High-order statistics in global sensitivity analysis: decomposition and model reduction

Gianluca Geraci, Pietro Marco Congedo, Remi Abgrall, Gianluca Iaccarino

High-order statistics in global sensitivity analysis: decomposition and model reduction

G. Geraci[a,*], P.M. Congedo[b,*], R. Abgrall[c], G. Iaccarino[a]

[a] *Flow Physics and Computational Engineering, Stanford University, 488 Escondido Mall, CA 94305-3035, USA*
[b] *INRIA Bordeaux–Sud-Ouest, Equipes Bacchus, 200, Avenue de la Vieille Tour, 33400 Talence, France*
[c] *Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*

Abstract

ANalysis Of VAriance (ANOVA) is a common technique for computing a ranking of the input parameters in terms of their contribution to the output variance. Nevertheless, the variance is not an universal criterion for ranking variables, since non symmetric outputs could require higher order statistics for their description and analysis. In this work, we illustrate how third and fourth-order moments, *i.e.* skewness and kurtosis, respectively, can be decomposed mimicking the ANOVA approach. It is also shown how this decomposition is correlated to a Polynomial Chaos (PC) expansion leading to a simple strategy to compute each term. New sensitivity indices, based on the contribution to the skewness and kurtosis, are proposed. The outcome of the proposed analysis is depicted by considering several test functions. Moreover, the ranking of the sensitivity indices is shown to vary according to their statistics order. Furthermore, the problem of formulating a truncated polynomial representation of the original function is treated. Both the reduction of the number of dimensions and the reduction of the order of interaction between parameters are considered. In both cases, the impact on the reduction is assessed in terms of statistics, namely the probability density function. Feasibility of the proposed analysis in a real-case is then demonstrated by presenting the sensitivity analysis of the performances of a turbine cascade in an Organic Rankine Cycles (ORCs), in the presence of complex thermodynamic models and multiple sources of uncertainty.

Keywords: sensitivity analysis, high-order statistics, skewness, kurtosis, polynomial chaos, model reduction, ORCs turbines.

1. Introduction

Optimization and design in the presence of uncertain operating conditions, material properties and manufacturing tolerances poses a tremendous challenge to the scientific computing community. Uncertainty quantification (UQ) approaches represent the inputs as random variables and seek to construct a statistical characterization of the quantities of interest. Several methodologies are proposed to tackle this issue, among those stochastic spectral methods [1, 2, 3, 4, 5], that can provide considerable speed-up in computational time when compared to Monte Carlo (MC) simulation. The presence of a large number of uncertain inputs leads to an exponential increase of the cost thus making these methodologies unfeasible [6]. This situation becomes even more challenging when robust design optimization is tackled [7, 8].

Several UQ methods have been developed with the objective of reducing the number of solutions required to obtain a statistical characterization of the quantity of interest, such as Sparse Grid techniques [9] or adaptive mesh generation [10, 11, 12, 13, 14]. These techniques can lead to a dramatical reduction of the quadrature points for moderate dimensional problem, provided that the function has some regularity properties. Classical sparse grids [9] are constructed from tensor products of one-dimensional quadrature formulas. Some Galerkin-based methods deal with multi-resolution wavelet expansions [15, 16], domain decomposition in the random space [17], adaptive h-refinement [3] for dealing with arbitrary probability distributions.

[*] *E-mail:* ggeraci@stanford.edu (Gianluca Geraci), pietro.congedo@inria.fr (Pietro Marco Congedo)

Among the collocation-based stochastic spectral methods, in [18] the authors proposed the use of sparse grid quadrature for stochastic collocation. Older studies show the errors and efficiency of sparse grid integration and interpolation [19, 20], Smolyak constructions based on one-dimensional nested Clenshaw-Curtis rules [19, 21] and the integration error of sparse grids based on one-dimensional Kronrod-Patterson rules [22].

An alternative solution is based on approaches attempting to identify the relative importance of the input uncertainties on the output. If some dimensions could be identified as negligible, they could be discarded in a reduced stochastic problem and better statistics estimations could be achieved with a lower computational cost.

Identifying the most influent parameters requires to determine the uncertain inputs which have the largest impact on the variability of the model output. In literature, Global sensitivity analysis (GSA) aims at quantifying how uncertainty in the input parameters of a model contributes to the uncertainty in its output (see for example [23]) where global sensitivity analysis techniques are applied to probabilistic safety assessment models). GSA classifies the inputs according to their importance on the output variations and it gives a hierarchy of the most important ones.

Traditionally, GSA is performed using methods based on the decomposition of the output variance [24], *i.e.* ANOVA. The ANOVA approach involves splitting a multi-dimensional function into its contributions from different groups of subdimensions. In 2001, Sobol used this formulation to define global sensitivity indices [24], displaying the relative variance contributions of different ANOVA terms. In [25], the authors introduced two High-Dimensional Model Reduction (HDMR) techniques to capture input-output relationships of physical systems with many input variables. These techniques are also based on ANOVA decompositions.

Several techniques have been developed to compute sensitivity indices at low computational cost [26]. In [27, 28, 29], generalized Polynomial Chaos Expansions (gPC) are used to build surrogate models for computing the Sobol's indices analytically as a post-processing of the PC coefficients. In [6], the authors combine multi-element polynomial chaos with an ANOVA functional decomposition to enhance the convergence rate of polynomial chaos in high dimensions and in problems with low stochastic regularity. In [30], the use of adaptive ANOVA decomposition is investigated as an effective dimension-reduction technique in modeling incompressible and compressible flows with high-dimension of random space. In [31], sparse Polynomial Chaos (PC) expansions are introduced in order to compute sensitivity indices: a PC-based metamodel (a computationally inexpensive polynomial approximation of the relation between inputs and output) which contains the significant terms whereas the PC coefficients are computed by least-square regression.

Other approaches are developed if the assumption of independence of the input parameters is not valid. For instance, new indices have been proposed in [32, 33] when a linear correlation exists. In [34], the authors introduce a global sensitivity indicator which looks at the influence of input uncertainty on the entire probability distribution without reference to a specific moment of the output (moment independence) and which can be defined also in the presence of correlations among the parameters. In [35], a gPC methodology to address global sensitivity analysis for this kind of problems is introduced, while in [36], a numerical procedure is proposed for moment-independent sensitivity methods.

The ANOVA-based analysis create a hierarchy of most important input parameters for a given output when variance is chosen as metrics. A strong limitation of this approach is the fact that the variance can not be considered a general indicator for a complete description of output variations. Roughly speaking, the variance measures the spread of a set of realizations for a random variable, therefore the information related to the direction of this spreading are completely missed. As a result, computation of Higher-Order (HO) statistics is of primary importance, for example the normalized third order, the *skewness* measures of the non-symmetry of the distribution or the normalized fourth order, the *kurtosis* measures whether the distribution is peaked or flat. Depending on the problem, a $n-$order statistics breakdown or, in the following decomposition, could be of interest. Moreover, it seems of primary importance to collect the set of interactions obtained from $n-$order statistics decomposition for a correct ranking of all the contributions.

For computing HO statistics, the most diffused methods are related to Monte Carlo and quasi-Monte Carlo approaches. Very few papers exist showing the application of polynomial-chaos techniques to the computation of HO statistics [37, 38].

First objective of this paper is to provide a general method in order to compute the decomposition of high-order statistics, then to formulate an approach similar to ANOVA but for *skewness and kurtosis*. One of the main results of the paper is to show how skewness and kurtosis can be written mimicking the variance decomposition in term of conditional statistics, *i.e.* an additive functional relationship always holds. The idea is to compute the most influential interactions not only for the variance but also for higher statistics permitting to improve the sensitivity analysis. This

is a fundamental step in order to formulate also innovative optimization methods for obtaining very robust designs by taking into account a complete description of the output statistics. Second objective is to illustrate the correlation between the high-order functional decomposition and the PC-based techniques, thus displaying how to compute each term from a numerical standpoint. Moreover, two reduction strategies, for the resulting polynomial metamodel, are considered. A classical truncation which neglects some non-significant stochastic dimensions and a reduction based on the analysis of the order of interactions. These strategies are evaluated with respect to the complete non-reduced model in terms of their effect on the probability density function of the output.

Several numerical test cases are proposed to demonstrate the interest of the proposed approach. Among them a thermodynamically complex flow in a ORCs turbine cascade, characterized by a significant uncertainty on the physical parameters and on the operating conditions at the turbine inlet [39], is analyzed. Three sources of uncertainties [40] are taken into account, namely the thermophysical properties, the inlet boundary conditions and the geometrical parameters of the blade.

The remaining paper is organized as follows. In Section 2, functional decomposition for variance, skewness and kurtosis are presented. In Section 3, the correlation between the functional decomposition and a Polynomial Chaos framework is depicted. Section 4 extends the sensitivity index definitions to high-order statistics and the reduction strategies for a polynomial metamodel are also presented. Several results, showing how the Polynomial Chaos expansion can be used practically to compute high-order statistics and the importance of considering skewness and kurtosis sensitivity indices when ranking the uncertainties/interactions, are presented in Section 5. Moreover in Section 5, the sensitivity analysis of a ORCs turbine is also presented. This numerical investigation aims to display the interest of the proposed approach on a real engineering problem. Eventually, conclusions and future perspectives are drawn in Section 6.

2. Functional decomposition

Let us consider a real function $f = f(\vec{\xi})$ with $\vec{\xi}$ a vector of independent and identically distributed random inputs $\vec{\xi} \in \Xi^d = \Xi_1 \times \cdots \times \Xi_n$ ($\Xi^d \subset \mathbb{R}^d$) and $\vec{\xi} \in \Xi^d \longmapsto f(\vec{\xi}) \in L^4(\Xi^d, \mathrm{d}\mathbb{P})$. In the following, we will assume the existence of a density function for the probability measure $\mathrm{d}\mathbb{P} = p(\vec{\xi})\mathrm{d}\vec{\xi}$, where $p(\vec{\xi}) = \prod_{i=1}^{d} p(\xi_i)$ is the joint probability density function. We note here that the Sobol' ANOVA decomposition, as well as the Polynomial Chaos approach, requires $f(\vec{\xi}) \in L^2(\Xi^d, \mathrm{d}\mathbb{P})$, however we are interested in defining statistics at least up to the fourth order. Moreover, the random inputs are assumed i.i.d. only for simplicity of exposure.

Let us introduce the definition of the Sobol functional decomposition [24] for the function f as

$$f(\vec{\xi}) = \sum_{i=0}^{N} f_{\vec{m}_i}(\vec{\xi} \cdot \vec{m}_i), \tag{1}$$

where the generic multi-index $\vec{m}$, of cardinality $card(\vec{m}) = d$, can contain only elements equal to 0 or 1. The total number of admissible multi-indices $\vec{m}_i$ is $N + 1 = 2^d$; this number represents the total number of contributions up to the dth-order of the stochastic variables $\vec{\xi}$. The scalar product between the stochastic vector $\vec{\xi}$ and $\vec{m}_i$ is a shorthand employed to identify the functional dependency of $f_{\vec{m}_i}$. In the following, the multi-index $\vec{m}_0 = (0, \ldots, 0)$, is associated to the mean term $f_0 = \mathbb{E}(f)$. The remaining N multi-indices are ordered as follows

$$\begin{aligned}
\vec{m}_1 &= (1, 0, \ldots, 0) \\
\vec{m}_2 &= (0, 1, \ldots, 0) \\
&\vdots \\
\vec{m}_d &= (0, \ldots, 1) \\
\vec{m}_{d+1} &= (1, 1, 0, \ldots, 0) \\
\vec{m}_{d+2} &= (1, 0, 1, 0, \ldots, 0) \\
&\vdots \\
\vec{m}_N &= (1, \ldots, 1).
\end{aligned} \tag{2}$$

This ordering is assumed for convenience and does not affect in any way the successive ANOVA functional decomposition.

The decomposition (1) is of ANOVA-type in the sense of Sobol [24] if all the members in Eq. (1) are orthogonal with respect to the joint pdf

$$\int_{\Xi^d} f_{\vec{\boldsymbol{m}}_i}(\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_i) f_{\vec{\boldsymbol{m}}_j}(\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_j) \mathrm{d}\mathbb{P} = 0 \quad \text{with} \quad \vec{\boldsymbol{m}}_i \neq \vec{\boldsymbol{m}}_j, \tag{3}$$

and for all the terms $f_{\vec{\boldsymbol{m}}_i}$, except f_0 (f_0 is used in the following as shorthand for $f_{\vec{\boldsymbol{m}}_0}$), it holds

$$\int_{\Xi_j} f_{\vec{\boldsymbol{m}}_i}(\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_i) p(\xi_j) \mathrm{d}\xi_j = 0 \quad \text{with} \quad \xi_j \in \left(\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_i\right). \tag{4}$$

Each term $f_{\vec{\boldsymbol{m}}_i}$ of (1) can be expressed as follows

$$f_{\vec{\boldsymbol{m}}_i}(\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_i) = \int_{\Xi^{d-\mathrm{nnz}_i}} f(\vec{\boldsymbol{\xi}}) p(\vec{\boldsymbol{\xi}}'_i) \mathrm{d}\vec{\boldsymbol{\xi}}'_i \quad - \sum_{\vec{\boldsymbol{m}}_j \subset \vec{\boldsymbol{m}}_i} f_{\vec{\boldsymbol{m}}_j}(\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_j), \tag{5}$$

where nnz_i is used as shorthand to count the number of non-null elements in $\vec{\boldsymbol{m}}_i$ and $\vec{\boldsymbol{\xi}}'_i$ represents the complement of $\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_i$ to $\vec{\boldsymbol{\xi}}$, *i.e.* $(\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_i) \cup \vec{\boldsymbol{\xi}}'_i = \vec{\boldsymbol{\xi}}$. Note that $\vec{\boldsymbol{m}}_0$ has null cardinality, therefore is always included in the series at the right hand side of Eq. (5).

Hereinafter, in order to substantially reduce the complexity of the notation, the integrals are written with respect the probability measure (relative to the multi-index $\vec{\boldsymbol{m}}_i$):

$$\mathrm{d}\mathbb{P}_i = p(\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_i) \mathrm{d}(\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_i), \tag{6}$$

where multiple, two or three, subscripts are adopted to denote the joint probability measure obtained by tensorization of the measures relative to each multi-indices, as for instance $\mathrm{d}\mathbb{P}_{ij} = \mathrm{d}\mathbb{P}_i \, \mathrm{d}\mathbb{P}_j$ or $\mathrm{d}\mathbb{P}_{ijk} = \mathrm{d}\mathbb{P}_i \, \mathrm{d}\mathbb{P}_j \, \mathrm{d}\mathbb{P}_k$.

The functional decomposition (1) is usually employed [24] to compute the contribution of each term to the overall variance, as shown in the next section.

2.1. Variance decomposition

ANOVA analysis is based on the variance decomposition in its conditional contributions. Variance can be written in terms of conditional expectation of f and f^2 as

$$\sigma^2 = \mathbb{E}(f^2) - \mathbb{E}(f)^2. \tag{7}$$

Starting from Eq. (1), it is easy to compute

$$f^2(\vec{\boldsymbol{\xi}}) = \sum_{i=0}^{N} f^2_{\vec{\boldsymbol{m}}_i}(\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_i) + 2 \sum_{i=0}^{N} \sum_{j=i+1}^{N} f_{\vec{\boldsymbol{m}}_i}(\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_i) f_{\vec{\boldsymbol{m}}_j}(\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_j). \tag{8}$$

If the equation (8) is integrated in the stochastic space and the orthogonality property (3) is applied, variance can be decomposed as

$$\sigma^2 = \sum_{i=1}^{N} \int_{\Xi^d} f^2_{\vec{\boldsymbol{m}}_i}(\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_i) \mathrm{d}\mathbb{P}_i = \sum_{i=1}^{N} \int_{\hat{\Xi}_i} f^2_{\vec{\boldsymbol{m}}_i}(\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_i) \mathrm{d}\mathbb{P}_i, \tag{9}$$

where the symbol $\hat{\Xi}_i$ is employed to indicate Ξ^{nnz_i} for brevity and $\mathbb{E}(f) = f_0$.

Hence, variance can be expressed as the summation of all the conditional contributions

$$\sigma^2 = \sum_{i=1}^{N} \sigma^2_{\vec{\boldsymbol{m}}_i}, \tag{10}$$

where

$$\sigma^2_{\vec{\boldsymbol{m}}_i} = \int_{\hat{\Xi}_i} f^2_{\vec{\boldsymbol{m}}_i}(\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_i) \mathrm{d}\mathbb{P}_i. \tag{11}$$

In the following, a similar decomposition is obtained for the skewness and kurtosis.

2.2. Skewness decomposition in conditional terms

In the case of skewness, the major drawbacks is the presence of an higher number of terms to compute with respect to the variance case, *i.e.* non-null mixed terms exist which, in the case of the variance, are zero due to orthogonality. The first step in order to obtain the skewness decomposition of the function $f(\vec{\boldsymbol{\xi}})$ is to raise the ANOVA series to the third power by employing the multinomial theorem and integrating over the space Ξ. Reducing the number of terms, exploiting the orthogonality of the basis, is possible and a more compact final expression can be obtained (the full derivation is reported in Appendix A)

$$\mu^3 = \sum_{p=1}^{N} \int_{\hat{\Xi}_p} f^3_{\vec{\boldsymbol{m}}_p} \mathrm{d}\mathbb{P}_p + 3 \sum_{\vec{\boldsymbol{m}}_p} \sum_{\vec{\boldsymbol{m}}_q \subset \vec{\boldsymbol{m}}_p} \int_{\hat{\Xi}_{pq}} f^2_{\vec{\boldsymbol{m}}_p} f_{\vec{\boldsymbol{m}}_q} \mathrm{d}\mathbb{P}_{pq} + 6 \sum_{p=1}^{N} \sum_{q=p+1}^{N} \sum_{\substack{r=q+1 \\ \vec{\boldsymbol{m}}_{pq} = \vec{\boldsymbol{m}}_r}}^{N} \int_{\hat{\Xi}_{pq}} f_{\vec{\boldsymbol{m}}_p} f_{\vec{\boldsymbol{m}}_q} f_{\vec{\boldsymbol{m}}_r} \mathrm{d}\mathbb{P}_{pqr}. \tag{12}$$

where $\hat{\Xi}_{ij} = \Xi^{\mathrm{nnz}_{ij}}$ and $\hat{\Xi}_{ijk} = \Xi^{\mathrm{nnz}_{ijk}}$. The notation is also simplified by omitting the explicit dependence of the function $f_{\vec{\boldsymbol{m}}_i}$ with respect to its coordinates, *i.e.* $f_{\vec{\boldsymbol{m}}_i} = f_{\vec{\boldsymbol{m}}_i}(\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_i)$.

Here, a special notation is introduced in order to compute multi-indices as $\vec{\boldsymbol{m}}_{ab\cdots z}$ as follows

$$\vec{\boldsymbol{m}}_{ab\cdots z} = \vec{\boldsymbol{m}}_a \boxplus \vec{\boldsymbol{m}}_b \boxplus \cdots \boxplus \vec{\boldsymbol{m}}_z = \left(\frac{m_{a_1} + m_{b_1} + \cdots m_{z_1}}{\max\left(1, \sum_{i=a}^{z} m_{i_1}\right)}, \ldots, \frac{m_{a_d} + m_{b_d} + \cdots + m_{z_d}}{\max\left(1, \sum_{i=a}^{z} m_{i_d}\right)} \right). \tag{13}$$

A useful outcome of this decomposition is the possibility to identify the conditional terms related to each single variable or group of variables as expressed for the variance by means of relation (11). In the case of skewness, the conditional terms have a more complex expression (except the first order terms, *i.e.* the terms related to the single variables). This complexity arises from the presence of mixed contributions. To obtain the additive form (as it holds for the variance)

$$\mu^3 = \sum_{i=1}^{N} \mu^3_{\vec{\boldsymbol{m}}_i}, \tag{14}$$

it is mandatory to identify all the set of indices whose interactions belong to an assigned multi-index $\vec{\boldsymbol{m}}_i$.

By considering that to each multi-index $\vec{\boldsymbol{m}}_i$ is associated a set of $2^{|\vec{\boldsymbol{m}}_i|} - 1$ sub-interactions and denoting this set as $\mathcal{P}_i$ (for instance if $\vec{\boldsymbol{m}}_i = (1,1)$ then the set $\mathcal{P}_i = \{(1,0),(0,1),(1,1)\}$), from the equation (12) it is possible to identify each contribution as follows

$$\mu^3_{\vec{\boldsymbol{m}}_i} = \int_{\hat{\Xi}_i} f^3_{\vec{\boldsymbol{m}}_i} \mathrm{d}\mathbb{P}_i + 3 \int_{\hat{\Xi}_i} f^2_{\vec{\boldsymbol{m}}_i} \sum_{\vec{\boldsymbol{m}}_q \in \mathcal{P}_{i,\neq}} f_{\vec{\boldsymbol{m}}_q} \mathrm{d}\mathbb{P}_i + 6 \sum_{\vec{\boldsymbol{m}}_p \in \mathcal{P}_{i,\neq}} \sum_{\substack{\vec{\boldsymbol{m}}_p \neq \vec{\boldsymbol{m}}_q \in \mathcal{P}_{i,\neq} \\ \vec{\boldsymbol{m}}_{pq} = \vec{\boldsymbol{m}}_i}} \int_{\hat{\Xi}_i} f_{\vec{\boldsymbol{m}}_i} f_{\vec{\boldsymbol{m}}_p} f_{\vec{\boldsymbol{m}}_q} \mathrm{d}\mathbb{P}_i, \tag{15}$$

where the following shorthand is used $\mathcal{P}_{i,\neq} = \mathcal{P}_i \setminus \{\vec{\boldsymbol{m}}_i\}$. It is easy to show that Eq. (15) reduces to a sum of zero contributions in the case of a function $f(\vec{\boldsymbol{\xi}})$ with Gaussian distribution; for instance assuming $f(\vec{\boldsymbol{\xi}}) = \sum_{i=1}^{3} \xi_i$ where $\xi_i \sim \mathcal{N}(0, 1/3)$, from the convolution of the pdfs we know $f(\vec{\boldsymbol{\xi}}) \sim \mathcal{N}(0,1)$, and $\mu^3 = 0$ with $\mu^3_{\vec{\boldsymbol{m}}_i} = 0 \; \forall \vec{\boldsymbol{m}}_i$. Note that Eq. (15) is also explicitly obtained in the Appendix A.

2.2.1. Kurtosis decomposition in conditional term

The decomposition of the kurtosis is presented in the following. The decomposition based on the functional form Eq. (1), after the application of the multinomial theorem and integration, leads to a final expression (for more details see the Appendix B) which reads

$$\begin{aligned} \mu^4 = & \sum_{p=1}^{N} \int_{\hat{\Xi}_p} f^4_{\vec{\boldsymbol{m}}_p} \mathrm{d}\mathbb{P}_p + 4 \sum_{\vec{\boldsymbol{m}}_p} \sum_{\vec{\boldsymbol{m}}_q \subset \vec{\boldsymbol{m}}_p} \int_{\hat{\Xi}_{pq}} f^3_{\vec{\boldsymbol{m}}_p} f_{\vec{\boldsymbol{m}}_q} \mathrm{d}\mathbb{P}_{pq} + 6 \sum_{p=1}^{N} \sum_{q=p+1}^{N} \int_{\hat{\Xi}_{pq}} f^2_{\vec{\boldsymbol{m}}_p} f^2_{\vec{\boldsymbol{m}}_q} \mathrm{d}\mathbb{P}_{pq} \\ & + 12 \sum_{p=1}^{N} \sum_{\substack{q=1 \\ q \neq p}}^{N} \sum_{\substack{r=q+1 \\ \vec{\boldsymbol{m}}_{qr} \setminus \cap_{qr} \subseteq \vec{\boldsymbol{m}}_p}}^{N} \int_{\hat{\Xi}_{pqr}} f^2_{\vec{\boldsymbol{m}}_p} f_{\vec{\boldsymbol{m}}_q} f_{\vec{\boldsymbol{m}}_r} \mathrm{d}\mathbb{P}_{pqr} + 24 \sum_{p=1}^{N} \sum_{q=p+1}^{N} \sum_{r=q+1}^{N} \sum_{\substack{t=r+1 \\ \vec{\boldsymbol{m}}_{pq} \setminus \cap_{pq} \subseteq \vec{\boldsymbol{m}}_{rt} \\ \vec{\boldsymbol{m}}_{rt} \subseteq \vec{\boldsymbol{m}}_{pq} \boxplus \cap_{rt}}}^{N} \int_{\hat{\Xi}_{pqrt}} f_{\vec{\boldsymbol{m}}_p} f_{\vec{\boldsymbol{m}}_q} f_{\vec{\boldsymbol{m}}_r} f_{\vec{\boldsymbol{m}}_t} \mathrm{d}\mathbb{P}_{pqrt}. \end{aligned} \tag{16}$$

Note that the operator of subtraction by set is employed with the standard notation $\setminus$, while $\cap_{rt}$ is a shorthand which indicates the intersection between the two multi-indices $\vec{\boldsymbol{m}}_{\boldsymbol{r}}$ and $\vec{\boldsymbol{m}}_{\boldsymbol{t}}$, *i.e.* it contains the variables which belong to both the multi-indices.

As for the skewness, the additive form for the kurtosis

$$\mu^4 = \sum_{i=1}^{N} \mu^4_{\vec{\boldsymbol{m}}_i}, \tag{17}$$

can be obtained when each specific multi-index is provided. Indeed, the conditional expression for $\mu^4_{\vec{\boldsymbol{m}}_i}$ is equal to (see the Appendix B for more details)

$$\begin{aligned}
\mu^4_{\vec{\boldsymbol{m}}_i} = & \int_{\hat{\Xi}_i} f^4_{\vec{\boldsymbol{m}}_i} \mathrm{d}\mathbb{P}_i + 4 \int_{\hat{\Xi}_i} f^3_{\vec{\boldsymbol{m}}_i} \sum_{\vec{\boldsymbol{m}}_q \in \mathcal{P}_{i,\neq}} f_{\vec{\boldsymbol{m}}_q} \mathrm{d}\mathbb{P}_i + 6 \sum_{\vec{\boldsymbol{m}}_p \in \mathcal{P}_i} \sum_{\substack{\vec{\boldsymbol{m}}_p \neq \vec{\boldsymbol{m}}_q \in \mathcal{P}_i \\ \vec{\boldsymbol{m}}_{pq} = \vec{\boldsymbol{m}}_i}} \int_{\hat{\Xi}_i} f^2_{\vec{\boldsymbol{m}}_p} f^2_{\vec{\boldsymbol{m}}_q} \mathrm{d}\mathbb{P}_i \\
& + 12 \sum_{\vec{\boldsymbol{m}}_p} \sum_{\vec{\boldsymbol{m}}_p \neq \vec{\boldsymbol{m}}_q \in \mathcal{P}_i} \sum_{\substack{\vec{\boldsymbol{m}}_r \in \mathcal{P}_i, r>q \\ \vec{\boldsymbol{m}}_p \boxplus \cap_{qr} = \vec{\boldsymbol{m}}_i}} \int_{\hat{\Xi}_i} f^2_{\vec{\boldsymbol{m}}_p} f_{\vec{\boldsymbol{m}}_q} f_{\vec{\boldsymbol{m}}_r} \mathrm{d}\mathbb{P}_i \\
& + 24 \sum_{\vec{\boldsymbol{m}}_p \in \mathcal{P}_i} \sum_{\vec{\boldsymbol{m}}_q \in \mathcal{P}_i, q>p} \sum_{\vec{\boldsymbol{m}}_r \in \mathcal{P}_i, r>q} \sum_{\substack{t>r, \vec{\boldsymbol{m}}_r \in \mathcal{P}_i \\ \vec{\boldsymbol{m}}_i \subseteq \vec{\boldsymbol{m}}_{pq} \boxplus \cap_{rt} \\ \vec{\boldsymbol{m}}_i \subseteq \vec{\boldsymbol{m}}_{rt} \boxplus \cap_{pq}}} \int_{\hat{\Xi}_i} f_{\vec{\boldsymbol{m}}_p} f_{\vec{\boldsymbol{m}}_q} f_{\vec{\boldsymbol{m}}_r} f_{\vec{\boldsymbol{m}}_t} \mathrm{d}\mathbb{P}_i.
\end{aligned} \tag{18}$$

3. Correlation with Polynomial Chaos Framework

This section is devoted to illustrate how variance, skewness and kurtosis functional decompositions can be computed within the polynomial chaos framework. In this context, an approximation $\tilde{f}$ of the function f is provided

$$f(\vec{\boldsymbol{\xi}}) \approx \tilde{f}(\vec{\boldsymbol{\xi}}) = \sum_{k=0}^{P} \beta_k \Psi_k(\vec{\boldsymbol{\xi}}), \tag{19}$$

where P is computed according to the order of the polynomial expansion n_0 and the stochastic dimension of the problem d

$$P + 1 = \frac{(n_0 + d)!}{n_0! d!}. \tag{20}$$

Each polynomial $\Psi_k(\vec{\boldsymbol{\xi}})$ of total degree n_o is a multivariate polynomial form which involves tensorization of 1D polynomials by using a multi-index $\vec{\boldsymbol{\alpha}}^k \in \mathbb{N}^d$, with $\sum_{i=1}^{d} \vec{\boldsymbol{\alpha}}^k_i \leq n_0$:

$$\Psi_k(\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}^{\star,k}) = \prod_{i=1}^{d} \psi_{\alpha^k_i}(\xi_i), \tag{21}$$

where the multi-index $\vec{\boldsymbol{m}}^{\star,k} = \vec{\boldsymbol{m}}^{\star,k}(\vec{\boldsymbol{\alpha}}^k) \in \mathbb{N}^d$ is a function of $\vec{\boldsymbol{\alpha}}^k$: $\vec{\boldsymbol{m}}^{\star,k} = (m^{\star,k}_1, \ldots, m^{\star,k}_d)$, with $m^{\star,k}_i = \alpha^k_i / \max\left(1, \alpha^k_i\right)$. It is important to note here that the number of multi-indices $\vec{\boldsymbol{\alpha}}^k$ is $P + 1$, see Eq. (20), which is function of the polynomial degree n_0. It follows that two multi-indices $\vec{\boldsymbol{m}}^{\star,i}(\vec{\boldsymbol{\alpha}}^i)$ and $\vec{\boldsymbol{m}}^{\star,j}(\vec{\boldsymbol{\alpha}}^j)$ can be equal even if $\vec{\boldsymbol{\alpha}}^i \neq \vec{\boldsymbol{\alpha}}^j$, *i.e.* they contain the same variables but the corresponding 1D polynomials are raised to a different power. By contrast, the ANOVA multi-indices are 2^d and the following condition always holds: $\vec{\boldsymbol{m}}_{\boldsymbol{i}} \neq \vec{\boldsymbol{m}}_{\boldsymbol{j}}$ for $i \neq j$.

For each polynomial basis, $\psi_0(\xi_i) = 1$ and then $\Psi_0(\vec{\boldsymbol{\xi}}) = 1$. Hence, the first coefficient β_0 is equal to the expected value of the function, *i.e.* $\mathbb{E}(f)$. 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 $p(\vec{\xi})$ of the inputs. Thanks to the orthogonality, the following relation holds

$$\int_{\Xi^d} \Psi_i(\vec{\xi})\Psi_k(\vec{\xi})\mathrm{d}\mathbb{P} = \delta_{ik}\langle\Psi_i(\vec{\xi}), \Psi_i(\vec{\xi})\rangle, \tag{22}$$

where $\langle\cdot,\cdot\rangle$ indicates the inner product and δ_{ij} is the Kronecker delta function.

The orthogonality can be used to compute the coefficients of the expansion in a non-intrusive PC framework as follows

$$\beta_k = \frac{\langle f(\vec{\xi}), \Psi_k(\vec{\xi})\rangle}{\langle\Psi_k(\vec{\xi})\Psi_k(\vec{\xi})\rangle} \quad \forall k, \tag{23}$$

where the integrals can be computed by means of any quadrature formula as described in [28].

3.1. Variance decomposition

The second order central moment $\mathbb{E}(f^2)$ can be evaluated as

$$\mathbb{E}(f^2) = \int_{\Xi^d} f(\vec{\xi})^2\mathrm{d}\mathbb{P} = \int_{\Xi^d}\left(\sum_{k=0}^{P}\beta_k\Psi_k(\vec{\xi})\right)^2\mathrm{d}\mathbb{P}. \tag{24}$$

This term can be computed exploiting the orthogonality

$$\mathbb{E}(f^2) = \int_{\Xi^d}\left(\sum_{k=0}^{P}\beta_k\Psi_k(\vec{\xi})\right)^2\mathrm{d}\mathbb{P} = \sum_{k=0}^{P}\beta_k^2\langle\Psi_k^2(\vec{\xi})\rangle. \tag{25}$$

As a consequence, variance is

$$\sigma^2 = \mathbb{E}(f^2) - \mathbb{E}(f)^2 = \sum_{k=1}^{P}\beta_k^2\langle\Psi_k^2(\vec{\xi})\rangle. \tag{26}$$

Finally, an explicit connection between the last expression and the Eq. (9) is found. Thus, each conditional term can be computed as

$$\sigma^2_{\vec{m}_i} = \sum_{k\in K_{\vec{m}_i}}\beta_k^2\langle\Psi_k^2(\vec{\xi})\rangle, \tag{27}$$

where $K_{\vec{m}_i}$ represents the set of indices associated to the variables included in the vector $(\vec{\xi}\cdot\vec{m}_i)$:

$$K_{\vec{m}_i} = \left\{k\in\{1,\dots,P\}\,|\,\vec{m}^{\star,k} = \vec{m}^{\star,k}(\vec{\alpha}^k) = \vec{m}_i\right\} \tag{28}$$

3.2. Skewness decomposition

The PC expansion, Eq. (19), raised to the third power by applying the multinomial theorem, and integrating over the whole stochastic space Ξ leads to a form which includes null terms thanks to the orthogonality properties of the PC basis. A compact form (the full derivation is reported in Appendix C) can be written as

$$\mu^3 = \sum_{p=1}^{P}\beta_p^3\langle\Psi_p^3(\vec{\xi})\rangle + 3\sum_{p=1}^{P}\beta_p^2\sum_{\substack{q=1\\q\neq p}}^{P}\beta_q\langle\Psi_p^2(\vec{\xi}), \Psi_q(\vec{\xi})\rangle\Delta_q^p + 6\sum_{p=1}^{P}\sum_{q=p+1}^{P}\sum_{r=q+1}^{P}\beta_p\beta_q\beta_r\langle\Psi_p(\vec{\xi}), \Psi_q(\vec{\xi})\Psi_r(\vec{\xi})\rangle\Delta_{pqr}, \tag{29}$$

where two functions are introduced for the selection. The first one Δ_q^p is defined as follows

$$\Delta_q^p = \begin{cases} 0 & \text{if} \quad \alpha_j^p = 0 \quad \text{and} \quad m_j^{\star,q} = 1 \\ 1 & \text{otherwise} \end{cases} \tag{30}$$

while the function Δ_{pqr} is defined as

$$\Delta_{pqr} = \begin{cases} 0 & \text{if} \quad m_j^{\star,p} + m_j^{\star,q} + m_j^{\star,r} = 1, 2 \\ 1 & \text{otherwise.} \end{cases} \tag{31}$$

For a specific $\vec{\boldsymbol{m}}_{\boldsymbol{i}}$, the previous expression reduces to

$$s_{\vec{\boldsymbol{m}}_{\boldsymbol{i}}} = \sum_{p \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{i}}}} \beta_p^3 \langle \Psi_p^3(\vec{\boldsymbol{\xi}}) \rangle + 3 \sum_{p \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{p}}}} \beta_p^2 \sum_{\substack{q \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{q}}} \\ \vec{\boldsymbol{m}}_{\boldsymbol{pq}} = \vec{\boldsymbol{m}}_{\boldsymbol{i}}}} \beta_q \langle \Psi_p^2(\vec{\boldsymbol{\xi}}), \Psi_q(\vec{\boldsymbol{\xi}}) \rangle \Delta_q^p + 6 \sum_{p \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{p}}}} \sum_{\substack{q \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{q}}} \\ q \geq p+1}} \sum_{\substack{r \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{r}}} \\ \vec{\boldsymbol{m}}_{\boldsymbol{pqr}} = \vec{\boldsymbol{m}}_{\boldsymbol{i}}}} \beta_p \beta_q \beta_r \langle \Psi_p(\vec{\boldsymbol{\xi}}), \Psi_q(\vec{\boldsymbol{\xi}}) \Psi_r(\vec{\boldsymbol{\xi}}) \rangle \Delta_{pqr}. \tag{32}$$

Note that the two functions Δ_q^p and Δ_{pqr} should be computed before computing the integral associated to each term for an efficient implementation. Only if this value is one, then the integral need to be truly computed.

3.3. Kurtosis decomposition

After the application of the multinomial theorem on Eq. (19) and its integration over Ξ, the expression for the kurtosis is obtained, in compact form as (the full form is derived in Appendix D)

$$\begin{aligned} \mu^4 &= \int_{\Xi^d} \left(f(\vec{\boldsymbol{\xi}}) - \beta_0 \right)^4 \mathrm{d}\mathbb{P} = \int_{\Xi^d} \left(\sum_{p=1}^{P} \beta_p \Psi_p(\vec{\boldsymbol{\xi}}) \right)^4 \mathrm{d}\mathbb{P} \\ &= \sum_{p=1}^{P} \beta_p^4 \langle \Psi_p^4(\vec{\boldsymbol{\xi}}) \rangle + 4 \sum_{p=1}^{P} \beta_p^3 \sum_{\substack{q=1 \\ q \neq p}}^{P} \beta_q \langle \Psi_p^3, \Psi_q \rangle \Delta_q^p + 6 \sum_{p=1}^{P} \beta_p^2 \sum_{q=p+1}^{P} \beta_q^2 \langle \Psi_p^2, \Psi_q^2 \rangle \\ &+ 12 \sum_{p=1}^{P} \beta_p^2 \sum_{\substack{q=1 \\ q \neq p}}^{P} \beta_q \sum_{\substack{r=q+1 \\ r \neq p}}^{P} \beta_r \langle \Psi_p^2, \Psi_q \Psi_r \rangle \Delta_{qr}^p + 24 \sum_{p=1}^{P} \sum_{q=p+1}^{P} \sum_{r=q+1}^{P} \sum_{t=r+1}^{P} \beta_p \beta_q \beta_r \beta_t \langle \Psi_p \Psi_q, \Psi_r \Psi_t \rangle \Delta_{pqrt}, \end{aligned} \tag{33}$$

where the function Δ_q^p is already introduced in (30), while the others two functions are defined as follows

$$\Delta_{qr}^p = \begin{cases} 0 & \text{if} \quad \alpha_j^p = 0 \quad \text{and} \quad m_j^{\star,q} + m_j^{\star,r} = 1, 2 \\ 1 & \text{otherwise} \end{cases} \quad \text{and} \quad \Delta_{pqrt} = \begin{cases} 0 & \text{if} \quad m_j^{\star,p} + m_j^{\star,q} + m_j^{\star,r} + m_j^{\star,t} = 1, 2 \\ 1 & \text{otherwise.} \end{cases} \tag{34}$$

The functions $\Delta_p^q, \Delta_{pqr}, \Delta_{qr}^p$ and Δ_{pqrt} can be computed efficiently before evaluating the integral terms of the expansion. A detail of a possible algorithm for their evaluation is reported in Appendix E.

In the case of the conditional contribution associated to a specific multi-index $\vec{\boldsymbol{m}}_{\boldsymbol{i}}$, it holds

$$\begin{aligned} \mu^4_{\vec{\boldsymbol{m}}_{\boldsymbol{i}}} &= \sum_{k \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{i}}}} \beta_k^4 \langle \Psi_k^4(\vec{\boldsymbol{\xi}}) \rangle + 4 \sum_{p \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{p}}}} \beta_p^3 \sum_{\substack{q \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{q}}} - \{p\} \\ \vec{\boldsymbol{m}}_{\boldsymbol{p}} \boxplus \vec{\boldsymbol{m}}_{\boldsymbol{q}} = \vec{\boldsymbol{m}}_{\boldsymbol{i}}}} \beta_q \langle \Psi_p^3, \Psi_q \rangle \Delta_q^p + 6 \sum_{p \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{p}}}} \beta_p^2 \sum_{\substack{q \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{q}}} - \{p\} \\ \vec{\boldsymbol{m}}_{\boldsymbol{p}} \boxplus \vec{\boldsymbol{m}}_{\boldsymbol{q}} = \vec{\boldsymbol{m}}_{\boldsymbol{i}}}} \beta_q^2 \langle \Psi_p^2, \Psi_q^2 \rangle \\ &+ 12 \sum_{p \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{p}}}} \beta_p^2 \sum_{q \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{q}}} - \{p\}} \beta_q \sum_{\substack{r \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{r}}} \\ r \geq q+1 \\ \vec{\boldsymbol{m}}_{\boldsymbol{pqr}} = \vec{\boldsymbol{m}}_{\boldsymbol{i}}}} \beta_r \langle \Psi_p^2, \Psi_q \Psi_r \rangle \Delta_{qr}^p + 24 \sum_{p \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{p}}}} \sum_{\substack{q \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{q}}} \\ q \geq p+1}} \sum_{\substack{r \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{r}}} \\ r \geq q+1}} \sum_{\substack{t \in K_{\vec{\boldsymbol{m}}_{\boldsymbol{t}}} \\ t \geq r+1 \\ \vec{\boldsymbol{m}}_{\boldsymbol{pqrt}} = \vec{\boldsymbol{m}}_{\boldsymbol{i}}}} \beta_p \beta_q \beta_r \beta_t \langle \Psi_p \Psi_q, \Psi_r \Psi_t \rangle \Delta_{pqrt}. \end{aligned} \tag{35}$$

4. Sensitivity indices and truncation error estimation

As introduced by Sobol [24], sensitivity indices for variance can be computed for each conditional contribution following Eq. (10):

$$\sigma^{2,\mathrm{SI}}_{\vec{\boldsymbol{m}}_{\boldsymbol{i}}} = \frac{\sigma^2_{\vec{\boldsymbol{m}}_{\boldsymbol{i}}}}{\sigma^2}. \tag{36}$$

Here, we introduce additional sensitivity indices based on the decomposition of skewness and kurtosis[1]. From the definition of the conditional terms in (15) and (18) it follows

$$\begin{aligned} s^{\mathrm{SI}}_{\vec{\boldsymbol{m}}_i} &= \frac{s_{\vec{\boldsymbol{m}}_i}}{s} \\ k^{\mathrm{SI}}_{\vec{\boldsymbol{m}}_i} &= \frac{k_{\vec{\boldsymbol{m}}_i}}{k}. \end{aligned} \tag{37}$$

In the case of the total sensitivity index (TSI) it is necessary to compute the overall influence of a variable. This sensitivity index can be computed summing up all the contributions in which the chosen variable is present

$$\begin{aligned} \mathrm{TSI}_j &= \sum_{\xi_j \in (\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_i)} \sigma^{2,\mathrm{SI}}_{\vec{\boldsymbol{m}}_i} \\ \mathrm{TSI}^s_j &= \sum_{\xi_j \in (\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_i)} s^{\mathrm{SI}}_{\vec{\boldsymbol{m}}_i} \\ \mathrm{TSI}^k_j &= \sum_{\xi_j \in (\vec{\boldsymbol{\xi}} \cdot \vec{\boldsymbol{m}}_i)} k^{\mathrm{SI}}_{\vec{\boldsymbol{m}}_i}. \end{aligned} \tag{38}$$

Moreover, following [41], it is possible to define two distinct effective dimensions for the ANOVA representation (1), namely the effective dimension δ_{tr} in the truncation sense

$$f(\vec{\boldsymbol{\xi}}) = \sum_{\vec{\boldsymbol{m}}_i} f_{\vec{\boldsymbol{m}}_i} \simeq \sum_{0<i\leq\delta_{\mathrm{tr}}} f_{\vec{\boldsymbol{m}}_i} = f_{\mathrm{tr}}, \tag{39}$$

and the effective dimension δ_{sup} in the superposition sense

$$f(\vec{\boldsymbol{\xi}}) = \sum_{\vec{\boldsymbol{m}}_i} f_{\vec{\boldsymbol{m}}_i} \simeq \sum_{\mathrm{nnz}_i\leq\delta_{\mathrm{sup}}} f_{\vec{\boldsymbol{m}}_i} = f_{\mathrm{sup}}. \tag{40}$$

The two integers δ_t and δ_s represent the number of terms which are needed in order to reach the required amount of variance, *i.e.* $\sigma^2(f_{\mathrm{tr}}) \leq C_{\mathrm{tr}}\sigma^2(f)$ and $\sigma^2(f_{\mathrm{sup}}) \leq C_{\mathrm{sup}}\sigma^2(f)$ respectively. Generally, constants C_{tr} and C_{sup} are very close to the unity; in [41] both are chosen to be 0.99. It is important to note here that the dimension δ_{tr} is dependent on the ordering of the terms employed in (1). From a practical point-of-view it is common to approximate the function of interest with a model including only first order interaction terms, *i.e.* terms which depend only on one single variable[2]. Hereinafter, the generic term metamodel is used to denote an approximation of the true function; it can assume any functional form when not differently specified.. This case is equivalent to choose $\delta_{\mathrm{sup}} = 1$, thus $\sum_{\mathrm{nnz}_i\leq 1} f_{\vec{\boldsymbol{m}}_i} = f_I$. When the superposition dimension is unitary, *i.e.* only the terms depending on each single variable are retained, it is possible to deduce explicitly the central moments corresponding to this metamodel

$$\int_{\Xi^d} (f(\vec{\boldsymbol{\xi}}) - f_0)^n \mathrm{d}\mathbb{P} \approx \int_{\Xi^d} (f_I(\vec{\boldsymbol{\xi}}) - f_0)^n \mathrm{d}\mathbb{P} = \int_{\Xi} \left(\sum_{\mathrm{nnz}_i=1} f_{\vec{\boldsymbol{m}}_i} \right)^n \mathrm{d}\mathbb{P}. \tag{41}$$

For $n = 2$ (variance) from the orthogonality of the ANOVA terms it follows

$$\int_{\Xi^d} (f_I(\vec{\boldsymbol{\xi}}) - f_0)^2 \mathrm{d}\mathbb{P} = \sum_{\mathrm{nnz}_i=1} \int_{\hat{\Xi}_i} f^2_{\vec{\boldsymbol{m}}_i} \mathrm{d}\mathbb{P}_i = \sigma^2_I = C_{\sigma^2}\sigma^2, \quad \text{with} \quad C_{\sigma^2} \leq 1. \tag{42}$$

The case of the third order is obtained with $n = 3$

$$\int_{\Xi^d} (f_I(\vec{\boldsymbol{\xi}}) - f_0)^3 \mathrm{d}\mathbb{P} = \sum_{\mathrm{nnz}_i=1} \int_{\hat{\Xi}_i} f^3_{\vec{\boldsymbol{m}}_i} \mathrm{d}\mathbb{P}_i = \mu^3_I = C_s\mu^3. \tag{43}$$

[1] In this case the ratios between the conditional central moment or the normalized one with the overall moment are equivalent.

[2] We employ for it the Roman number *I* to avoid confusion with the first term of the ANOVA expansion.

In this case the constant C_s can be either negative or zero. It is important to note that the terms of the kind $\int_{\hat{\Xi}_{ij}} f^2_{\vec{m}_i} f_{\vec{m}_j} \mathrm{d}\mathbb{P}_{ij}$, with $i \neq j$, are null for orthogonality. The skewness associated to the metamodel including only first order interaction terms has the same functional form of the variance, *i.e.* it is additive with respect to the contributes relates to each single variable.

The kurtosis results in a more complex form. By raising to the fourth power, $n = 4$, it follows

$$\int_{\Xi^d} (f_I(\vec{\xi}) - f_0)^4 \mathrm{d}\mathbb{P} = \sum_{\mathrm{nnz}_i=1} \int_{\hat{\Xi}_i} f^4_{\vec{m}_i} \mathrm{d}\mathbb{P}_i + 6 \sum_{\mathrm{nnz}_i=1} \sum_{\substack{\mathrm{nnz}_j=1 \\ j>i}} \int_{\hat{\Xi}_{ij}} f^2_{\vec{m}_i} f^2_{\vec{m}_j} \mathrm{d}\mathbb{P}_{ij} = \mu^4_I = C_k \mu^4, \quad \text{with} \quad C_k \geq 0. \tag{44}$$

The fourth central moment μ^4_I, in this case, includes the contributions associated to the second order interactions. Thus any of the terms $\int f^2_{\vec{m}_i} f^2_{\vec{m}_j} \mathrm{d}\mathbb{P}_{ij}$ belongs to $k_{\vec{m}_{ij}}$, where $\vec{m}_{ij} = \vec{m}_i \boxplus \vec{m}_j$. The kurtosis cannot be obtained only summing the first order contributions. Hence, this metamodel exhibits an higher kurtosis than the one corresponding only to the sum of its first order interaction terms.

If the usual definitions of skewness or kurtosis are needed, it is necessary to normalize the corresponding central moments by a factor, namely the third and fourth powers of the standard deviation. Hence, the error on the second central moment, *i.e.* the variance, needs to be taken into account as follows

$$\begin{aligned} s_I &= \frac{\mu^3_I}{(\sigma^2_I)^{3/2}} = \frac{C_s}{(C_{\sigma^2})^{3/2}} \frac{\mu^3}{(\sigma^2)^{3/2}} \\ k_I &= \frac{\mu^4_I}{(\sigma^2_I)^2} = \frac{C_k}{(C_{\sigma^2})^2} \frac{\mu^4}{(\sigma^2)^2} \end{aligned} \tag{45}$$

The estimations obtained in this section allow to examine the quality of a metamodel including only first order interaction terms by using additional information with respect the approach based only on the variance. In Section 5 these estimations are shown to be very useful to avoid situation in which a metamodel is accepted only based on the variance, while in reality it produces poor results in terms of high-order statistics. For instance, it is very common to obtain pdf which cannot preserve the asymmetry (skewness) properties of the original probability density function. Moreover, the previous estimations (45) are only based on the computation of the first order conditional terms (see (15) and (35) with $\mathrm{nnz}_i = 1$), thus their evaluation results to be always affordable numerically.

Moreover, the previous estimations can be adopted to select the best possible metamodel, in term of polynomial degree given the simulations at disposal. A classical non-intrusive implementation of a PC method starts from the knowledge of $n = 1, \ldots, N$ functional evaluations of the model. For each of these realization a weight w_n is associated according to the quadrature rule desired (it can be either a full tensorization of Gauss-based 1D quadrature rules, Sparse Grids or other more sophisticated choices, see also [28]). The pseudocode which follows is an example of a possible use of the estimations reported in Eq. 45 for the selection of the optimal polynomial degree of a metamodel including only first order interaction terms:

$n_0 = 1$, iter $= 1$ and $\varepsilon =$ `user_defined_tolerance`
while $\left(max(\Delta C_{\sigma^2}, \Delta C_s/(C_{\sigma^2})^{3/2}, \Delta C_k/(C_{\sigma^2})^2) > \varepsilon \quad \& \quad n_0 > 1\right)$ **do**
  for $i = 1, \ldots, d$ **do**
    for $k \in K_{\vec{m}_i}$ **do**
      $\beta_k = \frac{\langle f(\vec{\xi}), \Psi_k(\vec{\xi})\rangle}{\langle \Psi_k(\vec{\xi}), \Psi_k(\vec{\xi})\rangle}$
    end
  end
  for $n = 1, \ldots, N$ **do**
    $f^{(n)}_{\vec{m}_i} = \sum\limits_{k \in K_{\vec{m}_i}} \beta_k \Psi^{(n)}_k$
  end
  for $i = 1, \ldots, d$ **do**

$$\sigma^2_{I,i} = \sum_{n=1}^{N} w_n (f^{(n)}_{\vec{m}_i})^2, \quad \mu^3_{I,i} = \sum_{n=1}^{N} w_n (f^{(n)}_{\vec{m}_i})^3, \quad \text{and} \quad \mu^4_{I,i} = \sum_{n=1}^{N} w_n (f^{(n)}_{\vec{m}_i})^4 \tag{46}$$

    for $j = i + 1, \ldots, d$ **do**
      $\mu^4_{I,i} = \mu^4_{I,i} + 6 \sum\limits_{n=1}^{N} w_n (f^{(n)}_{\vec{m}_i})^2 (f^{(n)}_{\vec{m}_j})^2$
    end
  end
  Assembling the central moments:

$$\sigma^2_I = \sum_{i=1}^{d} \sigma^2_{I,i}, \quad \mu^3_I = \sum_{i=1}^{d} \mu^3_{I,i} \quad \text{and} \quad \mu^4_I = \sum_{i=1}^{d} \mu^4_{I,i} \tag{47}$$

  Evaluation of the central moments by collocation:

$$\begin{aligned} &\mu = \sum_{n=1}^{N} w_n f^{(n)}, \quad \sigma^2 = \sum_{n=1}^{N} w_n (f^{(n)})^2 - (\mu)^2, \\ &\mu^3 = \sum_{n=1}^{N} w_n (f^{(n)})^3 - 3\mu\sigma^2 - (\mu)^3, \quad \text{and} \quad \mu^4 = \sum_{n=1}^{N} w_n (f^{(n)})^4 - 4\mu^3\mu - 6(\mu)^2\sigma^2 - (\mu)^4. \end{aligned} \tag{48}$$

  Evaluation of the constants for the first order interaction metamodel:

$$C_{\sigma^2} = \frac{\sigma^2_I}{\sigma^2}, \quad C_s = \frac{\mu^3_I}{\mu^3} \quad \text{and} \quad C_k = \frac{\mu^4_I}{\mu^4} \tag{49}$$

  if iter > 1 **then**

$$\begin{aligned} \Delta C_{\sigma^2} &= \left| C_{\sigma^2}|_{\text{iter}} - C_{\sigma^2}|_{\text{iter}-1} \right| \\ \Delta C_s &= \left| \left.\frac{C_s}{(C_{\sigma^2})^{3/2}}\right|_{\text{iter}} - \left.\frac{C_s}{(C_{\sigma^2})^{3/2}}\right|_{\text{iter}-1} \right| \\ \Delta C_k &= \left| \left.\frac{C_k}{(C_{\sigma^2})^{2}}\right|_{\text{iter}} - \left.\frac{C_k}{(C_{\sigma^2})^{2}}\right|_{\text{iter}-1} \right| \end{aligned} \tag{50}$$

  end
  $n_0 = n_0 + 1$ and iter $=$ iter $+ 1$
end

Algorithm 1: Pseudocode for the choice of the optimal order for the first order interactions metamodel.

This algorithm requires only $n_0 d$ coefficients out of the $P+1 = (n_0+d)!/(n_0!d!)$ needed for a polynomial expansion of total order n_0. Indeed, the set $K_{\vec{m}_i}$ for a multi-index $\vec{m}_i$ with $\mathrm{nnz}_i = 1$ reduces to a set of n_0 elements, *i.e.* one for each polynomial order up to the total order n_0. Therefore, without requiring additional model evaluations the total degree n_0 can be raised up to reaching the convergence of the constants C_{σ^2}, C_s and C_k. The obtained polynomial degree represents the optimal degree for a metamodel including only first order interactions. In such a case, the metamodel obtained is the one including the minimum number of coefficients, however it is important to note that in complex models the metamodel including only first order contributions can furnish misleading interpretations of the tail of the distributions. In a case like that, the knowledge of the high-order moments decomposition is sufficient to evaluate, *a priori*, the quality of the metamodel. Relevant examples for this situation are reported in Section 5.5.

5. Numerical results

In this section, the importance of considering metrics based on high-order statistics for global sensitivity analysis is demonstrated through some numerical examples and a complex fluid flow application. The numerical test section is organized as follows. In Section 5.1, the previous relations for the computation of the high-order conditional terms are demonstrated numerically by showing the convergence properties of PC with respect to the analytical conditional high-order statistics. In Section 5.2, a comparison is performed between the information obtained by an analysis based on the variance or on high-order conditional contributions for numerical test problems with different kind of interactions between parameters. How reducing the model dimension in the truncation and in the superposition sense is described in Sections 5.3 and 5.4, respectively. Finally, in Section 5.5, the high-order decomposition analysis of a complex flow in a Organic Rankine Cycles (ORCs) turbine is presented, thus highlighting the importance of this kind of study during the design process of a complex system.

5.1. Computing conditional statistics by means of PC

The high-order statistics (the values of variance, skewness or kurtosis) can be computed using the same set of deterministic evaluation of the models, *i.e.* the same number of functional evaluations $f = f(\vec{\xi})$ in the same sample points, by evaluating the first order coefficients of the PC expansion for f, f^2, f^3 and f^4 corresponding to the expected values of the four functions. Hereinafter, this approach is referred to as *collocation*. It is important to remark that the collocation approach does not provide any kind of metamodel for the function f, nor the possibility to compute conditional terms. Then, this approach is employed only for a comparison with respect to the PC series and for assessing the convergence of the expansion.

Let consider the following function

$$f(\vec{\xi}) = \prod_{i=1}^{d} \sin(\pi\xi_i) \tag{51}$$

where each variable $\xi_i \sim \mathcal{U}(0,1)$ with an increasing dimension d up to three. In the following, statistical moments as well as sensitivity indices (relative) errors are systematically computed with respect to the analytical solution.

In Figure 1, convergence of the statistical moments is reported as a function of the number of functional evaluations for the dimension $d = 2$. A number of simulations equal to $N = 120$ is needed to reach a relative error of order $O(10^{-4})$ for the kurtosis. Also the collocation approach, as expected, converges faster, but, as already discussed, is limited to the computation of the full central moments.

Now, conditional statistics can be computed using a PC approach using Eqs. (29) and (33). In Figure 2, we show first-order interaction terms σ_1^2, μ_1^3 and μ_1^4 (where for symmetry $\sigma_1^2 = \sigma_2^2$, $\mu_1^3 = \mu_2^3$ and $\mu_1^4 = \mu_2^4$) and second order interaction terms (σ_{12}^2, μ_{12}^3, μ_{12}^4) errors computed with respect to the analytical solution. Statistics are well converged at $N = 120$. Then, the case with $d = 3$ is considered. In Figures 3 and 4, convergences of statistical moments and conditional statistics are reported, respectively. Convergence, for both statistical moments and conditional statistics with respect to their exact analytical counterparts, is attained at nearly $N = 1500$.

5.2. High-order indices analysis for global Sensitivity Analysis

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

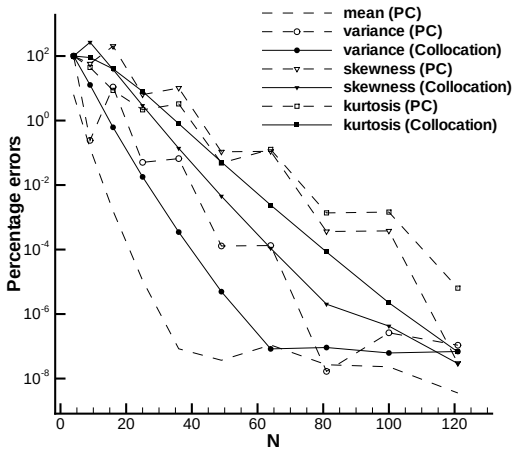


Figure 1: Statistical moments error vs number of function evaluations in the case $d = 2$.

Let us consider, the classical Sobol function (four dimensions)

$$f(\vec{\xi}) = \prod_{i=1}^{4} \frac{|4\xi_i - 2| + a_i}{1 + a_i}, \tag{52}$$

where $\xi_i \sim \mathcal{U}(0, 1)$. Two possible choices of the coefficients a_i are considered

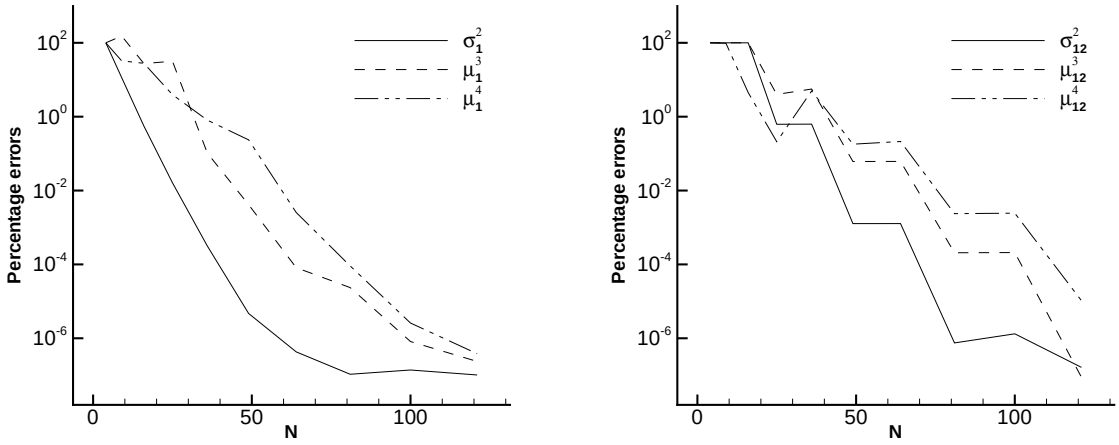
- $a_i = (i - 1)/2$, namely linear g-function f_{glin};
- $a_i = i^2$, namely quadratic g-function f_{gquad}.

In Figure 5, Sensitivity Indices (SI) for the linear g-function f_{glin} are reported. Several differences can be observed between the sensitivity indices computed on the variance or on high-order moments. The variance-based ranking illustrates that the first-order sensitivity indices, *i.e.* the indices relative to the first order interaction terms are higher than the second order ones, while these last ones are higher than the third and fourth order ones. This is not the case for skewness and kurtosis, where the second-order are the larger contributions. This behavior reveals that the variance is able to capture the ranking of the variables, but not the relative importance associated to higher-order interactions between variables. From a practical point of view, this error can lead to wrong decisions in a dimension reduction strategy as it will be shown in Sections 5.3 and 5.4. Quantitatively, the variance based only on first-order contributions exceeds 0.8, while skewness and kurtosis do not attain 0.1. In Table 1, the total sensitivity indices for the four variables are reported. It is evident that the ranking of variables is not influenced by the statistical moment chosen as metrics, but their relative importance is considerably different.

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

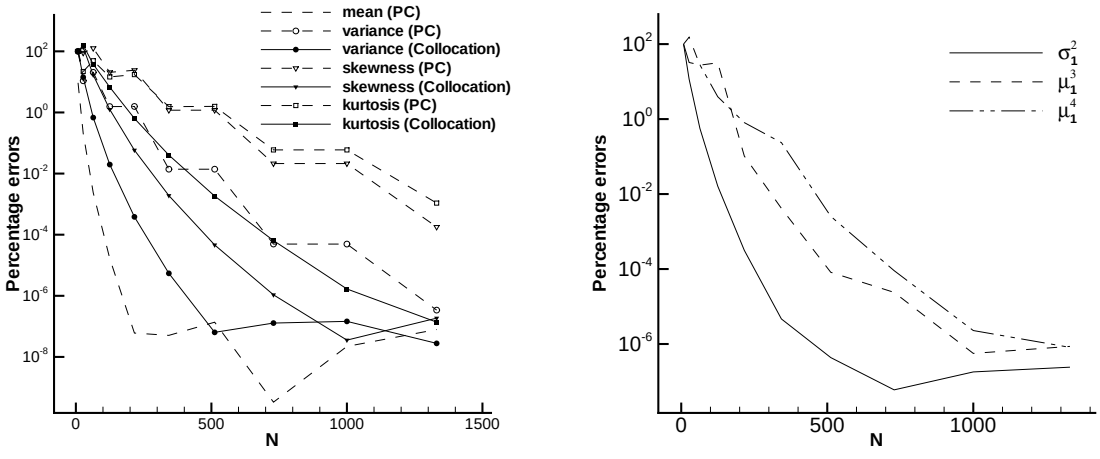
Table 1: Total sensitivity indices for the linear g-function function (52) based on a PC series with total degree $n_0 = 5$.

The same functional form, Eq. (52), can lead to slightly different results if the quadratic function coefficients are considered. In Figure 6, 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.



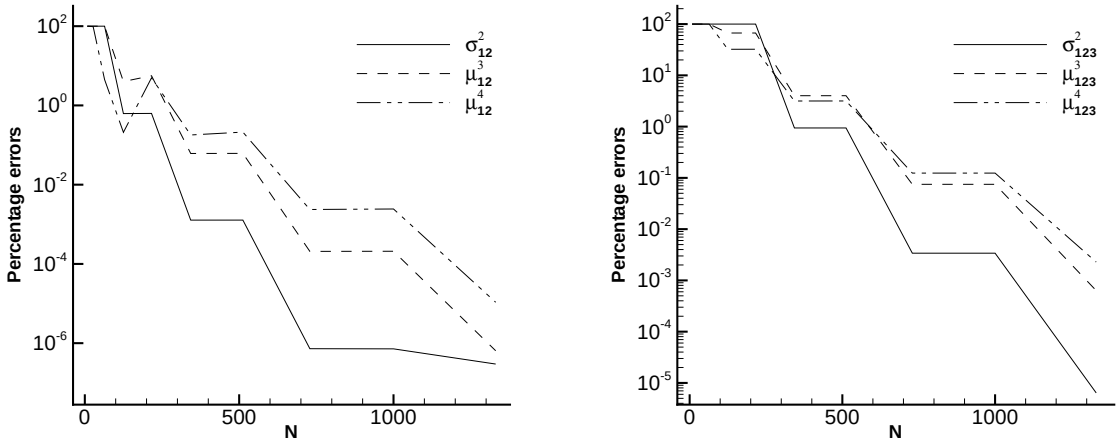
(a) (b)

Figure 2: Conditional statistics vs number of function evaluations in the case $d = 2$.



(a) (b)

Figure 3: Statistical moments (a) and conditional statistics (b) error vs number of function evaluations in the case $d = 3$.



(a) (b)

Figure 4: Sensitivity indices vs number of function evaluations in the case $d = 3$.

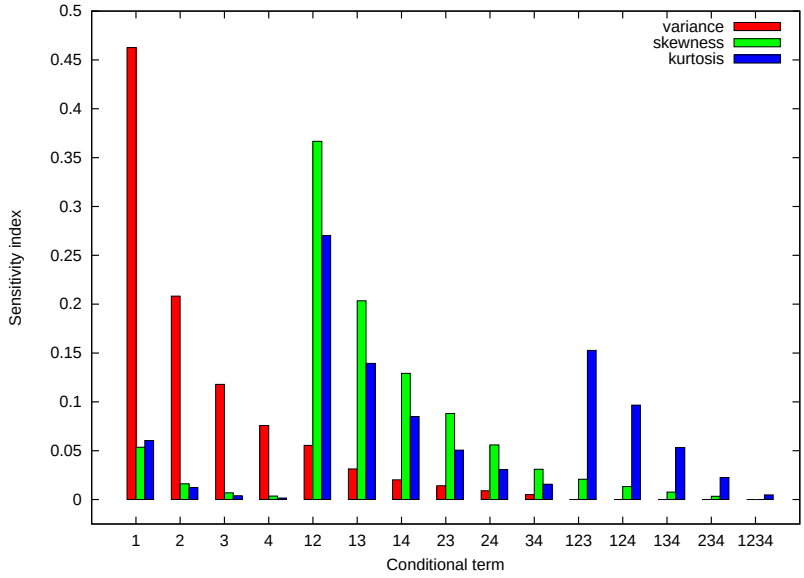


Figure 5: Sensitivity indices for the linear g-function f_{glin} (52) obtained with a PC series with total degree $n_0 = 5$.

By considering the variance, first-order contributions exceed 0.98, while a value larger than 0.5 is computed for high-order interactions for skewness or kurtosis. In this case, the contribution of the first variable exceeds 0.8, while for the skewness or kurtosis, in order to attain the same level, it is necessary to include also the contributions related to the interaction between the first and second variable. In Table 2, total sensitivity indices are reported for the four variables. In this case, variance contributions for both the third and 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 could be considered low. A low level of TSI for the variables ξ_3 and ξ_4 could suggest to reduce the model to the first two variables or neglect the contributions related to the order higher than one in a polynomial metamodel, as for instance the truncated PC series.

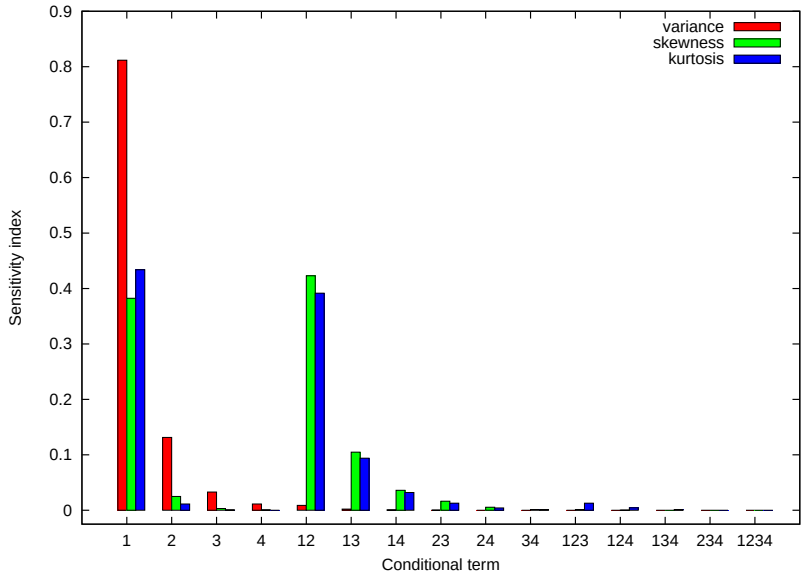


Figure 6: Sensitivity indices for the quadratic g-function f_{gquad} (52) obtained with a PC series with total degree $n_0 = 5$.

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} (52) based on a PC series with total degree $n_0 = 5$.

Let consider the following functions:

$$\begin{aligned} f_1 &= \xi_1 \exp\left(\frac{\xi_2}{\xi_3^2 + 1}\right) + \xi_1 \xi_2 \\ f_2 &= \prod_{i=1}^{3} \frac{2\xi_i + 1}{2}, \end{aligned} \tag{53}$$

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

Sensitivity indices associated to the first function f_1 are reported in Figure 7. For 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. On the contrary, 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, the inspection of the total sensitivity indices, reported in Table 3, suggests that the third variable ξ_3 is meaningless with respect to the variance. The TSI associated to ξ_3 are lower than the limit proposed in [42] to identify a negligible uncertainty. However, if this information is used together with the high-order total sensitivity indices information, the choice of neglecting the third variable should be considered more carefully. The results of a model reduction decision, totally based on variance measures, is further discussed in the following section.

Function f_2 is reported here to underline the difference between the measure of sensitivity associated to the variance and to the higher-order moments. In particular, the functional form of f_2, Eq. (53), includes an equal contribution of three variables. However, looking at Figure 8, it is possible to 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 contrary, are concentrated on second- and third-order interactions, respectively. Even if the sum of the first-order

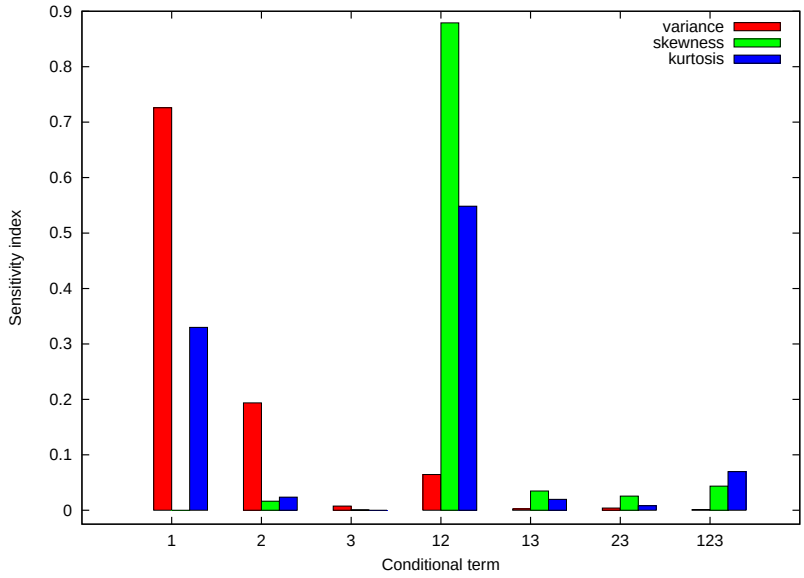


Figure 7: Sensitivity indices for the first function f_1 (53) obtained with a PC series with total degree $n_0 = 7$.

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 (53) based on a PC series with total degree $n_0 = 7$.

variance contributions exceeds 0.9, a reduction of the model in the superposition sense (*i.e.* by neglecting the high orders of interaction), could lead to wrong conclusions, as explained in Section 5.4. The skewness associated to a model including only first-order contributions does not include the skewness information about the probability distribution of the output.

Values for the total sensitivity indices are reported in Table 4 for this case. It is interesting to note that the sum of the total sensitivity indices over the three variables is higher for skewness and kurtosis than for the variance. Then, they reveal an intrinsically high-order interactions (see Eq. (53) for f_2 definition).

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 (53) based on a PC series with total degree $n_0 = 7$.

Numerical test cases presented in this section illustrate how information relative to variance-based sensitivity indices are insufficient in order to understand the true dependence of a model from its variables. Moreover, for a function that is known by points, *i.e.* for example experimental observations or runs of a numerical code, the sensitivity indices on the skewness and on the kurtosis could be very helpful to capture some interactions between subset of variables, much more than the variance.

This could be even more important if the aim is to build/reduce the dimensionality of the problem and to build an accurate metamodel.

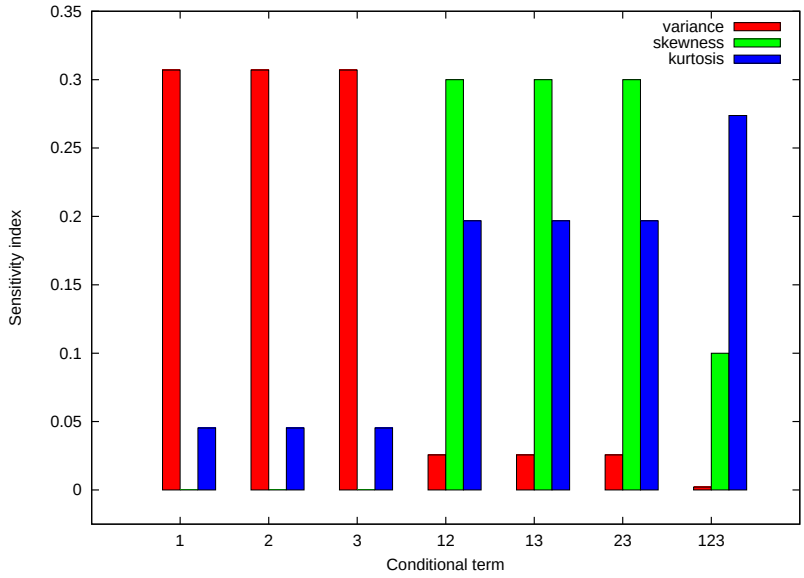


Figure 8: SI for the function f_2.

5.3. Dimensional reduction in the truncation sense

We start from the quadratic g-function, Eq. (52). From the analysis conducted in the previous section (see Table 2), the third and fourth variable seem to be meaningless for the variance-based indices. Their total sensitivity indices sum up to 0.05 for the variance, while exceed 0.15 for both skewness and kurtosis. Considering only the sensitivity indices computed on the variance, one could be tempted to neglect the variables ξ_3 and ξ_4 as follows

$$\begin{aligned} f_{G1} &= f_0 + f_1(\xi_1) + f_2(\xi_2) + f_{12}(\xi_1,\xi_2) \\ f_{G2} &= f_0 + f_1(\xi_1) + f_2(\xi_2) + f_{12}(\xi_1,\xi_2) + f_3(\xi_3) + f_{13}(\xi_1,\xi_3) + f_{23}(\xi_2,\xi_3) + f_{123}(\xi_1,\xi_2,\xi_3), \end{aligned} \tag{54}$$

where in the first case, f_{G1}, both variables are neglected; on the contrary for f_{G2} only ξ_4 is neglected. In this case, the ANOVA terms and the statistics can be computed analytically. In Table 5, the percentage errors, for the first four central moments, are reported with respect to the analytical exact solution for both the reduced models f_{G1} and f_{G2}.

Function	Variance	Skewness	Kurtosis
f_{G1}	4.7997	29.236	15.039
f_{G2}	1.2369	7.7705	4.0632

Table 5: Percentage $\left(\frac{abs(\mu^n - \mu^n_{ex})}{\mu^n_{ex}} \times 100\right)$ errors related to the reduced g-function f_{G1} and f_{G2}.

In Table 5, it is evident that an error of only 5% on the variance can correspond to a much larger error on the higher moments. This behavior is confirmed by at Figure 9, where the probability density function is computed for both f_{G1} and f_{G2} and compared with the complete function, Eq. (52). In this case, f_{G1} which includes only the first two variables can not reproduce the tails while a good approximation is attained elsewhere. In this case, the pdf is bounded between 0.4 and 1.8. If the third variable is included in the reduced model, both variance and kurtosis are computed with an error lower than 5%, while the error on the skewness remains lower than 8%. The total sensitivity indices associated to the fourth variable are reported in Table 2 and it is lower than 5% for the three moments. The improvement of the model given by including the third variable, f_{G2}, is evident in Figure 9, where the pdf of the reduced model better approximates the pdf of the complete function.

From a practical point-of-view, the dimension reduction is commonly accomplished by neglecting some inputs. On the contrary, for an analytical function, it is possible to compute the constant values to choose, for both ξ_3 and ξ_4,

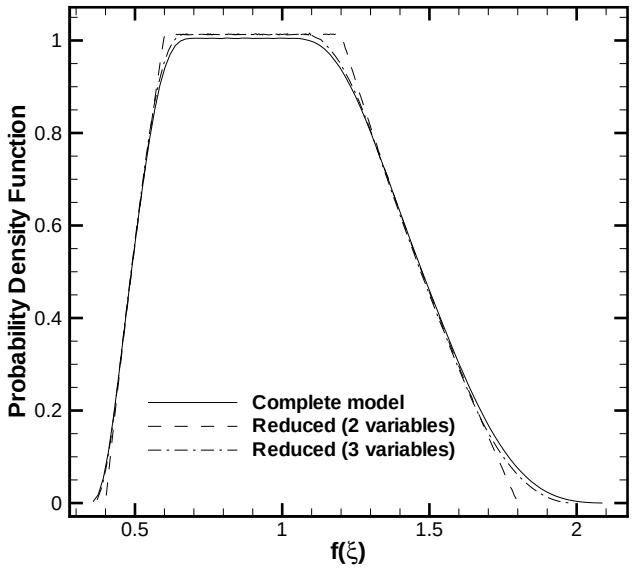


Figure 9: PDFs for the complete g-function and the reduced models (see equations 54).

in order to obtain a reduced model that preserves both the expected value and the variance of the original complete model. Of course, both requirements cannot be satisfied at the same time, but a set of values satisfying the mean and the variance can be obtained analytically requiring that

$$\begin{aligned} \mathbb{E}(f(\vec{\xi})) &= \frac{|4\bar{\xi_j} - 2| + a_j}{1 + a_j} = 1 \\ \mathbb{E}(f^2(\vec{\xi})) &= \left(\frac{|4\bar{\xi_j} - 2| + a_j}{1 + a_j}\right)^2 = \int_0^1 \left(\frac{|4\xi_j - 2| + a_j}{1 + a_j}\right)^2 \mathrm{d}\xi_j. \end{aligned} \tag{55}$$

The following values can be analytically computed for the two variables: $\xi_3 = \{1/4, 3/4, 91/120, 29/120\}$ and $\xi_4 = \{1/4, 3/4, 77/102, 25/102\}$.

In Figure 10, the pdf associated to the complete quadratic g-function with parameters ξ_3 and ξ_4 neglected is reported. In particular, all the possible choices of ξ_3 and ξ_4 are employed obtaining distinct probability functions which are very close each other. This behavior is an additional proof that the third and fourth variable are not significant and their values can only slightly affect the functional response.

From Figure 10, it is evident that neglecting some parameters in order to assure the correctness of the mean and the variance yields a pdf very close to that one obtained by neglecting entirely the ANOVA terms. From a practical point-of-view, the analysis of the reduced model can be carried out both with the ANOVA reduced model (if it is analytically possible to compute the integrals) or by imposing the parameters satisfying the expected value and variance. However, the main point here is that variance-based sensitivity analysis should be supplemented by high-order sensitivity analysis for building a reduced model which does not deteriorate the pdf of the results especially with respect to the tails.

5.4. Dimensional model reduction in the superposition sense

In this section, the problem of the truncation is analyzed from a different perspective, in a so-called *superposition sense*. This means that the dimension of the model is not reduced in terms of number of variables, but in terms of order of interaction between variables for the polynomial metamodel. Note also that, if the function is approximated by means of a PC series of total degree n_0, all the interactions of order $n_0 + 1$ are lost.

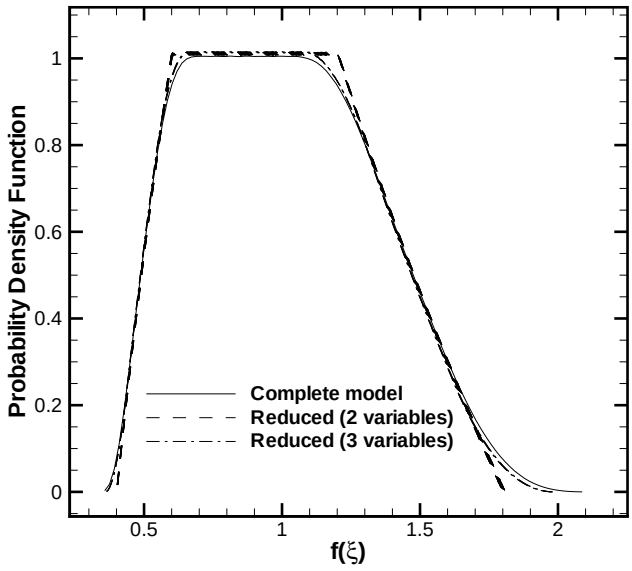


Figure 10: PDFs for the complete g-function and the reduced models.

From a practical standpoint, the error related to the truncation of the PC series is greater than the error of the truncation of the ANOVA approximation at a certain order. This is due to the approximation of each single term of the ANOVA expansion via a truncated polynomial series; the ANOVA functional decomposition contains 2^d terms which are approximated by a finite PC series. In the following, the reduced model is computed analytically. This case relies on a perfect knowledge of the reduced model. However, to obtain an equivalent result for a generic function a non truncated PC series would be necessary.

The first example considered is the linear g-function, Eq. (52). Results presented in Section 5.2 in terms of sensitivity indices (see Figure 5) and total sensitivity indices (see Table 1) illustrate some main features: the first order interaction seems to be important enough to entirely approximate the model; the contributions related to the first order interactions exceed 0.8 for the variance, but it is much less for both skewness and kurtosis. Two different reduced models are considered in this case: the first-order model f_{O1} and the second-order one f_{O2}, described by the following equations

$$\begin{aligned} f_{O1} &= f_0 + f_1(\xi_1) + f_2(\xi_2) + f_3(\xi_3) + f_4(\xi_4) \\ f_{O2} &= f_{O1} + f_{12}(\xi_1, \xi_2) + f_{13}(\xi_1, \xi_3) + f_{14}(\xi_1, \xi_4) + f_{23}(\xi_2, \xi_3) + f_{24}(\xi_2, \xi_4) + f_{34}(\xi_3, \xi_4). \end{aligned} \tag{56}$$

In Table 6, the contributions relative to each moment, for the two models, are reported, where the models are obtained by a truncated PC series and their exact counterparts, f_{O1}^{ex} and f_{O2}^{ex} are computed analytically. It is important to note here that the presence of the absolute value, in both the version of the g-function, prevents the spectral convergence of the PC expansion [28]. Hence, the difference between the PC and the analytical values in Table 6 depends on the PC series truncation.

Figure 11 illustrates the PDF for the complete and the reduced model. Note that including a large amount of variance could lead to a largely inaccurate metamodel if the information on the variance are not supplemented by those obtained by the analysis of high-order moments. From Table 6, it is evident that even if the variance related to the first order terms exceeds the 80% of the total variance, the corresponding skewness and kurtosis are very low. The situation is evident in the probability density function associated to the reduced model f_{O1} reported in Figure 11 where the pdf corresponding to f_{O1} does not contain any skewness (it is perfectly symmetric). The situation greatly improves including contributions up to the second order of interactions between variables. This is supported by the

Function	Variance	Skewness	Kurtosis
f_{O1}	86.46	8.02	7.83
f_{O2}	100.00	95.47	67.00
f_{O1}^{ex}	82.76	0.00	31.51
f_{O2}^{ex}	98.78	81.32	76.52

Table 6: Total contribution for the variance, skewness and kurtosis up to the first and second order (total degree 5) as computed in Section 5.2 and their analytical counterparts.

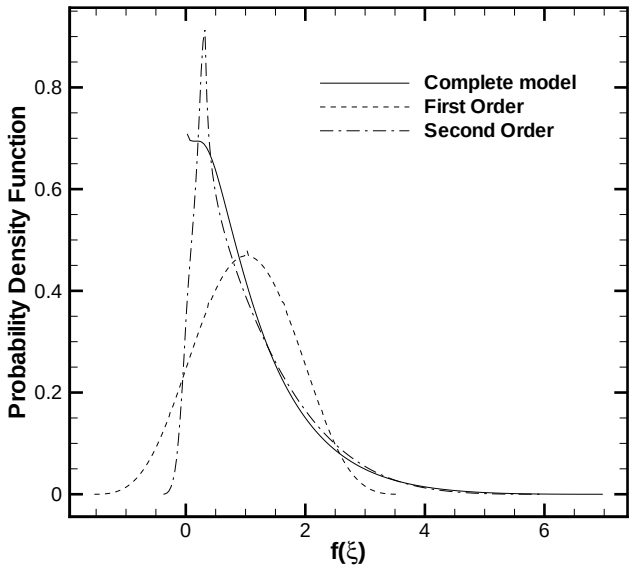


Figure 11: PDFs for the complete linear g-function f_{glin} (see equation (52)) and the reduced models f_{O1}^{ex} and f_{O2}^{ex} (see equations (56)).

statistics reported in Table 6 where, even if the improvement in terms of variance is reduced, a better approximation of both skewness and kurtosis is achieved.

The second example is the function f_2, Eq. (53). The results reported in Section 5.2 show than the first order terms represent more than 90% of the variance while they correspond to the 0% for the skewness and they contribute to less than 15% for the kurtosis. In this case, looking at the sensitivity indices relative to the variance, a model including only first order interaction terms could appear as a good approximation of the complete function. However, considering the first and the second order defined as follows

$$\begin{aligned} f_{O1} &= f_0 + f_1(\xi_1) + f_2(\xi_2) + f_3(\xi_3) \\ f_{O2} &= f_{O1} + f_{12}(\xi_1, \xi_2) + f_{13}(\xi_1, \xi_3) + f_{23}(\xi_2, \xi_3), \end{aligned} \tag{57}$$

the computation of the pdf reveals the importance of the high order interaction terms.

In Figure 12, 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.

5.5. Analysis of a ORCs turbine operating under uncertainty conditions

The last numerical example is a turbine blade of a two dimensional VKI LS-59 cascade, a configuration which has been widely studied [39, 43]. An unstructured CFD dense-gas solver is used to ensure the reliability of the computed results for dense gas flows (for more details see Ref. [39]). The two-dimensional flow domain is discretized by a

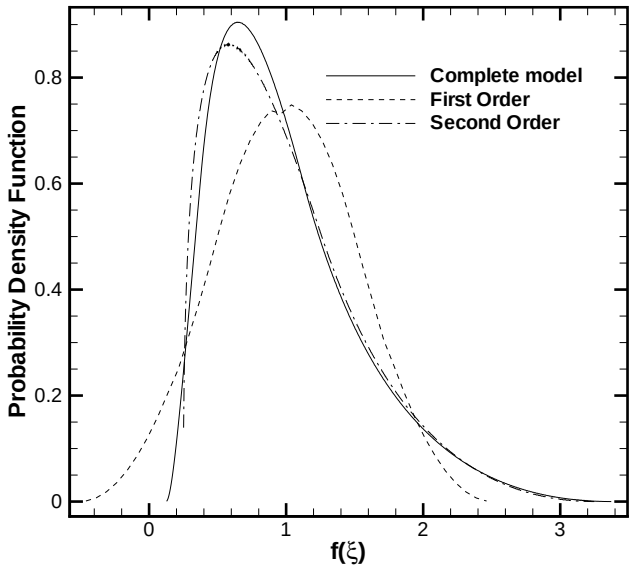


Figure 12: PDFs for the complete f_2 and the reduced models up to the first and second orders.

structured C-grid comprised of 192x16 cells (see Figure 13(a)). The boundary conditions are imposed as follows: at the inlet and outlet boundaries, non-reflecting boundaries are applied using the method of characteristics; a slip condition is imposed at the wall, which uses multi-dimensional linear extrapolation from interior points to calculate the wall pressure; periodicity conditions are prescribed at the inter-blade passage boundaries.

The siloxane *dodecamethylcyclohexasiloxane* ($C_{12}H_{36}Si_6O_6$), commercially known as D_6, is the fluid considered in this study. The physical properties of D_6 are reported in Table 7. The Peng-Robinson-Stryjek-Vera model (PRSV) equation is used as thermodynamic model for D6, which is provided in the non-dimensional form (critical point properties are used as reference) as follows:

$$P_r(T_r, v_r) = \frac{T_r/Z_c}{v_r - b_r} - \frac{a_r}{v_r^2 + 2bv - b^2}, \tag{58}$$

where a_r and b_r are substance-specific parameters related to the fluid critical-point properties:

$$a_r = 0.457235/Z_c^2\alpha(T_r), b_r = 0.077796/Z_c, \tag{59}$$

and Z_c is the critical compressibility factor, *i.e.* $Z_c = (P_c v_c)/(\mathcal{R}T_c)$. It can be computed by enforcing the passage of the critical isotherm through the critical point, and then it is the solution of a cubic equation. The correction factor α is given by

$$\alpha(T) = \left(1 + K(1 - T_r^{1/2})\right)^2, \tag{60}$$

with

$$K = \gamma_0 + \gamma_1\omega - \gamma_2\omega^2 + \gamma_3\omega^3, \tag{61}$$

where the parameter ω is the fluid acentric factor and the coefficients $(\gamma_0, \gamma_1, \gamma_2, \gamma_3)$ are equal to $(0.378893, 1.4897153, -0.17131848, 0.0196554)$, respectively.

To close the thermodynamics relations, a power law for the ideal gas specific heat at constant volume ($c_{v,\infty}$) is assumed

$$c_{v,\infty}(T) = c_{v,\infty}(T_c)\,(T_r)^n\,, \tag{62}$$

Table 7: Thermodynamic data for D_6, where M is the percentage molecular weight, and T_b is the boiling temperature at 1 atm. Properties are taken from Guardone et al.[45].

M (g/mole)	$T_c(K)$	$P_c(kPa)$	$T_b(K)$
444.9	645.8	961	518.1

where n a fluid-dependent parameter and T_c the critical temperature. The equation of state for the internal energy, e, is computed by exploiting the compatibility relations (see [44] for more details).

In summary, the PRSV model depends on the fluid acentric factor ω, the ideal-gas specific heat at the critical temperature $c_{v\infty}(T_c)$, and a fluid-dependent parameter n.

Three sources of uncertainties are considered in this study (globally four uncertainties parameters): uncertainty on the operating conditions, namely the inlet total temperature, T_{in}/T_c and inlet total pressure, p_{in}/p_c; uncertainty on the thermodynamic model, namely n (from previous studies, the large predominance of n with respect to ω, $c_{v\infty}$ has emerged [46, 40]); and uncertainties on geometrical parameters, namely the blade thickness ϕ. Following previous investigations [47], 3.0% of uncertainty for the temperature and pressure levels at the inlet conditions is considered. Since the D6 fluid is considered, the parameter n is assumed to be affected by 6% uncertainty (mean values equal to 0.5729) following [44, 47], with values of ω and $c_{v\infty}$ equal, respectively to 0.7361 and 105.86. An uncertainty of 2% for the thickness ϕ is considered.

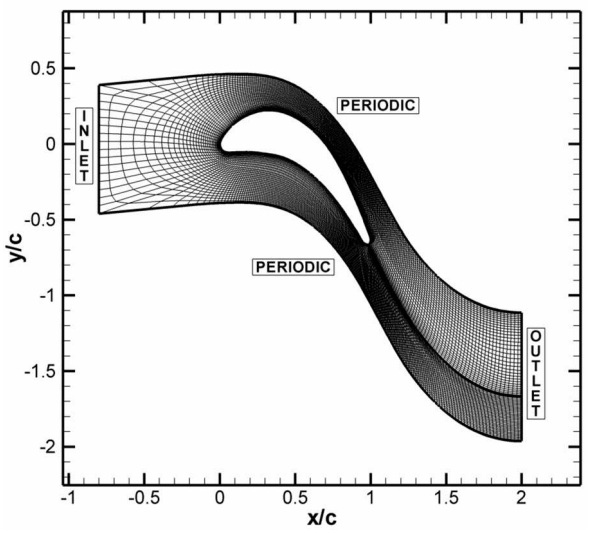
Performance of the turbine cascade can be evaluated by using several output criteria. Statistics decomposition is applied to some specific quantities of interest, which define global performances and efficiency of the turbine. In particular, the following criteria are selected:

- Δh, which is the enthalpy variation through turbine stage;
- The power output per unit depth (PO) expressed as $\Delta h \cdot \dot{m}/w_{mol}$ [W], where $\dot{m}$ is the mass flow rate and w_{mol} is the molecular weight;
- the relative temperature variation or Carnot factor $\Delta T/T_{inlet}$, where ΔT represents the variation of temperature between turbine inlet and outlet boundaries;
- the turbine isentropic efficiency $\Delta h/\Delta h_{ideal}$, where Δh represents the variation of the static enthalpy between turbine inlet and outlet boundaries, and Δh_{ideal} is the variation of enthalpy when an ideal isentropic transformation with the same initial conditions and pressure ratio as the real flow.

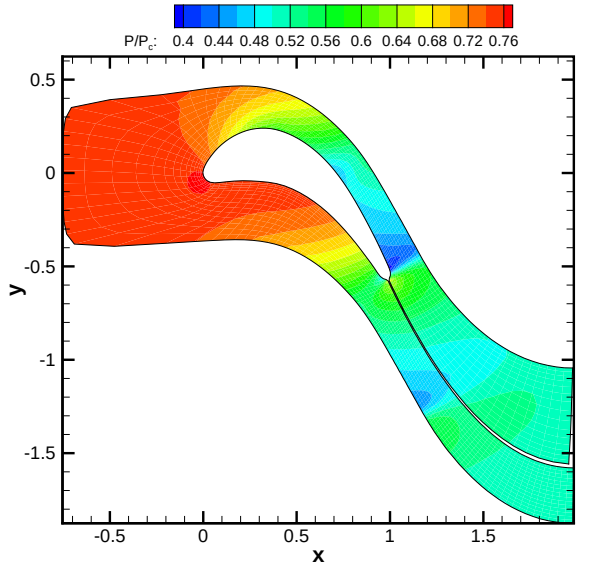
In Figures 14 and 15 the sensitivity indices are reported for the four outputs of interest. The results are based on 4096 simulations obtained at the locations of the tensorization of 1D Legendre polynomials with order 7 in each direction. The PC approximation has been chosen, after a convergence study not reported here for brevity, to use a total polynomial degree of order 5. For all the outputs the third order interaction terms are negligible for any statistic moments. The situation changes when first and second order contributions are considered. The first order terms contribute almost entirely to the overall variance, while for skewness and kurtosis their effect is as important as the second order terms. The interesting feature of the enthalpy variation (Figure 14(b)), the Carnot factor (Figure 15(a)) and the power output (Figure 15(b)) is the presence of negative contributions associated to the skewness. For the enthalpy variation and the power, the first order terms sum up to negative values of the order of the unity. Such a behavior means that the skewness associated to the first order metamodel shows a skewness roughly equal in magnitude to the complete function, but with an opposite sign. The effect of a similar situation is evident comparing the probability density functions of distinct metamodels associated to different superposition lengths.

In Table 8 the *a priori* linear metamodel estimations (see Section 4) are applied for the turbine case.

In Figures 16 and 17 the pdfs for the four outputs of interest are reported. For all the outputs it is evident that the metamodel including up to second order interaction terms is able to represent accurately the original function. For the efficiency, Figure 16(a), and the Carnot factor, Figure 17(a), the skewness associated to the model including only first order interaction terms has the same sign of the overall third moment. For the other outputs, namely the enthalpy variation, Figure 16(b), and the power output, Figure 17(b), the skewness of the first order model has the opposite sign and the effect is an inversion of the tails. However, we note here that the effect of the sign inversion is limited as it is

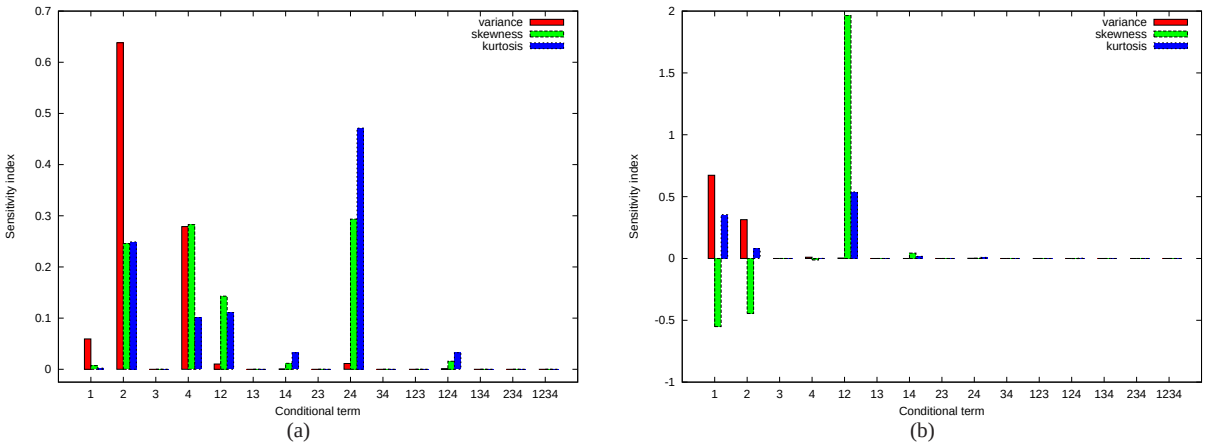


(a)



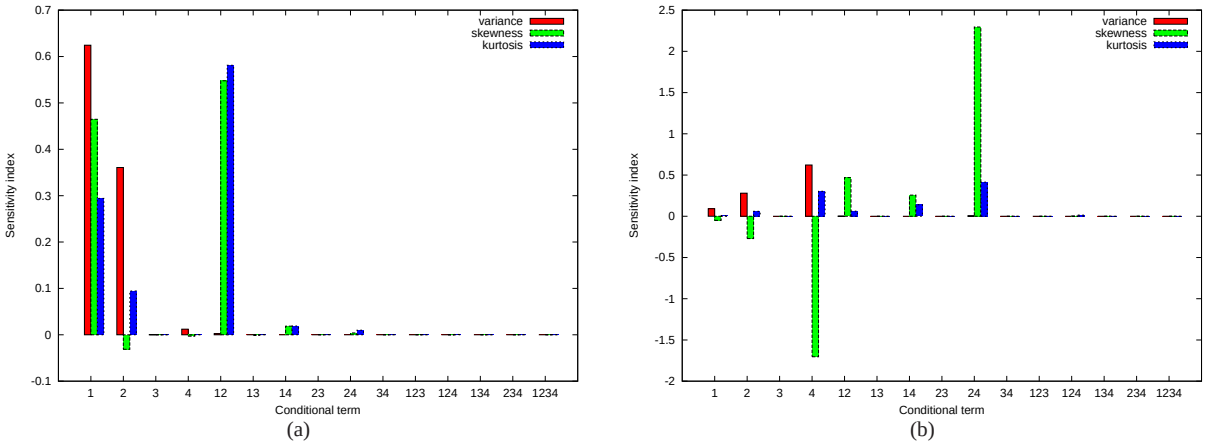
(b)

Figure 13: Computational grid for the turbine (a) and pressure contours for the reference case (b).



(a) (b)

Figure 14: Sensitivity indices for the isentropic efficiency (a) and the enthalpy variation (b).

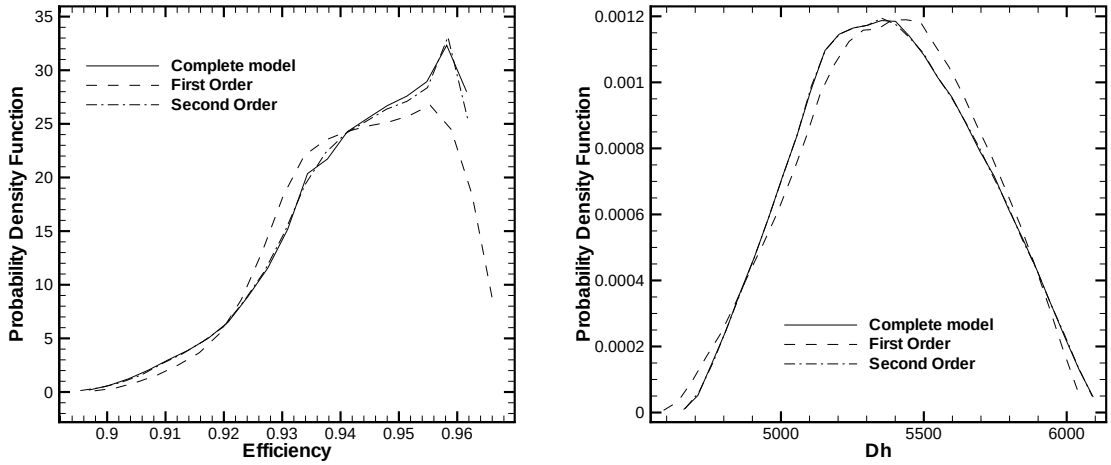


(a) (b)

Figure 15: Sensitivity indices for the Carnot factor (a) and the power output (b).

Output	μ_J^2	$\mathbf{C_{\sigma^2}}$	μ_J^3	C_s	μ_J^4	C_k	$\mathbf{C_s/(C_{\sigma^2})^{3/2}}$	$\mathbf{C_k/(C_{\sigma^2})^2}$
Efficiency	1.627E-4	**9.765E-1**	-8.277E-7	5.360E-1	6.833E-8	8.136E-1	**5.554E-1**	**8.533E-1**
Δh	8.280E4	**9.956E-1**	-1.961E6	-1.009	1.608E10	1.007	**-1.016**	**1.016**
Carnot	2.493E-7	**9.974E-1**	1.411E-11	4.304E-1	1.496E-13	9.594E-1	**4.321E-1**	**9.645E-1**
PO	8.188E-3	**9.931E-1**	-1.140E-4	-2.023	1.650E-4	1.011	**-2.045**	**1.026**

Table 8: Estimation errors for the linear metamodel following Section 4 for the turbine problem. The values relative to the amount of variance, skewness and kurtosis carried by the linear metamodel are reported in bold.



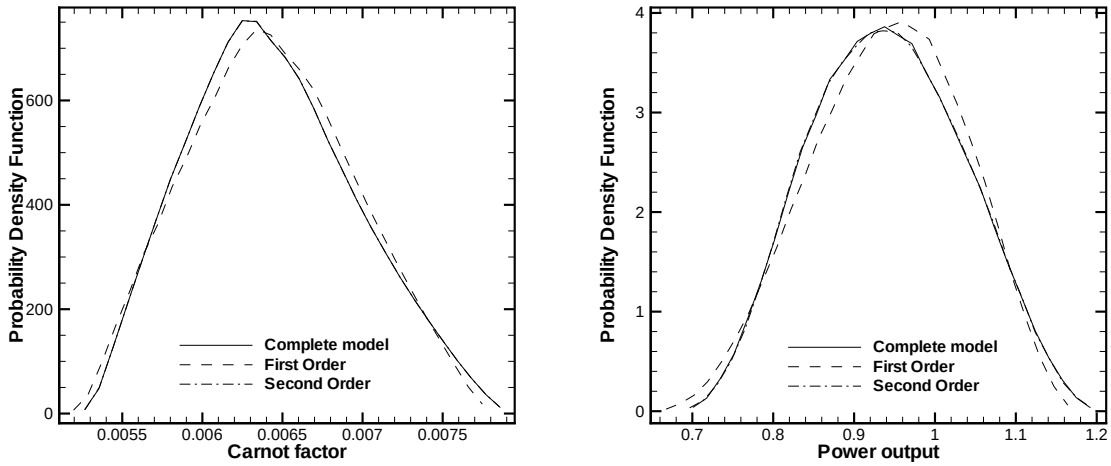
(a) (b)

Figure 16: Probability density functions for the metamodel including the first order interactions, second order interactions and the complete model. Efficiency is shown in (a) and the enthalpy variation in (b).

associated to small values of skewness for the complete model, namely -8.229×10^{-2} for the enthalpy variation and -0.1539 for the power output.

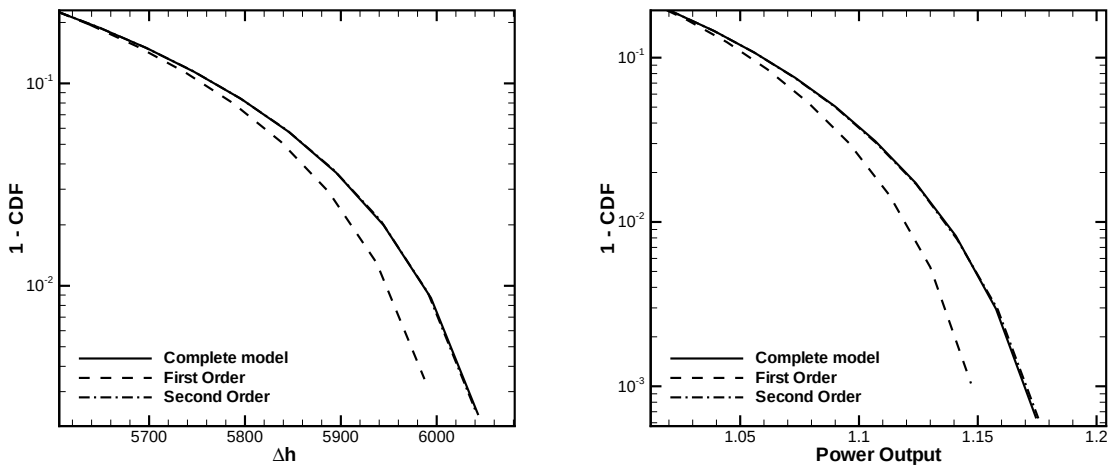
To further demonstrate the relevance of the high-order statistics, for the right tails of the enthalpy variation and power output, for which the metamodel including only first order interaction terms is providing poor approximations in term of skewness, the probability of exceeding a fixed value is reported in Figure 18. The probability of exceeding a fixed value is obtained as the difference between one and the cumulative density function (CDF) value. For instance, if the probability of exceeding the value of 1.5 for the power output is needed, then the metamodel including first order interactions can predict a probability of order 10^{-3}, while the complete model (or the more accurate metamodel) attain a value of order 10^{-2}.

In Section 4 a procedure for the optimal selection of the polynomial degree for a metamodel exploiting the information related to the high-order statistics has been introduced. In Figures 19 and 20 the convergence of the constants C_{σ^2}, $C_s/(C_{\sigma^2})^{3/2}$ and $C_k/(C_{\sigma^2})^2$ are reported as function of the number of coefficients of the metamodel, $n_0 d$, for the different outputs of the ORC turbine.The convergence is reported in term of absolute value of the difference between two successive polynomial degree and it shows how the variance is attained with low errors for low polynomial degrees. In all the cases, but the efficiency, an error of 10^{-2} is reached with order $n_0 = 2$ for the constant relative to the variance, while the convergence is slower for the constants related to the high-order statistics. In particular, if a threshold of 10^{-1} is fixed for the algorithm reported in Section 4 it is evident that an higher, $n_0 > 2$, is always needed. The outputs with slower convergence of the constants related to high-order statistics are the efficiency and the power output, reported in Figures 19(a) and 20(b), respectively. For these outputs, in Figures 21 and 22 the probability density function and the cumulative density function are reported. For both efficiency and power output the threshold of $\varepsilon = 10^{-1}$ is reached with $n_0 = 4$ for the high-order statistics, while $n_0 = 2$ is high enough to provide convergence for the constant related to the variance.



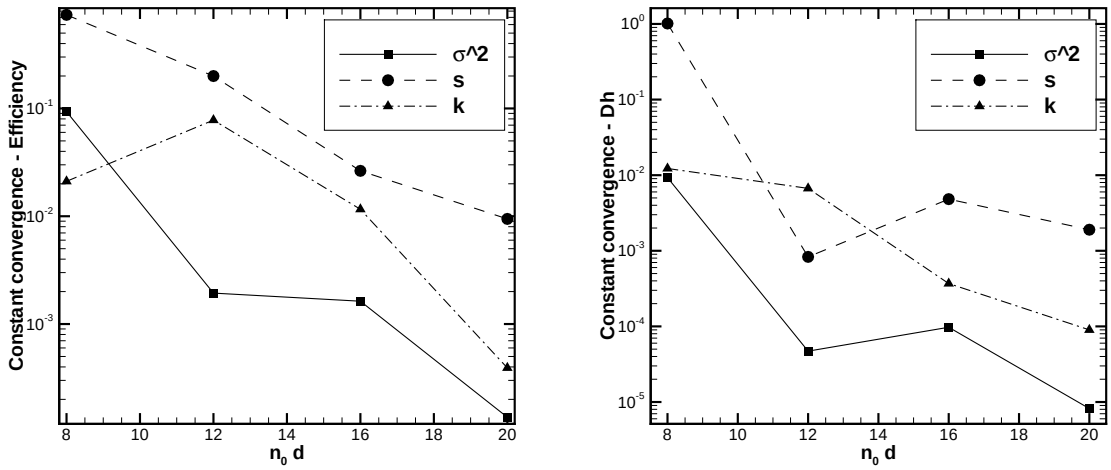
(a) (b)

Figure 17: Probability density functions for the metamodel including the first order interactions, second order interactions and the complete model. Carnot factor is shown in (a) and the power output in (b).



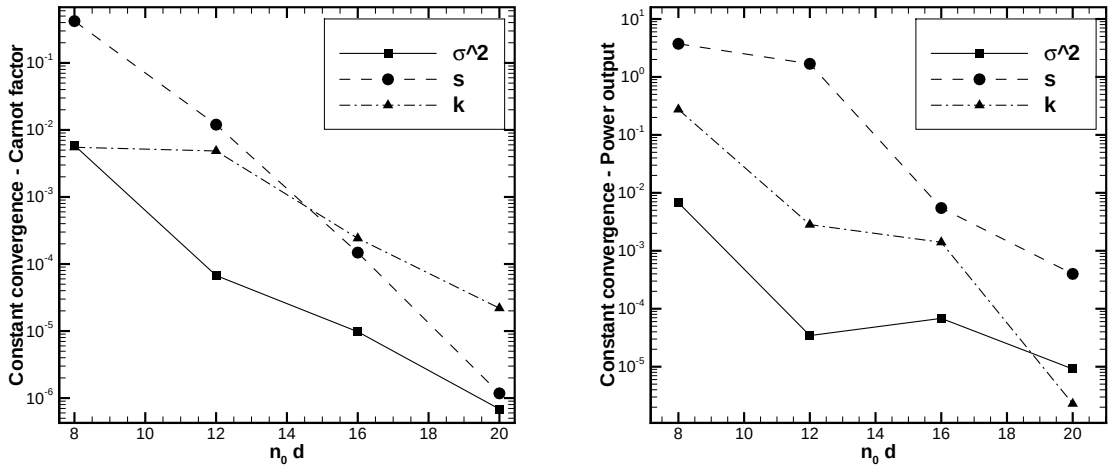
(a) (b)

Figure 18: Probability of exceeding a fixed value is computed as 1 – *CDF* for the enthalpy variation (a) and the power output (b).



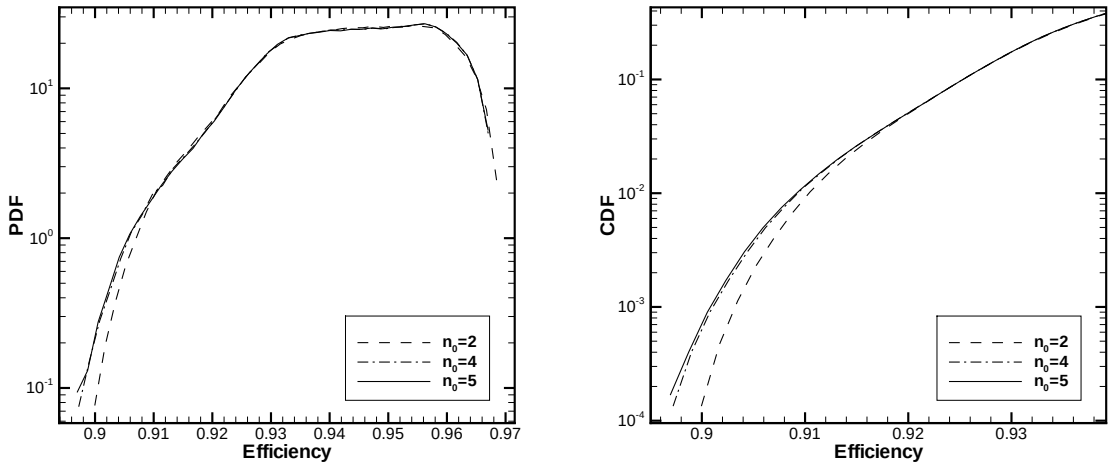
(a) (b)

Figure 19: Convergence of the constants C_{σ^2}, $C_s/(C_{\sigma^2})^{3/2}$ and $C_k/(C_{\sigma^2})^2$ varying the polynomial order n_0 for a metamodel including only first order interactions. Efficiency is reported in (a) while Δh is reported in (b).



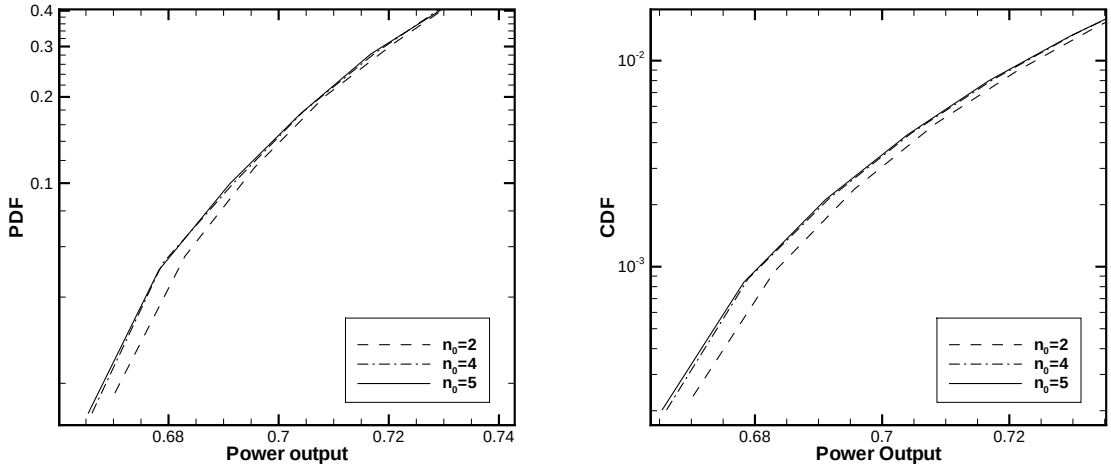
(a) (b)

Figure 20: Convergence of the constants C_{σ^2}, $C_s/(C_{\sigma^2})^{3/2}$ and $C_k/(C_{\sigma^2})^2$ varying the polynomial order n_0 for a metamodel including only first order interactions. The Carnot factor is reported in (a) while the power output is reported in (b).



(a) (b)

Figure 21: Convergence of the PDF (a) and CDF (b) for the efficiency varying the polynomial order n_0 for a metamodel including only first order interaction terms.



(a) (b)

Figure 22: Convergence of the PDF (a) and CDF (b) for the power output varying the polynomial order n_0 for a metamodel including only first order interaction terms.

6. Conclusions

We present a decomposition of high-order statistics and the importance of using this information for reducing the complexity of a uncertainty analysis.

We also present a correlation between the functional decomposition, as depicted by Sobol, and the polynomial chaos development which enables to identify each term of the decomposition, drawing also a practical way to compute all these terms. This procedure is assessed on several test-cases computing the convergence curves obtained by using PC with respect to the reference solution, that is the exact analytical one.

Moreover, sensitivity indices based on skewness and kurtosis decomposition are introduced. The importance of ranking the predominant contributions in terms not only of the variance but also of higher order moments (thus extending the ANOVA analysis to high order statistic moments), is demonstrated with several test functions.

Two different strategies for reducing the complexity of the problem are considered: reducing the number of dimensions or limiting the order of interactions between different variables. Considering high-order statistics is shown to be of fundamental importance for saving the statistics properties of the reduced problem with respect to the complete one.

Future works will be directed towards computationally effective strategies for the reduction of the global computational cost, and the use of high-order statistics in robust design optimization following similar approaches already introduced for the variance [46].

[1] D. Xiu, G. E. Karniadakis, The Wiener–Askey Polynomial Chaos for Stochastic Differential Equations, SIAM Journal on Scientific Computing 24 (2002) 619–644.

[2] D. Xiu, J. S. Hesthaven, High-Order Collocation Methods for Differential Equations with Random Inputs, SIAM Journal on Scientific Computing 27 (2005) 1118.

[3] X. Wan, G. E. Karniadakis, An adaptive multi-element generalized polynomial chaos method for stochastic differential equations, Journal of Computational Physics 209 (2005) 617–642.

[4] R. G. Ghanem, P. D. Spanos, Stochastic Finite Elements. A spectral approach, Springer Verlag, 1991.

[5] F. Nobile, R. Tempone, C. Webster, A sparse grid stochastic collocation method for partial differential equations with random input data, SIAM Journal on Numerical Analysis 46 (2008) 2309–2345.

[6] J. Foo, G. E. Karniadakis, Multi-element probabilistic collocation method in high dimensions, Journal of Computational Physics 229 (2010) 1536–1557.

[7] N. H. Kim, H. Wang, N. V. Queipo, Efficient Shape Optimization Under Uncertainty Using Polynomial Chaos Expansions and Local Sensitivities, Technical Report 5, 2006.

[8] M. Eldred, Recent Advances in Non-Intrusive Polynomial Chaos and Stochastic Collocation Methods for Uncertainty Analysis and Design , AIAA Paper 2009-2274 (2009) –37.

[9] S. A. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, Soviet Math. Dokl. (1963) 240–243.

[10] J. Foo, X. Wan, G. E. Karniadakis, The multi-element probabilistic collocation method (ME-PCM): Error analysis and applications, Journal of Computational Physics 227 (2008) 9572–9595.

[11] J. a. S. Witteveen, G. Loeven, H. Bijl, An adaptive Stochastic Finite Elements approach based on NewtonCotes quadrature in simplex elements, Computers & Fluids 38 (2009) 1270–1288.

[12] J. A. S. Witteveen, G. Iaccarino, Refinement Criteria for Simplex Stochastic Collocation with Local Extremum Diminishing Robustness, 2012.

[13] R. Abgrall, P. M. Congedo, G. Geraci, A one-time truncate and encode multiresolution stochastic framework, Journal of Computational Physics 257 (2014) 19–56.

[14] J. Tryoen, O. Maitre, A. Ern, Adaptive anisotropic spectral stochastic methods for uncertain scalar conservation laws, SIAM Journal on Scientific Computing 34 (2012).

[15] O. Le Maître, Uncertainty propagation using WienerHaar expansions, Journal of Computational Physics 197 (2004) 28–57.

[16] O. Le Maître, O. Knio, H. Najm, R. Ghanem, Multi-resolution analysis of Wiener-type uncertainty propagation schemes, Journal of Computational Physics 197 (2004) 502–531.

[17] I. Babuška, R. Tempone, G. Zouraris, Galerkin finite element approximations of stochastic elliptic differential equations, SIAM Journal on Numerical Analysis 42 (2004) 800–825.

[18] M. H. Keese A., Numerical methods and Smolyak quadrature for nonlinear partial differential equations, Informatikbericht 2003-5 (2003).

[19] E. Novak, K. Ritter, Simple cubature formulas with high polynomial exactness, Constr. Approx. (1999) 499–522.

[20] V. Barthelmann, E. Novak, K. Ritter, High dimensional polynomial interpolation on sparse grids, Adv. Comput. Math., Vol. 12, No. 4 (2000) 273–288.

[21] E. Novak, K. Ritter, High dimensional integration of smooth functions over cubes, Numerische Mathematik (1996) 79–97.

[22] K. Petras, On the Smolyak cubature error for analytic functions, Advances in Computational Mathematics 12 (2000) 71–93.

[23] E. Borgonovo, G. E. Apostolakis, S. Tarantola, A. Saltelli, Comparison of global sensitivity analysis techniques and importance measures in PSA, Reliability Engineering & System Safety 79 (2003) 175–185.

[24] I. Sobol, Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Mathematics and Computers in Simulation 55 (2001) 271–280.

[25] H. Rabitz, O. Alis, J. Shorter, K. Shim, Efficient input-output model representations., Comput. Phys. Commun. 117 (1999) 11–20.

[26] A. Saltelli, Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index., Comput. Phys. Commun. 181 (2010) 259–270.
[27] B. Sudret, Global sensitivity analysis using polynomial chaos expansions, Reliability Engineering and System Safety 93 (2008) 964–979.
[28] T. Crestaux, O. Le Maître, J.-M. Martinez, Polynomial chaos expansion for sensitivity analysis, Reliability Engineering & System Safety 94 (2009) 1161–1172.
[29] G. Blatman, B. Sudret, A comparison of three metamodel-based methods for global sensitivity analysis: GP modelling, HDMR and LAR-gPC, Procedia - Social and Behavioral Sciences 2 (2010) 7613–7614.
[30] X. Yanga, M. Choi, G. Lin, G. E. Karniadakis, Adaptive ANOVA Decomposition of Stochastic Incompressible and Compressible Flows, Journal of Computational Physics 231 (2012) 1587–1614.
[31] G. Blatman, B. Sudret, Efficient computation of global sensitivity indices using sparse polynomial chaos expansions, Reliability Engineering and System Safety 95 (2010) 1216–1229.
[32] C. Xu, G. Gertner, Extending a global sensitivity analysis technique to models with correlated parameters., Computational Statistics & Data Analysis 51 (2007) 5579–5590.
[33] C. Xu, G. Gertner, Uncertainty and sensitivity analysis for models with correlated parameters., Reliability Engineering & System Safety 93 (2008) 1563–1573.
[34] E. Borgonovo, A new uncertainty importance measure, Reliability Engineering & System Safety 92 (2007) 771–784.
[35] J. Y. Caniou, B. Sudret, Distribution-based global sensitivity analysis in case of correlated input parameters using polynomial chaos expansions, in: ICASP2011.
[36] E. Borgonovo, W. Castaings, S. Tarantola, Model emulation and moment-independent sensitivity analysis: An application to environmental modelling, Environmental Modelling & Software 34 (2012) 105–115.
[37] B. Sudret, Global sensitivity analysis using polynomial chaos expansions, in: Proc. 5th Int. Conf. on Comp. Stoch. Mech (CSM5), Rhodos, Greece.
[38] B. Sudret, Uncertainty propagation and sensitivity analysis in mechanical models Contributions to structural reliability and stochastic spectral methods, Habilitation à Diriger des Recherches, Université BLAISE PASCAL - Clermont II (2007) 1–252.
[39] P. Congedo, C. Corre, P. Cinnella, Numerical investigation of dense-gas effects in turbomachinery, Computers & Fluids 49 (2011) 290–301.
[40] P. M. Congedo, C. Corre, J.-M. Martinez, Shape optimization of an airfoil in a BZT flow with multiple-source uncertainties, Computer Methods in Applied Mechanics and Engineering 200 (2010) 216–232.
[41] R. E. Caflisch, W. Morokoff, A. B. Owen, Valuation of Mortgage Backed Securities Using Brownian Bridges to Reduce Effective Dimension, Journal of Computational Finance 1 (1997) 26–47.
[42] Z. Gao, J. S. Hesthaven, Efficient solution of ordinary differential equations with high-dimensional parametrized uncertainty, Communications in computational physics (2010) 1–33.
[43] R. Kiock, F. Lehthaus, N. C. Baines, C. H. Sieverding, The transonic flow through a plane turbine cascade as measured in four european wind tunnels, ASME J Eng Gas Turb Power 108 (1986) 277–284.
[44] P. Cinnella, P. M. Congedo, V. Pediroda, L. Parussini, Sensitivity analysis of dense gas flow simulations to thermodynamic uncertainties, Physics of Fluids 23 (2011) 116101:1–20.
[45] A. Guardone, C. Zamfirescu, P. Colonna, Maximum intensity of rarefaction shock waves for dense gases, Journal of Fluid Mechanics 642 (2010) 127–146.
[46] P. Congedo, G. Geraci, R. Abgrall, V. Pediroda, L. Parussini, TSI metamodels-based multi-objective robust optimization, Engineering Computations (Swansea, Wales) 30 (2013) 1032–1053.
[47] P. Colonna, J. Harinck, S. Rebay, A. Guardone, Real-gas effects in organic rankine cycle turbine nozzles, AIAA J Prop Power 24 (2008) 282–294.
[48] O. Le Maître, O. Knio, Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics, Springer Verlag, 2010.

Appendix A. Third central moment expression for the ANOVA functional decomposition

In this section, the skewness form Eq. (12) is computed using the expression obtained via the multinomial theorem applied to the functional ANOVA decomposition Eq. (1), and after the integration over Ξ it becomes

$$\mu^3 = \sum_{p=1}^{N} \int_{\hat{\Xi}_p} f^3_{\vec{m}_p} \mathrm{d}\mathbb{P}_p + 3 \sum_{p=1}^{N} \sum_{\substack{q=1 \\ q\neq p}}^{N} \int_{\hat{\Xi}_{pq}} f^2_{\vec{m}_p} f_{\vec{m}_q} \mathrm{d}\mathbb{P}_{pq} + 6 \sum_{p=1}^{N} \sum_{q=p+1}^{N} \sum_{r=q+1}^{N} \int_{\hat{\Xi}_{pqr}} f_{\vec{m}_p} f_{\vec{m}_q} f_{\vec{m}_r} \mathrm{d}\mathbb{P}_{pqr}. \tag{A.1}$$

The previous equation displays interactions between variables and sub-sets of variables. In particular, the second and the third terms of could be simplified to highlight the contributions that are always equal to zero (due to the orthogonality).

The second term on the right hand side can be simplified as follows

$$3 \sum_{p=1}^{N} \sum_{\substack{q=1 \\ q\neq p}}^{N} \int_{\hat{\Xi}_{pq}} f^2_{\vec{m}_p} f_{\vec{m}_q} \mathrm{d}\mathbb{P}_{pq} = 3 \sum_{\vec{m}_p} \sum_{\vec{m}_q \subset \vec{m}_p} \int_{\hat{\Xi}_{pq}} f^2_{\vec{m}_p} f_{\vec{m}_q} \mathrm{d}\mathbb{P}_{pq}. \tag{A.2}$$

Proof. This term expresses the interaction between two multi-indices $\vec{m}_p$ and $\vec{m}_q$, where one of them is raised to the second power. The two multi-indices should be different for construction, *i.e.* $\vec{m}_p \neq \vec{m}_q$: if $\vec{m}_q$ is not a subset of $\vec{m}_p$, then a set of coordinates $\vec{m}_q \setminus \cap_{pq}$ belongs only to $\vec{m}_q$ (if the set is totally disjointed, then the term $\cap_{pq}$ is the null set). Note that the symbol $\cap_{pq}$ indicates the coordinates contained in both the multi-indices p and q.

The integral reformulated into the form

$$\int_{\hat{\Xi}_{pq}} f^2_{\vec{m}_p} f_{\vec{m}_q} \mathrm{d}\mathbb{P}_{pq} = \int_{\hat{\Xi}_{p+\cap pq}} f^2_{\vec{m}_p} \left(\int_{\hat{\Xi}_{q\setminus\cap pq}} f_{\vec{m}_q} \mathrm{d}\mathbb{P}_{q\setminus\cap pq} \right) \mathrm{d}\mathbb{P}_{p+\cap_{pq}} = 0, \tag{A.3}$$

where the internal integral is equal to zero due to the orthogonality of the ANOVA contributions (see equation (4)). □

The case of the identification of the contributions to a specific multi-index $\vec{m}_i$ is now addressed. The resulting is the set of the variables $\vec{m}_{pq} = \vec{m}_i$. If all the sub-sets of variables and their interactions of $\vec{m}_i$ are collected in the set $\mathcal{P}_i$, the contributions to $\vec{m}_i$ are a sub set of the $2^{|\mathcal{P}_i|} - 1$ simple combinations of class two. However, not all the combinations of sums $\vec{m}_p \boxplus \vec{m}_q$ contribute to the conditional term $\mu^3_{\vec{m}_i}$, for instance $(1,0,0) \boxplus (1,1,0) \neq \vec{m}_i$. Thus, requiring simultaneously $\vec{m}_q \subset \vec{m}_p$ and $\vec{m}_{pq} = \vec{m}_i$ allows to identify the non-null contributions as

$$3 \sum_{\vec{m}_p \in \mathcal{P}_i} \sum_{\substack{\vec{m}_{pq} = \vec{m}_i \\ \mathcal{P}_i \ni \vec{m}_q \subset \vec{m}_p}} \int_{\hat{\Xi}_{pq}} f^2_{\vec{m}_p} f_{\vec{m}_q} \mathrm{d}\mathbb{P}_{pq} = 3 \int_{\hat{\Xi}_i} f^2_{\vec{m}_i} \sum_{\vec{m}_q \in \mathcal{P}_{i,\neq}} f_{\vec{m}_q} \mathrm{d}\mathbb{P}_i, \tag{A.4}$$

where $\mathcal{P}_{i,\neq}$ is employed as shorthand for $\mathcal{P}_{i,\neq} = \mathcal{P}_i - \{\vec{m}_i\}$.

The last term of (A.1) can be written as follows

$$6 \sum_{p=1}^{N} \sum_{q=p+1}^{N} \sum_{r=q+1}^{N} \int_{\hat{\Xi}_{pqr}} f_{\vec{m}_p} f_{\vec{m}_q} f_{\vec{m}_r} \mathrm{d}\mathbb{P}_{pqr} = 6 \sum_{p=1}^{N} \sum_{q=p+1}^{N} \sum_{\substack{r=q+1 \\ \vec{m}_{pq} = \vec{m}_r}}^{N} \int_{\hat{\Xi}_{pq}} f_{\vec{m}_p} f_{\vec{m}_q} f_{\vec{m}_r} \mathrm{d}\mathbb{P}_{pq}. \tag{A.5}$$

Proof. This case can be demonstrated extending what already done for the dyadic interaction between multi-indices:

$$\int_{\hat{\Xi}_{pqr}} f_{\vec{m}_p} f_{\vec{m}_q} f_{\vec{m}_r} \mathrm{d}\mathbb{P}_{pqr} = \int_{\hat{\Xi}_{pq+\cap pqr}} f_{\vec{m}_p} f_{\vec{m}_q} \left(\int_{\hat{\Xi}_{r\setminus\cap pqr}} f_{\vec{m}_r} \mathrm{d}\mathbb{P}_{r\setminus\cap pqr} \right) \mathrm{d}\mathbb{P}_{pq+\cap_{pqr}} = 0, \tag{A.6}$$

by using the orthogonality property. □

If a specific index $\vec{m}_i$ is of interest, the conditional contribution is identified requiring $\vec{m}_{pqr} = \vec{m}_i$. Thus, the condition $\vec{m}_{pq} = \vec{m}_r = \vec{m}_i$ identifies the non-null contributions

$$\mu^3_{\vec{m}_i} = 6 \sum_{\vec{m}_p \neq \vec{m}_i} \sum_{\substack{\vec{m}_p \neq \vec{m}_q \neq \vec{m}_i \\ \vec{m}_{pq} = \vec{m}_i}} \int_{\hat{\Xi}_i} f_{\vec{m}_p} f_{\vec{m}_q} f_{\vec{m}_i} \mathrm{d}\mathbb{P}_i. \tag{A.7}$$

In conclusion, the final form for the third-order central moment μ^3 is equal to

$$\mu^3 = \sum_{p=1}^{N} \int_{\hat{\Xi}_p} f^3_{\vec{m}_p} \mathrm{d}\mathbb{P}_p + 3 \sum_{\vec{m}_p} \sum_{\vec{m}_q \subset \vec{m}_p} \int_{\hat{\Xi}_{pq}} f^2_{\vec{m}_p} f_{\vec{m}_q} \mathrm{d}\mathbb{P}_{pq} + 6 \sum_{p=1}^{N} \sum_{q=p+1}^{N} \sum_{\substack{r=q+1 \\ \vec{m}_{pq} = \vec{m}_r}}^{N} \int_{\hat{\Xi}_{pq}} f_{\vec{m}_p} f_{\vec{m}_q} f_{\vec{m}_r} \mathrm{d}\mathbb{P}_{pq},$$

where each conditional contribution ($\mu^3 = \sum \mu^3_{\vec{m}_i}$) is

$$\mu^3_{\vec{m}_i} = \int_{\hat{\Xi}_i} f^3_{\vec{m}_i} \mathrm{d}\mathbb{P}_i + 3 \int_{\hat{\Xi}_i} f^2_{\vec{m}_i} \sum_{\vec{m}_q \in \mathcal{P}_{i,\neq}} f_{\vec{m}_q} \mathrm{d}\mathbb{P}_i + 6 \sum_{\vec{m}_p \in \mathcal{P}_{i,\neq}} \sum_{\substack{\vec{m}_p \neq \vec{m}_q \in \mathcal{P}_{i,\neq} \\ \vec{m}_{pq} = \vec{m}_i}} \int_{\hat{\Xi}_i} f_{\vec{m}_i} f_{\vec{m}_p} f_{\vec{m}_q} \mathrm{d}\mathbb{P}_i.$$

In the previous expression we note that the second and third term should include the product of $\vec{m}_i$ and all the multi-indices relative to the sub-interactions, i.e. $\mathcal{P}_i$. Moreover, $\mathcal{P}_{i,\neq}$ is used as shorthand for $\mathcal{P}_i \setminus \vec{m}_i$.

Appendix B. Fourth central moment expression for the ANOVA functional decomposition

In this section, Eq. (18) is obtained starting from the application of the multinomial theorem to Eq. (1)

$$\begin{aligned}\mu^4 = &\sum_{p=1}^{N}\int_{\hat{\Xi}_p} f_{\vec{\boldsymbol{m}}_p}^4 \mathrm{d}\mathbb{P}_p + 4\sum_{p=1}^{N}\sum_{\substack{q=1\\q\neq p}}^{N}\int_{\hat{\Xi}_{pq}} f_{\vec{\boldsymbol{m}}_p}^3 f_{\vec{\boldsymbol{m}}_q}\mathrm{d}\mathbb{P}_{pq} + 6\sum_{p=1}^{N}\sum_{q=p+1}^{N}\int_{\hat{\Xi}_{pq}} f_{\vec{\boldsymbol{m}}_p}^2 f_{\vec{\boldsymbol{m}}_q}^2\mathrm{d}\mathbb{P}_{pq}\\ &+ 12\sum_{p=1}^{N}\sum_{\substack{q=1\\q\neq p}}^{N}\sum_{\substack{r=j+1\\r\neq p}}^{N}\int_{\hat{\Xi}_{pqr}} f_{\vec{\boldsymbol{m}}_p}^2 f_{\vec{\boldsymbol{m}}_q} f_{\vec{\boldsymbol{m}}_r}\mathrm{d}\mathbb{P}_{pqr} + 24\sum_{p=1}^{N}\sum_{q=p+1}^{N}\sum_{r=q+1}^{N}\sum_{t=r+1}^{N}\int_{\hat{\Xi}_{pqrt}} f_{\vec{\boldsymbol{m}}_p} f_{\vec{\boldsymbol{m}}_q} f_{\vec{\boldsymbol{m}}_r} f_{\vec{\boldsymbol{m}}_t}\mathrm{d}\mathbb{P}_{pqrt}.\end{aligned} \tag{B.1}$$

The first three terms are easy to handle. The first and the third one, on the right side, cannot be further simplified, while the second one can be manipulated as shown in the previous section. Thus, it is possible to write

$$4\sum_{p=1}^{N}\sum_{\substack{q=1\\q\neq p}}^{N}\int_{\hat{\Xi}_{pq}} f_{\vec{\boldsymbol{m}}_p}^3 f_{\vec{\boldsymbol{m}}_q}\mathrm{d}\mathbb{P}_{pq} = 4\sum_{\vec{\boldsymbol{m}}_p}\sum_{\vec{\boldsymbol{m}}_q\subset\vec{\boldsymbol{m}}_p}\int_{\hat{\Xi}_{pq}} f_{\vec{\boldsymbol{m}}_p}^3 f_{\vec{\boldsymbol{m}}_q}\mathrm{d}\mathbb{P}_{pq}. \tag{B.2}$$

as already demonstrated for the skewness term.

More attention should be devoted to the last two terms. The first one of them can be written as

$$12\sum_{p=1}^{N}\sum_{\substack{q=1\\q\neq p}}^{N}\sum_{\substack{r=j+1\\r\neq p}}^{N}\int_{\hat{\Xi}_{pqr}} f_{\vec{\boldsymbol{m}}_p}^2 f_{\vec{\boldsymbol{m}}_q} f_{\vec{\boldsymbol{m}}_r}\mathrm{d}\mathbb{P}_{pqr} = 12\sum_{p=1}^{N}\sum_{\substack{q=1\\q\neq p}}^{N}\sum_{\substack{r=q+1\\\vec{\boldsymbol{m}}_{qr}\setminus\cap_{qr}\subseteq\vec{\boldsymbol{m}}_p}}^{N}\int_{\hat{\Xi}_{pqr}} f_{\vec{\boldsymbol{m}}_p}^2 f_{\vec{\boldsymbol{m}}_q} f_{\vec{\boldsymbol{m}}_r}\mathrm{d}\mathbb{P}_{pqr}. \tag{B.3}$$

Proof. If the multi-indices $\vec{\boldsymbol{m}}_q$ and $\vec{\boldsymbol{m}}_r$ are totally independent from the variables contained in $\vec{\boldsymbol{m}}_p$, the condition $\int_{\hat{\Xi}_{qr}} f_{\vec{\boldsymbol{m}}_q} f_{\vec{\boldsymbol{m}}_r}\mathrm{d}\mathbb{P}_{qr} = 0$ holds (due to the orthogonality of the ANOVA functional components). In the general case, when $\vec{\boldsymbol{m}}_{qr}\cap\vec{\boldsymbol{m}}_p \neq 0$, the existence of a null integral is related to the presence of variables in the multi-indices $\vec{\boldsymbol{m}}_q$ or $\vec{\boldsymbol{m}}_r$ which are not contained in $\vec{\boldsymbol{m}}_p$. If $\vec{\boldsymbol{m}}_{qr}\setminus\cap_{qr}\nsubseteq\vec{\boldsymbol{m}}_p$ then it is possible to write as follows

$$\int_{\hat{\Xi}_{p+\cap_{qr}}} f_{\vec{\boldsymbol{m}}_p}^2\left(\int_{\hat{\Xi}_{\vec{\boldsymbol{m}}_{qr}\setminus\cap_{qr}}} f_{\vec{\boldsymbol{m}}_q} f_{\vec{\boldsymbol{m}}_r}\mathrm{d}\mathbb{P}_{qr\setminus\cap_{qr}}\right)\mathrm{d}\mathbb{P}_p = 0. \tag{B.4}$$

Note that the internal integral is carried out with respect to a variable contained only in one of $\vec{\boldsymbol{m}}_q$ or $\vec{\boldsymbol{m}}_r$, then is always zero due to orthogonality. Obviously, the case with the subsets related to $\vec{\boldsymbol{m}}_{qr}$ and $\vec{\boldsymbol{m}}_p$ totally disjointed, is included in the previous condition. □

If a specific multi-index $\vec{\boldsymbol{m}}_i$ is provided, then in this case the contribution of this term is computed as follows

$$12\sum_{\vec{\boldsymbol{m}}_p}\sum_{\vec{\boldsymbol{m}}_p\neq\vec{\boldsymbol{m}}_q\in\mathcal{P}_i}\sum_{\substack{\vec{\boldsymbol{m}}_r\in\mathcal{P}_i, r>q\\\vec{\boldsymbol{m}}_p\boxplus\cap_{qr}=\vec{\boldsymbol{m}}_i}}\int_{\hat{\Xi}_i} f_{\vec{\boldsymbol{m}}_p}^2 f_{\vec{\boldsymbol{m}}_q} f_{\vec{\boldsymbol{m}}_r}\mathrm{d}\mathbb{P}_i. \tag{B.5}$$

Proof. The previous equation can be obtained considering the requirements $\vec{\boldsymbol{m}}_{pqr} = \vec{\boldsymbol{m}}_i$ and $\vec{\boldsymbol{m}}_{qr}\setminus\cap_{qr}\subseteq\vec{\boldsymbol{m}}_p$, with $\vec{\boldsymbol{m}}_p, \vec{\boldsymbol{m}}_q, \vec{\boldsymbol{m}}_r \in \mathcal{P}_i$. It is easy to verify that if the second equation is true, then it must follow $\vec{\boldsymbol{m}}_p\boxplus\vec{\boldsymbol{m}}_{qr}\setminus\cap_{qr}\subseteq\vec{\boldsymbol{m}}_p\boxplus\vec{\boldsymbol{m}}_p = \vec{\boldsymbol{m}}_p$, from which $\vec{\boldsymbol{m}}_i\setminus\cap_{qr}\subseteq\vec{\boldsymbol{m}}_p$. Finally, $\vec{\boldsymbol{m}}_i = \vec{\boldsymbol{m}}_p\boxplus\cap_{qr}$ holds (the equality sign follows from $\vec{\boldsymbol{m}}_p, \vec{\boldsymbol{m}}_q$ and $\vec{\boldsymbol{m}}_r\in\mathcal{P}_i$). Remark that great attention must be paid in manipulating expressions with the summation of multi-indices $\boxplus$. Generally, consider that $\vec{\boldsymbol{m}}_p\setminus\vec{\boldsymbol{m}}_q = \vec{\boldsymbol{m}}_r \Rightarrow \vec{\boldsymbol{m}}_p = \vec{\boldsymbol{m}}_r\boxplus\vec{\boldsymbol{m}}_q$ holds but the contrary is not guaranteed. □

The last term of the kurtosis can be also reformulated as follows

$$24\sum_{p=1}^{N}\sum_{q=p+1}^{N}\sum_{r=q+1}^{N}\sum_{t=r+1}^{N}\int_{\hat{\Xi}_{pqrt}} f_{\vec{\boldsymbol{m}}_p} f_{\vec{\boldsymbol{m}}_q} f_{\vec{\boldsymbol{m}}_r} f_{\vec{\boldsymbol{m}}_t}\mathrm{d}\mathbb{P}_{pqrt} = 24\sum_{p=1}^{N}\sum_{q=p+1}^{N}\sum_{\substack{r=q+1\\\vec{\boldsymbol{m}}_{pq}\setminus\cap_{pq}\subseteq\vec{\boldsymbol{m}}_{rt}\subseteq\vec{\boldsymbol{m}}_{pq}\boxplus\cap_{rt}}}^{N}\sum_{t=r+1}^{N}\int_{\hat{\Xi}_{pqrt}} f_{\vec{\boldsymbol{m}}_p} f_{\vec{\boldsymbol{m}}_q} f_{\vec{\boldsymbol{m}}_r} f_{\vec{\boldsymbol{m}}_t}\mathrm{d}\mathbb{P}_{pqrt}. \tag{B.6}$$

Proof. This term can be obtained with some constraints on the multi-indices: they should share, two by two, some sets of coordinates: $\vec{\boldsymbol{m}}_{\boldsymbol{pq}} \setminus \cap_{pq} \subseteq \vec{\boldsymbol{m}}_{\boldsymbol{rt}}$ and $\vec{\boldsymbol{m}}_{\boldsymbol{rt}} \setminus \cap_{rt} \subseteq \vec{\boldsymbol{m}}_{\boldsymbol{pq}}$. The fulfillment of the previous conditions assures that the integral cannot be divided into a product between integrals of orthogonal contributions. These conditions could be applied choosing randomly two couple of indices. Using the second constraint, $\vec{\boldsymbol{m}}_{\boldsymbol{rt}} \subseteq \vec{\boldsymbol{m}}_{\boldsymbol{pq}} \boxplus \cap_{rt}$ is obtained. Using this relation with the first requirement, $\vec{\boldsymbol{m}}_{\boldsymbol{pq}} \setminus \cap_{pq} \subseteq \vec{\boldsymbol{m}}_{\boldsymbol{rt}} \subseteq \vec{\boldsymbol{m}}_{\boldsymbol{pq}} \boxplus \cap_{rt}$ holds. □

If a set of variables is specified by using the multi-index $\vec{\boldsymbol{m}}_{\boldsymbol{i}}$, then the conditional contribution that arises from the previous term can be identified as follows

$$24 \sum_{\vec{\boldsymbol{m}}_{\boldsymbol{p}} \in \mathcal{P}_i} \sum_{\vec{\boldsymbol{m}}_{\boldsymbol{q}} \in \mathcal{P}_i, q>p} \sum_{\vec{\boldsymbol{m}}_{\boldsymbol{r}} \in \mathcal{P}_i, r>q} \sum_{\substack{t>r,\, \vec{\boldsymbol{m}}_{\boldsymbol{r}} \in \mathcal{P}_i \\ \vec{\boldsymbol{m}}_{\boldsymbol{i}} \subseteq \vec{\boldsymbol{m}}_{\boldsymbol{pq}} \boxplus \cap_{rt} \\ \vec{\boldsymbol{m}}_{\boldsymbol{i}} \subseteq \vec{\boldsymbol{m}}_{\boldsymbol{rt}} \boxplus \cap_{pq}}} \int_{\hat{\Xi}_i} f_{\vec{\boldsymbol{m}}_{\boldsymbol{p}}} f_{\vec{\boldsymbol{m}}_{\boldsymbol{q}}} f_{\vec{\boldsymbol{m}}_{\boldsymbol{r}}} f_{\vec{\boldsymbol{m}}_{\boldsymbol{t}}} \mathrm{d}\mathbb{P}_i. \tag{B.7}$$

Proof. The set of all the possible sub-sets of variables relative to a multi-index $\vec{\boldsymbol{m}}_{\boldsymbol{i}}$ is represented by $\mathcal{P}_i$, but not all the possible combinations of four elements selected from $\mathcal{P}_i$ are leading to the multi-index $\vec{\boldsymbol{m}}_{\boldsymbol{i}}$. Thus, the first condition is to require that $\vec{\boldsymbol{m}}_{\boldsymbol{pqrt}} = \vec{\boldsymbol{m}}_{\boldsymbol{i}}$. In analogy with the previous proof, the non-null elements need to satisfy. The two requirements $\vec{\boldsymbol{m}}_{\boldsymbol{pq}} \setminus \cap_{pq} \subseteq \vec{\boldsymbol{m}}_{\boldsymbol{rt}}$ and $\vec{\boldsymbol{m}}_{\boldsymbol{rt}} \setminus \cap_{rt} \subseteq \vec{\boldsymbol{m}}_{\boldsymbol{pq}}$. If the two latter requirements are manipulated as $\vec{\boldsymbol{m}}_{\boldsymbol{rt}} \boxplus \vec{\boldsymbol{m}}_{\boldsymbol{pq}} \setminus \cap_{pq} \subseteq \vec{\boldsymbol{m}}_{\boldsymbol{rt}} \boxplus \vec{\boldsymbol{m}}_{\boldsymbol{rt}}$ and $\vec{\boldsymbol{m}}_{\boldsymbol{pq}} \boxplus \vec{\boldsymbol{m}}_{\boldsymbol{rt}} \setminus \cap_{rt} \subseteq \vec{\boldsymbol{m}}_{\boldsymbol{pq}} \boxplus \vec{\boldsymbol{m}}_{\boldsymbol{pq}}$, the following conditions can be written

$$\begin{aligned} \vec{\boldsymbol{m}}_{\boldsymbol{i}} \setminus \cap_{pq} \subseteq \vec{\boldsymbol{m}}_{\boldsymbol{rt}} &\Rightarrow \vec{\boldsymbol{m}}_{\boldsymbol{i}} \subseteq \vec{\boldsymbol{m}}_{\boldsymbol{rt}} \boxplus \cap_{pq} \\ \vec{\boldsymbol{m}}_{\boldsymbol{i}} \setminus \cap_{rt} \subseteq \vec{\boldsymbol{m}}_{\boldsymbol{pq}} &\Rightarrow \vec{\boldsymbol{m}}_{\boldsymbol{i}} \subseteq \vec{\boldsymbol{m}}_{\boldsymbol{pq}} \boxplus \cap_{rt}. \end{aligned} \tag{B.8}$$

□

In conclusion, summing up all the contributions, the final form for the kurtosis can be written as

$$\begin{aligned} k = &\sum_{p=1}^{N} \int_{\hat{\Xi}_p} f^4_{\vec{\boldsymbol{m}}_{\boldsymbol{p}}} \mathrm{d}\mathbb{P}_p + 4 \sum_{\vec{\boldsymbol{m}}_{\boldsymbol{p}}} \sum_{\vec{\boldsymbol{m}}_{\boldsymbol{q}} \subset \vec{\boldsymbol{m}}_{\boldsymbol{p}}} \int_{\hat{\Xi}_{pq}} f^3_{\vec{\boldsymbol{m}}_{\boldsymbol{p}}} f_{\vec{\boldsymbol{m}}_{\boldsymbol{q}}} \mathrm{d}\mathbb{P}_{pq} + 6 \sum_{p=1}^{N} \sum_{q=p+1}^{N} \int_{\hat{\Xi}_{pq}} f^2_{\vec{\boldsymbol{m}}_{\boldsymbol{p}}} f^2_{\vec{\boldsymbol{m}}_{\boldsymbol{q}}} \mathrm{d}\mathbb{P}_{pq} \\ &+ 12 \sum_{p=1}^{N} \sum_{\substack{q=1 \\ q \neq p}}^{N} \sum_{r=q+1}^{N} \int_{\hat{\Xi}_{pqr}} f^2_{\vec{\boldsymbol{m}}_{\boldsymbol{p}}} f_{\vec{\boldsymbol{m}}_{\boldsymbol{q}}} f_{\vec{\boldsymbol{m}}_{\boldsymbol{r}}} \mathrm{d}\mathbb{P}_{pqr} + 24 \sum_{p=1}^{N} \sum_{q=p+1}^{N} \sum_{r=q+1}^{N} \sum_{t=r+1}^{N} \int_{\hat{\Xi}_{pqrt}} f_{\vec{\boldsymbol{m}}_{\boldsymbol{p}}} f_{\vec{\boldsymbol{m}}_{\boldsymbol{q}}} f_{\vec{\boldsymbol{m}}_{\boldsymbol{r}}} f_{\vec{\boldsymbol{m}}_{\boldsymbol{t}}} \mathrm{d}\mathbb{P}_{pqrt}, \end{aligned} \tag{B.9}$$

where each conditional contribution, with respect to a fixed multi-index $\vec{\boldsymbol{m}}_{\boldsymbol{i}}$, is equal to

$$\begin{aligned} k_{\vec{\boldsymbol{m}}_{\boldsymbol{i}}} = &\int_{\hat{\Xi}_i} f^4_{\vec{\boldsymbol{m}}_{\boldsymbol{i}}} \mathrm{d}\mathbb{P}_i + 4 \int_{\hat{\Xi}_i} f^3_{\vec{\boldsymbol{m}}_{\boldsymbol{i}}} \sum_{\vec{\boldsymbol{m}}_{\boldsymbol{q}} \in \mathcal{P}_{i,\neq}} f_{\vec{\boldsymbol{m}}_{\boldsymbol{q}}} \mathrm{d}\mathbb{P}_i + 6 \sum_{\vec{\boldsymbol{m}}_{\boldsymbol{p}} \in \mathcal{P}_i} \sum_{\substack{\vec{\boldsymbol{m}}_{\boldsymbol{p}} \neq \vec{\boldsymbol{m}}_{\boldsymbol{q}} \in \mathcal{P}_i \\ \vec{\boldsymbol{m}}_{\boldsymbol{pq}} = \vec{\boldsymbol{m}}_{\boldsymbol{i}}}} \int_{\hat{\Xi}_i} f^2_{\vec{\boldsymbol{m}}_{\boldsymbol{p}}} f^2_{\vec{\boldsymbol{m}}_{\boldsymbol{q}}} \mathrm{d}\mathbb{P}_i \\ &+ 12 \sum_{\vec{\boldsymbol{m}}_{\boldsymbol{p}}} \sum_{\vec{\boldsymbol{m}}_{\boldsymbol{p}} \neq \vec{\boldsymbol{m}}_{\boldsymbol{q}} \in \mathcal{P}_i} \sum_{\substack{\vec{\boldsymbol{m}}_{\boldsymbol{r}} \in \mathcal{P}_i, r>q \\ \vec{\boldsymbol{m}}_{\boldsymbol{p}} \boxplus \cap_{qr} = \vec{\boldsymbol{m}}_{\boldsymbol{i}}}} \int_{\hat{\Xi}_i} f^2_{\vec{\boldsymbol{m}}_{\boldsymbol{p}}} f_{\vec{\boldsymbol{m}}_{\boldsymbol{q}}} f_{\vec{\boldsymbol{m}}_{\boldsymbol{r}}} \mathrm{d}\mathbb{P}_i \\ &+ 24 \sum_{\vec{\boldsymbol{m}}_{\boldsymbol{p}} \in \mathcal{P}_i} \sum_{\vec{\boldsymbol{m}}_{\boldsymbol{q}} \in \mathcal{P}_i, q>p} \sum_{\vec{\boldsymbol{m}}_{\boldsymbol{r}} \in \mathcal{P}_i, r>q} \sum_{\substack{t>r,\, \vec{\boldsymbol{m}}_{\boldsymbol{r}} \in \mathcal{P}_i \\ \vec{\boldsymbol{m}}_{\boldsymbol{i}} \subseteq \vec{\boldsymbol{m}}_{\boldsymbol{pq}} \boxplus \cap_{rt} \\ \vec{\boldsymbol{m}}_{\boldsymbol{i}} \subseteq \vec{\boldsymbol{m}}_{\boldsymbol{rt}} \boxplus \cap_{pq}}} \int_{\hat{\Xi}_i} f_{\vec{\boldsymbol{m}}_{\boldsymbol{p}}} f_{\vec{\boldsymbol{m}}_{\boldsymbol{q}}} f_{\vec{\boldsymbol{m}}_{\boldsymbol{r}}} f_{\vec{\boldsymbol{m}}_{\boldsymbol{t}}} \mathrm{d}\mathbb{P}_i. \end{aligned} \tag{B.10}$$

Appendix C. Skewness from the PC expansion

In this section, the skewness is obtained from the PC series expansion. By applying the multinomial theorem, the skewness can be written as a sum of contributions generated by the interactions of the polynomial basis components

$$\begin{aligned} s &= \int_{\Xi^d} \left(f(\vec{\boldsymbol{\xi}}) - \beta_0\right)^3 \mathrm{d}\mathbb{P} = \int_{\Xi^d} \left(\sum_{p=1}^{P} \beta_p \Psi_p(\vec{\boldsymbol{\xi}})\right)^3 \mathrm{d}\mathbb{P} \\ &= \sum_{p=1}^{P} \beta_p^3 \langle \Psi_p^3(\vec{\boldsymbol{\xi}})\rangle + 3\sum_{p=1}^{P}\beta_p^2 \sum_{\substack{q=1\\q\neq p}}^{P} \beta_q \langle \Psi_p^2(\vec{\boldsymbol{\xi}}), \Psi_q(\vec{\boldsymbol{\xi}})\rangle + 6\sum_{p=1}^{P}\sum_{q=p+1}^{P}\sum_{r=q+1}^{P} \beta_p\beta_q\beta_r \langle \Psi_p(\vec{\boldsymbol{\xi}}), \Psi_q(\vec{\boldsymbol{\xi}})\Psi_r(\vec{\boldsymbol{\xi}})\rangle. \end{aligned} \tag{C.1}$$

At the first glance, it seems that no orthogonal contributions are present, because the interactions involve only polynomial forms raised to a power higher than one or *triadic* interaction. However, the second and third terms should be further investigated.

Following from the definition of each polynomial term (21), the product between two polynomial terms of the basis, where the first one is raised to the power n with $1 < n \in \mathbb{N}$, can be written as follows

$$\begin{aligned} \langle \Psi_p^n(\vec{\boldsymbol{\xi}}), \Psi_q(\vec{\boldsymbol{\xi}})\rangle &= \int_{\Xi^d} \left(\prod_{i=1}^{d} \psi^n_{\alpha_i^p}(\xi_i)\right)\left(\prod_{i=1}^{d} \psi_{\alpha_i^q}(\xi_i)\right) \mathrm{d}\mathbb{P} \\ &= \int_{\Xi^d} \left(\prod_{i=1}^{d} \psi^n_{\alpha_i^p}(\xi_i)\,\psi_{\alpha_i^q}(\xi_i)\right) \mathrm{d}\mathbb{P} = \prod_{i=1}^{d} \int_{\Xi_i} \psi^n_{\alpha_i^p}(\xi_i)\,\psi_{\alpha_i^q}(\xi_i)\,p(\xi_i)\mathrm{d}\xi_i. \end{aligned} \tag{C.2}$$

Due to the orthogonality of the PC basis with respect to $\Psi_0 = 1$, it follows that if $\alpha_i^p = 0$ then the integral with respect to the variable ξ_i becomes

$$\int_{\Xi_i} \psi_{\alpha_i^q}(\xi_i)\,p(\xi_i)\mathrm{d}\xi_i = 0 \quad \text{for} \quad \alpha_i^q \neq 0. \tag{C.3}$$

From this relation, the orthogonality of $\langle \Psi_p^n(\vec{\boldsymbol{\xi}}), \Psi_q(\vec{\boldsymbol{\xi}})\rangle$ follows. The non-null existence of the corresponding skewness (and kurtosis term) can be efficiently identified by means of the function Δ_q^p defined in §3.2.

The third term of the skewness from the PC series involves *triadic* interaction of polynomial terms raised to the power one

$$\langle \Psi_p(\vec{\boldsymbol{\xi}}), \Psi_q(\vec{\boldsymbol{\xi}})\Psi_r(\vec{\boldsymbol{\xi}})\rangle = \int_{\Xi^d} \left(\prod_{i=1}^{d} \psi_{\alpha_i^p}(\xi_i)\,\psi_{\alpha_i^q}(\xi_i)\,\psi_{\alpha_i^r}(\xi_i)\right) \mathrm{d}\mathbb{P} = \prod_{i=1}^{d} \int_{\Xi_i} \psi_{\alpha_i^p}(\xi_i)\,\psi_{\alpha_i^q}(\xi_i)\,\psi_{\alpha_i^r}(\xi_i)\,p(\xi_i)\mathrm{d}\xi_i. \tag{C.4}$$

This term can be analyzed after the inspection of the relative multi-indices $\vec{\boldsymbol{m}}^{\star,p}$, $\vec{\boldsymbol{m}}^{\star,q}$ and $\vec{\boldsymbol{m}}^{\star,r}$. If the sum of the respective components of the multi-indices, *i.e.* $m_i^{\star,p} + m_i^{\star,q} + m_i^{\star,r}$, is equal to zero, then the variable is not present and no information can be obtained (the previous integral would be equal to 1 in such a case). If the sum $m_i^{\star,p} + m_i^{\star,q} + m_i^{\star,r}$ is equal to 1, this means that the variable is present in only one polynomial term between ψ_p, ψ_q and ψ_r, while it should not be present in the others (the relative coefficient $\alpha_i = 0$). This leads to a null integral due to the orthogonality of the basis with respect to the probability density function. However, another possibility can be associated to a null integral: if the sum $m_i^{\star,p} + m_i^{\star,q} + m_i^{\star,r} = 2$, the orthogonality between two polynomial terms guarantees that the integral is zero. The previous results can be resumed in the function Δ_{pqr} introduced in Section 3.2.

Appendix D. Kurtosis from the PC expansion

In this section, as already shown for the skewness in Appendix C, the kurtosis structure relying on the PC series is described. By applying the multinomial theorem to the PC series expansion, the kurtosis is computed as follows

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

The first and third term are not null in general, while the second one has already been analyzed for the skewness (the presence of the third power instead of the second one is not relevant). In addition, the last two terms need a separate analysis. The first contains the interaction between three polynomial terms, where the first of them is raised to the second power. In the general case of $1 < n \in \mathbb{N}$, it holds that

$$\langle \Psi_p^n(\vec{\boldsymbol{\xi}}), \Psi_q(\vec{\boldsymbol{\xi}})\Psi_r(\vec{\boldsymbol{\xi}})\rangle = \int_{\Xi^d} \left(\prod_{i=1}^{d} \psi^n_{\alpha_i^p}(\xi_i)\, \psi_{\alpha_i^q}(\xi_i)\, \psi_{\alpha_i^r}(\xi_i) \right) \mathrm{d}\mathbb{P} = \prod_{i=1}^{d} \int_{\Xi_i} \psi^n_{\alpha_i^p}(\xi_i)\, \psi_{\alpha_i^q}(\xi_i)\, \psi_{\alpha_i^r}(\xi_i)\, p(\xi_i)\mathrm{d}\xi_i. \tag{D.2}$$

From the last equation, it is possible to note that the orthogonality between polynomial term can be advocated if the term raised to the $n-$th power is zero. If $\alpha_i^p = 0$ and the sum of the remaining terms $m_i^{\star,q} + m_i^{\star,r} \neq 0$, then a null integral exist irrespectively of the coefficients α_i^p and α_i^q

$$\int_{\Xi_i} \psi_{\alpha_i^q}(\xi_i)\, \psi_{\alpha_i^r}(\xi_i)\, p(\xi_i)\mathrm{d}\xi_i = 0. \tag{D.3}$$

Note that the function Δ_{qr}^p has been introduced in §3.3.

The last term of the kurtosis expansion involves the interaction of four polynomials terms. This case represents an extension of the term already analyzed for the skewness (see Appendix C) where the interaction between three polynomial terms has been discussed. By inspecting the sum of the coefficients of the multi-indices $\vec{\boldsymbol{m}}^{\star,p}$, $\vec{\boldsymbol{m}}^{\star,q}$, $\vec{\boldsymbol{m}}^{\star,r}$ and $\vec{\boldsymbol{m}}^{\star,t}$, it is possible to determine if the integral is equal to zero. In particular, if the sum relative to the ith coordinates, *i.e.* $m_i^{\star,p} + m_i^{\star,q} + m_i^{\star,r} + m_i^{\star,t}$, is equal to 1 or 2, the orthogonal properties with respect to the pdf holds (this is true irrespectively of the values of the α_i^k coefficients). This result has been used in Section 3.3 to define the function Δ_{pqrt}.

Appendix E. Computing the functions Δ_p^q, Δ_{pqr}, Δ_{qr}^p and Δ_{pqrt}

The third and fourth-order central moments can be efficiently computed by means of the PC expansion if the functions Δ_p^q, Δ_{pqr}, Δ_{qr}^p and Δ_{pqrt} are evaluated before computing the integrals. Hereinafter, we refer to these functions as selecting functions. The first step is obtaining, in the pre-processing stage, the multi-indices $\vec{\boldsymbol{\alpha}}^k \in \mathbb{N}^d$ for $k = 0, \ldots, P$. This task can be accomplished using the algorithm proposed in [48]. Evaluating the selecting functions requires the computation of the normalized multi-indices $\vec{\boldsymbol{m}}^{\star,k} = \vec{\boldsymbol{m}}^{\star,k}(\vec{\boldsymbol{\alpha}}^k)$. In the following we recalled the algorithm proposed in [48] in which we add the computation of $\vec{\boldsymbol{m}}^{\star,k}$:

1. Set $\alpha_i^0 = 0$, for $i = 1, \ldots, d$

   1.bis $\vec{\boldsymbol{m}}^{\star,0} = 0$

2. Set $\alpha_j^i = \delta_{ij}$, for $1 \leq i, j \leq d$

   2.bis $\vec{\boldsymbol{m}}^{\star,i} = \vec{\boldsymbol{\alpha}}^i$

3. Set $P = d$

4. Set $h_i(1) = 1$ for $i = 1, \ldots, d$

5. For $k = 2, \ldots, n_0$

   a. Set $L = P$

   b. Set $h_i(k) = \sum_{\ell=i}^{N} h_\ell(k-1)$

   c. For $j = 1, \ldots, d$:

      * For $\ell = L - h_j(k) + 1, \ldots, L$

        · Set $P = P + 1$

        · Set $\alpha_i^P = \alpha_i^\ell$ for $i = 1, \ldots, d$

            - Set $\alpha_j^P = \alpha_j^P + 1$
            - Set $m_j^{\star,P} = \alpha_j^P / \max\left(1, \alpha_j^P\right)$.

Let us start with the evaluation of the function Δ_q^p:

- Set $\Delta_q^p = 1$
- For $j = 1, \ldots, d$
    - If $\alpha_j^p = 0$
        - For $i = 1, \ldots, d$
            - If $m_j^{\star,q} = 1$ then $\Delta_q^p = 0$.

The selecting function Δ_{pqr} is evaluated as

- Set $\Delta_{qpr} = 1$
- For $j = 1, \ldots, d$
    - If $m_j^p + m_j^q + m_j^r = 1$ or $m_j^p + m_j^q + m_j^r = 2$ then $\Delta_{qpr} = 0$.

The remaining functions, Δ_{qr}^p and Δ_{pqrt}, are computed in analogy with the ones described above.