11th World Congress on Computational Mechanics (WCCM XI)
5th European Conference on Computational Mechanics (ECCM V)
6th European Conference on Computational Fluid Dynamics (ECFD VI)
E. Oñate, J. Oliver and A. Huerta (Eds)

ON THE USE OF HIGH-ORDER STATISTICS IN ROBUST DESIGN OPTIMIZATION

P.M. Congedo[1] and G. Geraci[2] and G. Iaccarino[2]

[1] INRIA Bordeaux Sud-Ouest, Team Bacchus, 200 avenue de la Vieille Tour, 33405 Talence, France. pietro.congedo@inria.fr, gianluca.geraci@inria.fr.
[2] Mechanical Engineering Dept, Center for Turbulence Research, Stanford University, Bld 500, CA 94305-3035, USA. jops@stanford.edu.

Key words: High-order statistics, Robust design optimization, CFD.

Abstract. ANOVA analysis is a very common numerical technique for computing a hierarchy of most important input parameters for a given output when variations are computed in terms of variance. This second central moment can not be retained as an universal criterion for ranking some variables, since a non-gaussian output could require higher order (more than second) statistics for a complete description and analysis. In this work, we illustrate how third and fourth-order statistic moments, i.e. skewness and kurtosis, respectively, can be decomposed. It is shown that this decomposition is correlated to a polynomial chaos expansion, permitting to easily compute each term. Then, new sensitivity indices are proposed basing on the computation of the kurtosis. Some test-cases are introduced showing the importance of high-order statistics in robust design optimization.

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. In many industry-relevant situations the performance metrics depend in a complex, non-linear fashion on those factors and the construction of an accurate representation of this relationship is difficult. Probabilistic uncertainty quantification (UQ) approaches represent the inputs as random variables and seek to construct a statistical characterization of few quantities of interest. Several methodologies are proposed to tackle this problem; polynomial chaos (PC) methods [1] can provide considerable speed-up in computational time when compared to MC. In realistic situations however, the presence of a large number of uncertain inputs leads to an exponential increase of the cost thus making these methodologies unfeasible [2]. This situation becomes even more challenging when robust design optimization is tackled [3]. Robust optimization processes may require a prohibitive computational cost when dealing with a large number of uncertainties and a highly non-linear fitness function. Efforts in the development of numerical method are directed mainly to reduce the number of deterministic evaluations necessary for solving the optimization problem and for the uncertainty quantification (UQ) of the performances of interest. The overall cost is typically the product of the cost of the two approaches because the stochastic analysis and the optimization strategy are completely decoupled. Decoupled approaches are simple but more expensive than necessary.

Several UQ methods have been developed with the objective of reducing the number of solution required to obtain a statistical characterization of the quantity of interest. An alternative solution is based on approaches attempting to identify the relative importance of the input uncertainties on the output. A well known methodology is based on a decomposition of the variance of the quantity of interest in contributions closely connected to each of the input uncertainties (main effects) or combined inputs [4].

In this work, we illustrate the impact of high-order statistics in robust design efficiency and global computational cost. The aim is to provide some useful indications for obtaining a good trade-off between the high-quality information given by high-order statistics and the feasibility of the whole optimization loop. Several test-case will be presented and analyzed. Moreover, an efficient multi-objective optimization method taking into account high-order statistic moments, such as the third and fourth-order statistic moments, i.e. skewness and kurtosis, respectively, will be proposed. These moments can be easily computed by means of a Polynomial Chaos (PC) method.

2 HIGH-ORDER STATISTICS DEFINITION

Let us consider a real function $f = f(\boldsymbol{\xi})$ with $\boldsymbol{\xi}$ a vector of random inputs $\boldsymbol{\xi} \in \Xi^d = \Xi_1 \times \cdots \times \Xi_n$ $(\Xi \subset \mathbb{R}^d)$ and $\boldsymbol{\xi} \in \Xi^d \longmapsto f(\boldsymbol{\xi}) \in L^2(\Xi^d, p(\boldsymbol{\xi}))$, where $p(\boldsymbol{\xi}) = \prod_{i=1}^{d} p(\xi_i)$ is the probability density function of $\boldsymbol{\xi}$.

We can define the central statistical moment of f of order n as

$$\mu^n(f) = \int_{\Xi^d} (f(\boldsymbol{\xi}) - E(f))^n p(\boldsymbol{\xi}) \mathrm{d}\xi, \tag{1}$$

where $E(f)$ indicates the expected value of f

$$E(f) = \int_{\Xi^d} f(\boldsymbol{\xi}) p(\boldsymbol{\xi}) \mathrm{d}\xi. \tag{2}$$

In the following, we indicate with σ^2, the variance (second-order moment), with s the skewness (third-order), and with k the kurtosis (fourth-order). We recall here some convenient formulas for skewness s and kurtosis k. They can be computed as follows

$$\begin{aligned} s &= E(f^3) - 3E(f^2)E(f) + 2E(f)^3 \\ s &= E(f^3) - 3\sigma^2 E(f) - E(f)^3, \end{aligned} \tag{3}$$

$$\begin{aligned} k &= E(f^4) - 4E(f^3)E(f) + 6E(f^2)E(f)^2 - 3E(f)^4 \\ k &= E(f^4) - 4sE(f) - 6\sigma^2 E(f)^2 - E(f)^4. \end{aligned} \tag{4}$$

These expressions are used for the functional decomposition shown in the following sections.

3 FUNCTIONAL DECOMPOSITION

Let us apply the definition of the Sobol functional decomposition to the function f as follows

$$f(\boldsymbol{\xi}) = \sum_{i=0}^{N} f_{\boldsymbol{m_i}}(\boldsymbol{\xi} \cdot \boldsymbol{m_i}), \tag{5}$$

where the multi-index $\boldsymbol{m}$, of cardinality $card(\boldsymbol{m}) = d$, can contain only elements equal to 0 or 1. Clearly, the total number of admissible multi-indices $\boldsymbol{m_i}$ is $N+1 = 2^d$; this number represent the total number of contributes up to the dth-order of the stochastic variables $\boldsymbol{\xi}$. The scalar product between the stochastic vector $\boldsymbol{\xi}$ and $\boldsymbol{m_i}$ is employed to identify the functional dependences of $f_{\boldsymbol{m_i}}$. In this framework, the multi-index $\boldsymbol{m}_0 = (0, \ldots, 0)$, is associated to the mean term $f_{\boldsymbol{m}_0} = \int_{\Xi^d} f(\boldsymbol{\xi}) p(\boldsymbol{\xi}) \mathrm{d}\xi$. As a consequence, $f_{\boldsymbol{m}_0}$ is equal to the expectancy of f, *i.e.* $E(f)$. Let us assume, in the following, to order the N multi-indices $\boldsymbol{m_i}$ in the following way:

$$\begin{aligned} \boldsymbol{m}_1 &= (1, 0, \ldots, 0) \\ \boldsymbol{m}_2 &= (0, 1, \ldots, 0) \\ &\vdots \\ \boldsymbol{m}_d &= (0, \ldots, 1) \\ \boldsymbol{m}_{d+1} &= (1, 1, 0, \ldots, 0) \\ \boldsymbol{m}_{d+2} &= (1, 0, 1, 0, \ldots, 0) \\ &\vdots \\ \boldsymbol{m}_N &= (1, \ldots, 1). \end{aligned} \tag{6}$$

Except the term $\boldsymbol{m}_0$, that should be the first in the series, the remaining N multi-indices $\boldsymbol{m_i}$ should be classified with respect to a prescribed criterion. However, this criterion does not affect in any way the successive ANOVA functional decomposition.

The decomposition (5) is of ANOVA-type in the sense of Sobol if all the members in Eq. (5) are orthogonal, *i.e.* as follows

$$\int_{\Xi^d} f_{\boldsymbol{m_i}}(\boldsymbol{\xi} \cdot \boldsymbol{m_i}) f_{\boldsymbol{m}_j}(\boldsymbol{\xi} \cdot \boldsymbol{m}_j) p(\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi} = 0 \quad \text{with} \quad \boldsymbol{m_i} \neq \boldsymbol{m}_j, \tag{7}$$

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

$$\int_{\Xi^d} f_{\boldsymbol{m_i}}(\boldsymbol{\xi}\cdot\boldsymbol{m_i})p(\xi_j)\mathrm{d}\xi_j = 0 \quad \text{with} \quad \xi_j \in (\boldsymbol{\xi}\cdot\boldsymbol{m_i})\,. \tag{8}$$

Each term $f_{\boldsymbol{m_i}}$ of (5) can be expressed as

$$f_{\boldsymbol{m_i}}(\boldsymbol{\xi}\cdot\boldsymbol{m_i}) = \int_{\Xi^{d-\mathrm{card}(\hat{\boldsymbol{m}}_i)}} f_{\boldsymbol{m_i}}(\boldsymbol{\xi}\cdot\boldsymbol{m_i})p(\bar{\boldsymbol{\xi}}_i)\bar{\boldsymbol{\xi}}_i \quad - \sum_{\substack{\boldsymbol{m}_j\neq\boldsymbol{m_i}\\ \mathrm{card}(\hat{\boldsymbol{m}}_j)<\mathrm{card}(\boldsymbol{m_i})}} f_{\boldsymbol{m}_j}(\boldsymbol{\xi}\cdot\boldsymbol{m}_j), \tag{9}$$

where the multi-indexes, $\hat{\boldsymbol{m}}_{\boldsymbol{i}}$, have cardinality equal to the number of non-null elements in $\boldsymbol{m_i}$ and $\bar{\boldsymbol{\xi}}_i$ contains all the variables not contained in $(\boldsymbol{\xi}\cdot\boldsymbol{m_i})$, *i.e.* $(\boldsymbol{\xi}\cdot\boldsymbol{m_i})\cup\bar{\boldsymbol{\xi}}_i=\boldsymbol{\xi}$.

The functional decomposition (5) is usually employed to compute the contribution of each term to the overall variance, as shown in the next section.

3.1 Variance decomposition

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

$$\sigma^2 = E(f^2) - E(f)^2. \tag{10}$$

As a consequence, the problem is to compute $E(f^2)$, seeing that $E(f)$ is known and equal to $f_{\boldsymbol{m}_0}$. Starting from Eq. (5), it is easy to compute

$$f^2(\boldsymbol{\xi}) = \sum_{i=0}^{N} f^2_{\boldsymbol{m_i}}(\boldsymbol{\xi}\cdot\boldsymbol{m_i}) + 2\sum_{i=0}^{N}\sum_{j=i+1}^{N} f_{\boldsymbol{m_i}}(\boldsymbol{\xi}\cdot\boldsymbol{m_i})f_{\boldsymbol{m}_j}(\boldsymbol{\xi}\cdot\boldsymbol{m}_j). \tag{11}$$

If the equation (11) is integrated in the stochastic space and the orthogonality property (7) is applied, variance can be decomposed as

$$\sigma^2 = \sum_{i=1}^{N}\int_{\Xi^d} f^2_{\boldsymbol{m_i}}(\boldsymbol{\xi}\cdot\boldsymbol{m_i})p(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi} = \sum_{i=1}^{N}\int_{\hat{\Xi}_i} f^2_{\boldsymbol{m_i}}(\boldsymbol{\xi}\cdot\boldsymbol{m_i})p(\boldsymbol{\xi}\cdot\boldsymbol{m_i})\mathrm{d}(\boldsymbol{\xi}\cdot\boldsymbol{m_i}), \tag{12}$$

where the symbol $\hat{\Xi}_i$ is employed to indicate $\Xi^{\mathrm{card}(\hat{\boldsymbol{m}}_i)}$ for brevity.

Variance can be expressed as the summation of all the conditional contributions

$$\sigma^2 = \sum_{i=1}^{N}\sigma^2_{\boldsymbol{m_i}}. \tag{13}$$

So, a comparison with the equation (12) shows that

$$\sigma^2_{\boldsymbol{m_i}} = \int_{\hat{\Xi}_i} f^2_{\boldsymbol{m_i}}(\boldsymbol{\xi}\cdot\boldsymbol{m_i})p(\boldsymbol{\xi}\cdot\boldsymbol{m_i})\mathrm{d}(\boldsymbol{\xi}\cdot\boldsymbol{m_i}). \tag{14}$$

Then, the same type of analysis is applied to skewness and kurtosis.

3.2 Skewness decomposition in conditional terms

In this section, the same procedure already presented in the previous section for the variance, is extended to the computation of the skewness. In this case, the major drawbacks is the presence of an higher number of terms to compute with respect to the variance case. In the case of the variance, due to the properties of the ANOVA terms, all the mixed contributions are zero due to orthogonality. This is not the case of the mixed contribution for the skewness. However some terms can be identified as orthogonal, as well as the case of the variance reducing the overall number of terms to compute.

The first step in order to obtain the skewness in terms of the ANOVA components of the function $f(\boldsymbol{\xi})$ is to raise the ANOVA functional decomposition of the function $f(\boldsymbol{\xi})$ to the third power by employing the multinomial theorem as follows

$$\begin{aligned} s = (f(\boldsymbol{\xi}) - f_0)^3 &= \left(\sum_{i=1}^{N} f_{\boldsymbol{m_i}}(\boldsymbol{\xi} \cdot \boldsymbol{m_i}) \right)^3 \\ &= \sum_{i=1}^{N} \int_{\hat{\Xi}_i} f^3_{\boldsymbol{m_i}}(\boldsymbol{\xi} \cdot \boldsymbol{m_i}) \mathrm{d}\mu_i + 3 \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \int_{\hat{\Xi}_{ij}} f^2_{\boldsymbol{m_i}}(\boldsymbol{\xi} \cdot \boldsymbol{m_i}) f_{\boldsymbol{m_j}}(\boldsymbol{\xi} \cdot \boldsymbol{m_j}) \mathrm{d}\mu_{ij} \\ &+ 6 \sum_{i=1}^{N} \sum_{j=i+1}^{N} \sum_{k=j+1}^{N} \int_{\hat{\Xi}_{ijk}} f_{\boldsymbol{m_i}}(\boldsymbol{\xi} \cdot \boldsymbol{m_i}) f_{\boldsymbol{m_j}}(\boldsymbol{\xi} \cdot \boldsymbol{m_j}) f_{\boldsymbol{m}_k}(\boldsymbol{\xi} \cdot \boldsymbol{m}_k) \mathrm{d}\mu_{ijk}, \end{aligned} \tag{15}$$

where $\hat{\Xi}_{ij} = \Xi^{\mathrm{card}(\hat{\boldsymbol{m_{ij}}})}$ and $\hat{\Xi}_{ijk} = \Xi^{\mathrm{card}(\hat{\boldsymbol{m_{ijk}}})}$. In the following, the notation is simplified by omitting the explicit dependence of the function $f_{\boldsymbol{m_i}}$ with respect to its coordinates, *i.e.* $f_{\boldsymbol{m_i}} = f_{\boldsymbol{m_i}}(\boldsymbol{\xi} \cdot \boldsymbol{m_i})$.

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

$$\boldsymbol{m_{ab\cdots z}} = \boldsymbol{m_a} \boxplus \boldsymbol{m_b} \boxplus \cdots \boxplus \boldsymbol{m_z} = \left(\frac{m_{a_1} + m_{b_1} + \cdots m_{z_1}}{\Big|\Big| m_{a_1} + m_{b_1} + \cdots + m_{z_1} \Big|\Big|_{\neq 0}}, \ldots, \frac{m_{a_d} + m_{b_d} + \cdots + m_{z_d}}{\Big|\Big| m_{a_d} + m_{b_d} + \cdots + m_{z_d} \Big|\Big|_{\neq 0}} \right), \tag{16}$$

where the norm $\Big|\Big| \cdot \Big|\Big|_{\neq 0}$ is defined as

$$\Big|\Big| \alpha \Big|\Big|_{\neq 0} = \begin{cases} |\alpha| & \text{if} \quad \alpha \neq 0 \\ 1 & \text{if} \quad \alpha = 0. \end{cases} \tag{17}$$

The expression presented in (15) includes some terms always equal to zero due to the orthogonality of the ANOVA functional components. In particular, a more compact final expression can be obtained as:

$$s = \sum_{p=1}^{N} \int_{\hat{\Xi}_p} f^3_{\boldsymbol{m_p}} \mathrm{d}\mu_p + 3 \sum_{\boldsymbol{m_p}} \sum_{\boldsymbol{m_q} \subset \boldsymbol{m_p}} \int_{\hat{\Xi}_{pq}} f^2_{\boldsymbol{m_p}} f_{\boldsymbol{m_q}} \mathrm{d}\mu_{pq} + 6 \sum_{p=1}^{N} \sum_{q=p+1}^{N} \sum_{\substack{r=q+1 \\ \boldsymbol{m_{pq}}=\boldsymbol{m_r}}}^{N} \int_{\hat{\Xi}_{pq}} f_{\boldsymbol{m_p}} f_{\boldsymbol{m_q}} f_{\boldsymbol{m_r}} \mathrm{d}\mu_{pq}. \tag{18}$$

One of the most important contribution of this kind of approach 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 (14). 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 contribution. For obtaining an additional form of the kind

$$s = \sum_{i=1}^{N} s_{\boldsymbol{m_i}}, \tag{19}$$

it is mandatory to identify all the set of indexes whose interactions become part of an assigned multi-index $\boldsymbol{m_i}$.

Considering that to each multi-index $\boldsymbol{m_i}$ is associated a set of $2^{|\boldsymbol{m_i}|}-1$ sub-interactions and denoting this set as $\mathcal{P}_i$ (for instance if $\boldsymbol{m_i} = (1,1)$ then the set $\mathcal{P}_i = \{(1,0),(0,1),(1,1)\}$) holds, from the equation (18) it is possible to identify each contribution as follows

$$s_{\boldsymbol{m_i}} = \int_{\hat{\Xi}_i} f^3_{\boldsymbol{m_i}} \mathrm{d}\mu_i + 3 \int_{\hat{\Xi}_i} f^2_{\boldsymbol{m_i}} \sum_{\boldsymbol{m_q} \in \mathcal{P}_{i,\neq}} f_{\boldsymbol{m_q}} \mathrm{d}\mu_i + 6 \sum_{\boldsymbol{m_p} \in \mathcal{P}_{i,\neq}} \sum_{\substack{\boldsymbol{m_p} \neq \boldsymbol{m_q} \in \mathcal{P}_{i,\neq} \\ \boldsymbol{m_{pq}}=\boldsymbol{m_i}}} \int_{\hat{\Xi}_i} f_{\boldsymbol{m_i}} f_{\boldsymbol{m_p}} f_{\boldsymbol{m_q}} \mathrm{d}\mu_i. \tag{20}$$

3.2.1 Kurtosis decomposition in conditional term

In this section, how decomposing the kurtosis is described. The functional decomposition based on the functional Sobol form (Eq. (5)), after the application of the multinomial theorem, is equal to

$$k = (f(\boldsymbol{\xi}) - f_0)^4 = \left(\sum_{i=1}^{N} f_{\boldsymbol{m_i}}(\boldsymbol{\xi} \cdot \boldsymbol{m_i}) \right)^4 \tag{21}$$

As already made for the skewness, the previous expression includes some terms always equal to zero thanks to the orthogonality properties of the ANOVA contributions.

Let us provide the relations to identify the conditional contribution related to a variable or a set of variables. In particular, if a specific multi-index $\boldsymbol{m_i}$ is provided, then the conditional expression for the kurtosis $k_{\boldsymbol{m_i}}$ is equal to

$$
\begin{aligned}
k_{\boldsymbol{m_i}} = & \int_{\hat{\Xi}_i} f_{\boldsymbol{m_i}}^4 \mathrm{d}\mu_i + 4 \int_{\hat{\Xi}_i} f_{\boldsymbol{m_i}}^3 \sum_{\boldsymbol{m_q} \in \mathcal{P}_{i,\neq}} f_{\boldsymbol{m_q}} \mathrm{d}\mu_i + 6 \sum_{\boldsymbol{m_p} \in \mathcal{P}_i} \sum_{\substack{\boldsymbol{m_p} \neq \boldsymbol{m_q} \in \mathcal{P}_i \\ \boldsymbol{m_{pq}} = \boldsymbol{m_i}}} \int_{\hat{\Xi}_i} f_{\boldsymbol{m_p}}^2 f_{\boldsymbol{m_q}}^2 \mathrm{d}\mu_i \\
& + 12 \sum_{\boldsymbol{m_p}} \sum_{\boldsymbol{m_p} \neq \boldsymbol{m_q} \in \mathcal{P}_i} \sum_{\substack{\boldsymbol{m_r} \in \mathcal{P}_i, r>q \\ \boldsymbol{m_p} \boxplus \cap_{qr} = \boldsymbol{m_i}}} \int_{\hat{\Xi}_i} f_{\boldsymbol{m_p}}^2 f_{\boldsymbol{m_q}} f_{\boldsymbol{m_r}} \mathrm{d}\mu_i \\
& + 24 \sum_{\boldsymbol{m_p} \in \mathcal{P}_i} \sum_{\boldsymbol{m_q} \in \mathcal{P}_i, q>p} \sum_{\boldsymbol{m_r} \in \mathcal{P}_i, r>q} \sum_{\substack{t>r, \boldsymbol{m_r} \in \mathcal{P}_i \\ \boldsymbol{m_i} \subseteq \boldsymbol{m_{pq}} \boxplus \cap_{rt} \\ \boldsymbol{m_i} \subseteq \boldsymbol{m_{rt}} \boxplus \cap_{pq}}} \int_{\hat{\Xi}_i} f_{\boldsymbol{m_p}} f_{\boldsymbol{m_q}} f_{\boldsymbol{m_r}} f_{\boldsymbol{m_t}} \mathrm{d}\mu_i
\end{aligned}
\tag{22}
$$

4 CORRELATION WITH POLYNOMIAL CHAOS FRAMEWORK

This section is devoted to illustrate how variance, skewness and kurtosis from the functional decomposition are correlated with the polynomial chaos framework. If a polynomial chaos formulation is used, an approximation $\tilde{f}$ of the function f is provided

$$f(\boldsymbol{\xi}) \approx \tilde{f}(\boldsymbol{\xi}) = \sum_{k=0}^{P} \beta_k \Psi_k(\boldsymbol{\xi}), \tag{23}$$

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

$$P + 1 = \frac{(n_0 + d)!}{n_0! d!}. \tag{24}$$

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

$$\Psi_k(\boldsymbol{\xi} \cdot \boldsymbol{m}^{\star,k}) = \prod_{i=1}^{d} \psi_{\alpha_i^k}(\xi_i) \tag{25}$$

where the multi index $\boldsymbol{m}^{\star,k} = \boldsymbol{m}^{\star,k}(\boldsymbol{\alpha}^k) \in \mathbb{N}^d$ is a function of $\boldsymbol{\alpha}^k$: $\boldsymbol{m}^{\star,k} = (m_1^{\star,k}, \dots, m_d^{\star,k})$, with $m_i^{\star,k} = \alpha_i^k / \left|\left|\alpha_i^k\right|\right|_{\neq 0}$.

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

$$\int_{\Xi} \Psi_i(\boldsymbol{\xi}) \Psi_k(\boldsymbol{\xi}) p(\boldsymbol{\xi}) \mathrm{d}\xi = \delta_{ij} \langle \Psi_i(\boldsymbol{\xi}), \Psi_i(\boldsymbol{\xi}) \rangle \tag{26}$$

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

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

$$\beta_k = \frac{\langle f(\boldsymbol{\xi}), \Psi_k(\boldsymbol{\xi}) \rangle}{\langle \Psi_k(\boldsymbol{\xi}), \Psi_k(\boldsymbol{\xi}) \rangle}, \quad \forall k. \tag{27}$$

5 INTRODUCING MORE SENSITIVITY INDICES

As introduced by Sobol, sensitivity indexes for variance can be computed for each conditional contribution following Eq. (13):

$$\sigma^{2,\mathrm{SI}}_{\boldsymbol{m_i}} = \frac{\sigma^2_{\boldsymbol{m_i}}}{\sigma^2}. \tag{28}$$

Here, we introduce new sensitivity indexes, basing on the decomposition of kurtosis and using the definition of the conditional term, as follows

$$k^{\mathrm{SI}}_{\boldsymbol{m_i}} = \frac{k_{\boldsymbol{m_i}}}{k}. \tag{29}$$

If a total sensitivity index is needed, *i.e.* it is necessary to compute the overall influence of a variable, it can be computed summing up all the contributions in which the variable is present

$$\begin{aligned} \mathrm{TSI}_j &= \sum_{\xi_j \in (\boldsymbol{\xi} \cdot \boldsymbol{m_i})} \sigma^{2,\mathrm{SI}}_{\boldsymbol{m_i}} \\ \mathrm{TSI}^k_j &= \sum_{\xi_j \in (\boldsymbol{\xi} \cdot \boldsymbol{m_i})} k^{\mathrm{SI}}_{\boldsymbol{m_i}}. \end{aligned} \tag{30}$$

Remark that some indexes could be introduced also for the skewness, but in this case the positivity of each term is not guaranteed.

6 SOME NUMERICAL RESULTS

6.1 Importance of Skewness in decomposition

This paragraph is devoted to show how important is to control the skewness during an optimization process. Let us consider the following polynomial function :

$$f = a\,(xz + xy) + b\left(x^2 + z^2\right) + (cba)\,y^2 \tag{31}$$

where x, y and z vary between 0 and 1 with an uniform pdf. Parameters a, b and c are design parameters that vary between -5 and 5. For this function, it is possible to compute analytically high-order statistics, as functions of the design parameters. Basing on Eqs. previously developed, formulas for mean, variance and skewness can be computed exactly

$$\mu = \frac{1}{3}cba + \frac{1}{2}a + \frac{2}{3}b, \tag{32}$$

$$\sigma^2 = \frac{8}{45}b^2 + \frac{1}{4}ab + \frac{5}{36}a^2 + \frac{1}{12}cba^2 + \frac{4}{45}c^2b^2a^2, \tag{33}$$

$$\begin{aligned} s = \frac{11}{120}ab^2 + \frac{13}{120}a^2b + \frac{32}{945}b^3 + \frac{1}{24}a^3 + \frac{1}{24}cb^2a^2 \\ + \frac{17}{360}a^3cb + \frac{1}{60}c^2b^2a^3 + \frac{16}{945}c^3b^3a^3. \end{aligned} \tag{34}$$

The skewness can be decomposed in conditional contributions, as follows

$$\begin{aligned} s_x &= \frac{1}{15}\left(\frac{1}{2} + \frac{1}{2}b\right)^2 b + \frac{1}{3780}b^3 \\ s_y &= \frac{1}{15}\left(\frac{1}{2}cba + \frac{1}{4}a\right)^2 cba + \frac{1}{3780}c^3b^3a^3 \\ s_z &= \frac{1}{15}\left(\frac{1}{4}a + \frac{1}{2}b\right)^2 b + \frac{1}{3780}b^3 \\ s_{xy} &= \frac{1}{45}a^2b + \frac{31}{720}a^3cb + \frac{1}{48}a^3 + \frac{1}{24}cb^2a^2 \\ s_{xz} &= \frac{1}{360}a^2b + \left(\frac{1}{6}\left(\frac{1}{2}a + \frac{1}{2}b\right)\right)\left(\frac{1}{4}a + \frac{1}{2}b\right)a \\ s_{yz} &= 0 \\ s_{xyz} &= 0 \end{aligned} \tag{35}$$

In order to show the importance to take into account also the high-order statistics in the robust optimization, different types of optimization are performed using several objective functions.

First, a classical bi-objective optimization is performed, where the mean of the function (Eq. (32)) is maximized and its variance (Eq. (33)) minimized. The Pareto front is reported in Figure 1. No measures of skewness have been used during the optimization process, then the Pareto front is constituted by various designs displaying a very large variation of skewness.

Now, let us consider a three-objectives optimization, *i.e.* consisting in the maximization of the mean (Eq. (32)), the minimization of the variance (Eq. (33)) and the minimization of the absolute value of the conditional skewness s_{xy} (Eq. (35)). In this case, the Pareto front is no more constituted by a curve, but by a surface in a 3D plan. The Pareto front

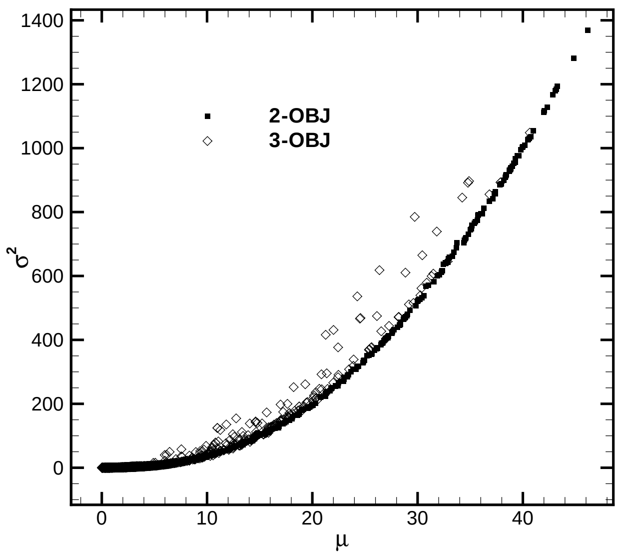

Figure 1: Pareto front in the plan $\mu - \sigma^2$ for the bi-objectives and three-objectives problem.

is represented by means of 2D representation in the Figures 1 and 2 with projections on the plans μ-σ^2, μ-s_{xy} and σ^2-s_{xy}, respectively. As shown in Figure 2, designs belonging to the Pareto front display a large variation of the conditional skewness.

Now, let us compare the results obtained with both optimizations. In Figure 1, we show Pareto fronts in the plan $\mu - \sigma^2$. Designs obtained with the three-objectives optimization are dominated (with respect to only μ and σ^2) by the designs coming from the bi-objectives optimization. This is reasonable seeing that designs from bi-objective optimization are not influenced by the skewness s_{xy} during the optimization.

In Figure 3, curves associated to the three-objectives optimization are obtained by the 3D Pareto front regarding only the designs having a skewness lower than 0.0001. Remark that individuals of this Pareto front take values of μ lower than 3.2 and values of σ^2 lower than 4.4. Moreover, they could be dominated in terms of μ and σ^2 by some individuals of the Pareto front obtained from the bi-objectives optimization. Here, the interest is to get a Pareto front that is not sensitive to large variation in the skewness, since designs obtained from bi-objective optimization could present large skewness values. This displays the great interest to estimate high-order statistics during optimization.

7 CONCLUSIONS

This paper deals with the decomposition of high-order statistics, *i.e.* such as skewness and kurtosis. The main contributions are the following:

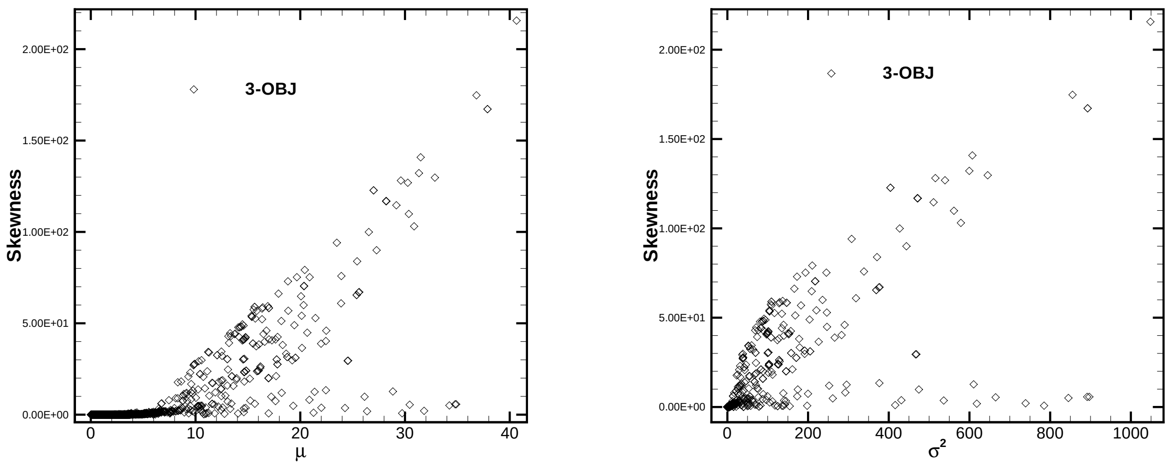

Figure 2: Pareto front in the plan $\mu - s_{xy}$ (on the left) and $\sigma^2 - s_{xy}$ (on the right) for the three-objectives optimization.

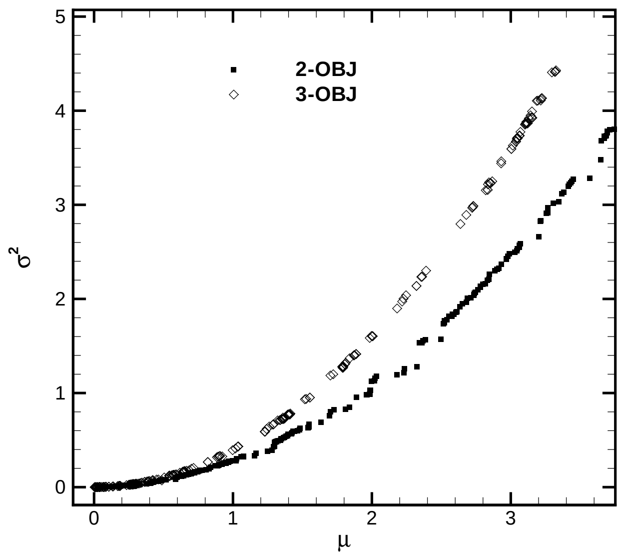

Figure 3: Pareto front in the plan $\mu - \sigma^2$ for the three-objectives optimization (extracted by the complete one considering only skewness inferior to 0.0001) and the bi-objective optimization.

- A correlation was found between the functional decomposition, as depicted by Sobol, and the polynomial chaos development. This permitted to identify clearly each term of the decomposition, drawing also a practical way to compute all these terms.

- Computing skewness was shown to be of great importance for an exhaustive and complete stochastic analysis.
- Sensitivity indices bases on kurtosis decomposition were introduced. The importance of ranking the predominant uncertainties in terms not only of the variance but also of higher order moments (then extending the ANOVA analysis also to higher order statistic moments), was demonstrated with an algebraic function, where all the decomposition terms can be calculated analytically.

Future works will be oriented towards adaptive strategies for the reduction of the global computational cost.

REFERENCES

[1] Xiu, Dongbin and Karniadakis, George Em. The Wiener–Askey Polynomial Chaos for Stochastic Differential Equations. *SIAM Journal on Scientific Computing*, Vol. **24**, 619–644, 2002.

[2] Foo, Jasmine and Karniadakis, George Em. Multi-element probabilistic collocation method in high dimensions. *Journal of Computational Physics*, Vol. **229**, 1536–1557, 2010.

[3] M. S. Eldred, C. G. Webster, P. G. Constantine. Design Under Uncertainty Employing Stochastic Expansion Methods. *AIAA Paper 20086001*, 2008.

[4] Congedo, P.M. and Geraci, G. and Abgrall, R. and Pediroda, V. and Parussini, L.. TSI metamodels-based multi-objective robust optimization. *Engineering Computations*, Vol. **30-8**, 1032–1053, 2013.