for financial support.

REFERENCES

ABGRALL, R., CONGEDO, P. M., GERACI, G. & IACCARINO, G. 2012 *Decomposition of high-order statistics.* Technical Report, INRIA.

BORGONOVO, E., APOSTOLAKIS, G. E., TARANTOLA, S. & SALTELLI, A. 2003 Comparison of global sensitivity analysis techniques and importance measures in PSA. *Reliab. Eng. Syst. Safe.* **79**, 175–185.

CAFLISCH, R. E., MOROKOFF, W. & OWEN, A. B. 1997 Valuation of Mortgage Backed Securities Using Brownian Bridges to Reduce Effective Dimension. *J. Comput. Finance* **1**, 26–47.

CONGEDO, P. M., GERACI, G., ABGRALL, R., PEDIRODA, V. & PARUSSINI, L. 2013

TSI metamodels-based multi-objective robust optimization. *Eng. Computation* **30**, 1032–1053.

Eldred, M. S. 2009 Recent advances in non-intrusive polynomial chaos and stochastic collocation methods for uncertainty analysis and design. *AIAA Paper 2009–2274*.

Foo, J. & Karniadakis, G. E. 2010 Multi-element probabilistic collocation method in high dimensions. *J. Comput. Phys.* **229**, 1536–1557.

Gao, Z. & Hesthaven, J. S. 2011 Efficient solution of ordinary differential equations with high-dimensional parametrized uncertainty. *Commun. Comput. Phys.* **10**, 253–278.

Kim, N. H., Wang, H. & Queipo, N. V. 2006 Efficient shape optimization under uncertainty using polynomial chaos expansions and local sensitivities. *AIAA J.* **44**, 1112–1115.

Sobol, I. 2001 Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comp. Simulat.* **55**, 271–280.