INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN, ICED'07
28 - 31 AUGUST 2007, CITÉ DES SCIENCES ET DE L'INDUSTRIE, PARIS, FRANCE

COMPARATIVE ANALYSIS OF UNCERTAINTY PROPAGATION METHODS FOR ROBUST ENGINEERING DESIGN

Mattia Padulo[1], Michele Sergio Campobasso[2] and Marin Dimitrov Guenov[1]
[1]School of Engineering, Cranfield University
[2]Department of Aerospace Engineering, Glasgow University

ABSTRACT
One of the main stages of robust Engineering Design is the propagation of uncertainty in the computer analysis system. This paper presents a comparative study of various methods for uncertainty propagation, including some recently proposed derivative-free techniques called Sigma-point methods. The latter are assessed against more conventional techniques such as Monte Carlo simulation and moment propagation based on truncated Taylor series. The accuracy and computational efficiency of all methods is investigated by means of mathematical analyses and numerical examples. The solution of a robust optimization problem based on the presented methods is included. Theoretical analyses and numerical results highlight that the accuracy of robust optimization using Sigma-point methods can be enhanced with respect to that using moment propagation based on linearization without altering significantly the computational cost of a single optimization step. The paper also discusses the optimal deployment of Automatic Differentiation in conjunction with the uncertainty analysis technique chosen for the solution of the robust optimization problem and highlights the computational benefits that this yields to.

Keywords: Robust Design, Uncertainty Propagation, Automatic Differentiation, Sigma-point methods

1 INTRODUCTION
Modern Engineering Design of complex products such as aircraft and aircraft propulsive systems is increasingly based on sophisticated multidisciplinary analysis and optimization tools. One difficulty associated with this design task is how to handle a variety of non-deterministic factors associated with the modelling stage (e.g. uncertainty of physical and numerical modelling), the manufacturing process (e.g. manufacturing and assembly tolerances) and the operation environment (e.g. weather variability). Robust design optimization (RDO) aims at designing products which not only have an optimal performance for given design specifications and constraints, but also exhibit minimal sensitivity with respect to probabilistically varying factors such as those mentioned above. The original methodology is due to Taguchi, and was based on direct experimentation [20]. It was then extended to simulation-based design, and gradually improved by using nonlinear constrained optimization techniques [18]. The constrained RDO strategy is made up of three main parts. The first stage consists of identifying, qualifying and quantifying the sources of uncertainty associated with the design input and the analysis modules. This is usually done by means of stochastic models. The second phase consists of propagating the uncertainty through the analysis system, to obtain a probabilistic description of the objective functions and constraints. The probabilistic state is often defined by suitable robust forms of objectives and constraints depending on expectation and variance of their deterministic counterparts. Finally, the third stage consists of optimizing the robust objectives subject to the robust constraints. The robust optimal design is such that both the expectation of the objectives is optimized and their variances are minimized.
This paper focuses on the second stage of RDO, which is critical for both the accuracy and the efficiency of the methodology. Several approaches to uncertainty analysis are available, including Monte Carlo methods (MCS), method of moments based on truncated Taylor series (MM) and Gaussian quadrature (GQ). Among these techniques, uncertainty analysis based on first-order derivatives has been often used in RDO due to its simplicity and computational speed. The accuracy of this method, however, is not satisfactory for highly non-linear functions or widely spread input

variables, and the method cannot handle discontinuous problems or discrete variables. In this paper, we compare the aforementioned methods with two advanced techniques, the Unscented Transform [15] and the Divided Difference Filter [17]. These methods belong to a class called Sigma-point methods (SP) and have been proposed in Control Theory in the past few years. One of the objectives of this work is to demonstrate their suitability for robust Engineering Design. As discussed in later sections, the computational performance of the SP methods in robust gradient-based optimization can be greatly enhanced by the use of Automatic Differentiation (AD) [1, 24]. AD tools can be used to obtain derivatives of functions defined by arbitrarily complex computer programs, and such derivatives are exact with respect to the code-defined functions from which they are obtained. This means that, unlike Finite Difference estimates, they have no truncation error and are thus accurate to within machine accuracy. AD software is usually based on either *source transformation* or *operator overloading* [2]. In both cases, the source code of the analysis program is required. The code differentiation can be done in *forward* or *reverse* (also known as *adjoint*) mode. The latter route is computationally much faster than the former when the number of input variables is much larger than that of the output variables. This is because the gradient of each output with respect to all input variables can be obtained at a cost comparable to that of a single analysis when using the adjoint mode [1, 2]. AD software is currently available for all common programming languages such as FORTRAN and C/C^{++}. The numerical examples reported in this paper are implemented in MATLAB, and we have chosen to use MAD [9, 10] for AD.
The paper is structured as follows: Section 2 introduces the problem of robust optimization and uncertainty propagation; the uncertainty analysis based on MCS, GQ and MM is summarized in Section 3, whereas the SP methods are presented in Section 4. This section also reports an error analysis based on a comparison with the estimates of the Taylor series. In Section 5, we use the aforementioned methods to perform uncertainty analysis of selected test functions and compare the accuracy and computational performance of the methods. A robust optimization problem is also considered. The conclusions of this study are finally reported in Section 6.

2 ROBUST OPTIMIZATION AND UNCERTAINTY PROPAGATION

Engineering design is commonly performed with the aid of computer programs, which model the different disciplines characterizing the behaviour of the system of interest. Let us consider a scalar function $y = f(\boldsymbol{x})$, with $\boldsymbol{x}$ being an array of n design variables. The deterministic design optimization problem consists of finding the vector $\boldsymbol{x}$ which minimizes y subject to r constraints $g_i(\boldsymbol{x}) \le 0$. In RDO both the objective function and the constraints become probabilistic functions of the stochastic variable $\boldsymbol{x}$, which here is assumed to have a known probability distribution. The robust attribute of the objective function can be achieved by reducing its sensitivity to the random fluctuation of the design variables $\boldsymbol{x}$, that is, minimizing the variance σ_y^2 while optimizing its expectation $\overline{y}$. The robustness of the constraints is instead guaranteed by accounting for their probabilistic satisfaction within the range of variation of the design variables. The robust optimization problem can be formally stated as follows:

Find $\boldsymbol{x} \in R^n$ to minimize $F(\overline{y}(\boldsymbol{x}), \sigma_y(\boldsymbol{x}))$,

subject to $g_{c_i}(\overline{g}_i(\boldsymbol{x}), \sigma_{gi}(\boldsymbol{x})) \le 0 \quad i = 1, ..., r$,

$\boldsymbol{x}_L \le \boldsymbol{x} \le \boldsymbol{x}_U$.

The target of uncertainty analysis is to obtain mean and variance of y starting from the knowledge of the uncertainty affecting $\boldsymbol{x}$. If all the variables are continuous, the first two moments are:

$$\overline{y} = E[f(\boldsymbol{x})] = \int_{-\infty}^{+\infty} f(\boldsymbol{x}) p_X(\boldsymbol{x}) d\boldsymbol{x}, \qquad (1)$$

$$\sigma_y^2 = \int_{-\infty}^{+\infty} \left\{ f(\boldsymbol{x}) - E[f(\boldsymbol{x})] \right\}^2 p_X(\boldsymbol{x}) d\boldsymbol{x}. \qquad (2)$$

Here p_X is the joint probability function corresponding to the mentioned distributions. Since a closed-form solution of these integrals can be obtained only in a few cases of practical interest, uncertainty

propagation is usually performed in an approximate fashion. Existing approaches include Monte Carlo methods [13, 14, 26], Taylor-based method of moments [7, 11, 12, 19], metamodels [22], polynomial chaos expansion [27] and reliability-based techniques [28].
The next section concerns the first two methods along with GQ. These three approaches provide the theoretical background required for discussing certain features of the SP propagation technique.

3 BACKGROUND

3.1 Monte Carlo Methods

Simulation methods like Monte Carlo simulation have been intensively used in uncertainty analysis since their introduction in the 1940s. The underlying principle is frequentist: probability distribution of the output of a process induced by the probability distribution of stochastic inputs is obtained by performing *m* repetitions of the process. Each time the sampling point ξ_i of the input space is drawn according to the known (or assumed) distribution of the inputs. When using simple random sampling, unbiased estimators of the integrals in Eq. (1) and Eq. (2) are given by:

$$\overline{y}_{MCS} = \frac{1}{m}\sum_{i=1}^{m} y_i \,, \qquad (3)$$

$$\sigma^2_{y_{MCS}} = \frac{1}{m-1}\sum_{i=1}^{m} (y_i - \overline{y}_{MCS})^2 \,, \qquad (4)$$

where $y_i = f(\xi_i)$. The standard deviation of the mean estimate from its exact value is $O(m^{-1/2})$, and is independent of the number of inputs *n*. One major disadvantage of MCS with respect to other uncertainty analysis methods is that the number of samples required for a sufficiently accurate estimate of the mean and the variance can be very large, thus requiring a large computational effort. For this reason, several other sampling techniques have been developed to reduce the number of required samples with respect to that of random sampling. The class of alternative methods include descriptive sampling [26], stratified and Latin Hypercube sampling [14]. These techniques reduce the number of samples required for a given accuracy, but the number of analyses still remains too high for practical use, particularly in the case of computationally demanding analysis codes. Another downside of any Monte Carlo technique to be used for RDO is the repeatability of the prediction [4]. For a given input distribution, the mean and variance obtained with different simulations may differ significantly even when using the same number of samples, and this spread decreases with *m*. In order to avoid insufficiently accurate derivatives when using MCS in gradient-based robust optimization, one may then have to increase *m* above the threshold required for the desired accuracy of mean and variance. This leads in turn to a very high computational effort. Figure 1 shows the mean and variance estimates obtained trough random sampling of the function $y = \sin(x)$. The variable *x* is normally distributed with mean $\overline{x} = \pi/4$ and unit variance. Note that the number of samples needed for a sufficient reduction of the mean and variance spread is $O(10^6)$. In the framework of a gradient-based RDO, this number of analyses ought to be performed at each step of the optimization.

3.2 Gaussian Quadrature

Quadrature formulas give the approximation of the integral of a function $f(\boldsymbol{x})$ on a domain $D \subset R^n$ by a properly weighted sum of particular functional values $f(x_i)$, where the points x_i's are suitably selected in *D*, and are also called nodes. When *f* depends on a single scalar input, GQ formulas are known to perform better than many other quadrature formulas [3]. Their straightforward extension to the multivariate integration is a product rule, which consists of applying the one-dimensional formulas to each of the *n* dimensions of *D*. Let us assume that $D = D_1 \times D_2 \times ... \times D_n$, where $\times$ denotes the Cartesian product, and $D_i \subset R$, for $i = 1...n$. Applying the one-dimensional integration rule with *N* nodes and given weights W_i to each D_i, the integrals of Eqs. (1) and (2) can be approximated as follows:

$$\overline{y}_{GQ1} = \sum_{i_1=1}^{N} W_{i_1} \left(\sum_{i_2=1}^{N} W_{i_2} \left(\ldots \sum_{i_n=1}^{N} W_{i_n} f(\boldsymbol{x}_{i_1,i_1,\ldots,i_n}) \right) \right), \tag{5}$$

$$\sigma^2_{y\,GQ1} = \sum_{i_1=1}^{N} W_{i_1} \left(\sum_{i_2=1}^{N} W_{i_2} \left(\ldots \sum_{i_n=1}^{N} W_{i_n} \left[f(\boldsymbol{x}_{i_1,i_1,\ldots,i_n}) - \overline{y}_{GQ1} \right]^2 \right) \right). \tag{6}$$

Three-node formulas (N=3) are usually adopted to obtain a sufficient accuracy. The arrays of design variables in Eqs. (6) and (7) are distributed at the vertices of an hypergrid of dimension 3^n. The vector associated with $i_1 = i_2 = ... = i_n = 2$ corresponds to the mean of the input variables $\overline{\boldsymbol{x}}$. The states with one or more subscript i_p equal to 1 have the corresponding components perturbed by $h^- \sigma_{i_p}$ with respect to their mean, where σ_{i_p} is the standard deviation of x_{i_p} and h^- is a suitable scalar. The states with one or more subscript i_p equal to 3, on the other hand, have the corresponding components perturbed by $h^+ \sigma_{i_p}$ with respect to their mean, where h^+ is again a suitable scalar. The three weights W_i and the values of h^+ and h^- have now to be determined. For the case of independent Gaussian variables, Taguchi [25] proposed a solution based on equal weights $W_i = 1/3$ with $h^+ = h^- = h_{GQ1} = \sqrt{3}/2$. D'Errico and Zaino [6] improved this approach suggesting the adoption of $h_{GQ1} = \sqrt{3}$ and distinctive weights $W_i = \{1/6, 2/3, 1/6\}$. Seo and Kwak [21] finally generalized it to consider non-Gaussian distributions. Those modified Taguchi methods for statistical tolerance design have an accuracy of $O(\sigma_x{}^6)$ (for symmetrical input distributions), but are very demanding for computational RDO since the number of required function evaluations is $N = 3^n$.

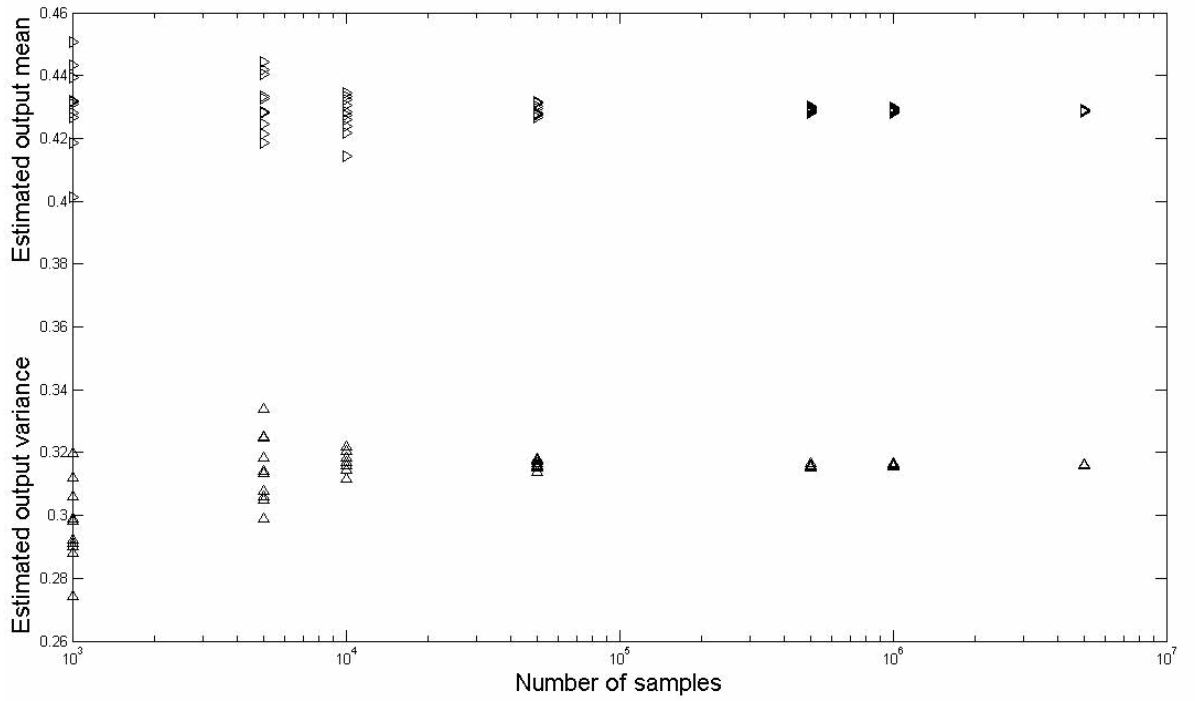

Figure 1. Convergence of mean and variance estimate for random sampling MCS

Computationally more affordable alternatives to full factorial design have been suggested. For example, Evans [8] proposed an improved integration technique for which the approximated mean and variance in the case of independent input variables are:

$$\overline{y}_{GQ2} = f(\boldsymbol{x}_0) + \sum_{p=1}^{n} W_p \left[\frac{f(\boldsymbol{x}_{p+}) - f(\boldsymbol{x}_0)}{h_p^+} - \frac{f(\boldsymbol{x}_{p-}) - f(\boldsymbol{x}_0)}{h_p^-} \right] +$$
$$+ \sum_{p=1}^{n} \sum_{\substack{q=1 \\ q \neq p}}^{n} W_{pq} \left[\frac{f(\boldsymbol{x}_{p+q+}) - f(\boldsymbol{x}_0)}{2h_p^+ h_q^+} - \frac{f(\boldsymbol{x}_{p+q-}) - f(\boldsymbol{x}_0)}{h_p^- h_q^+} + \frac{f(\boldsymbol{x}_{p-q-}) - f(\boldsymbol{x}_0)}{2h_p^- h_q^-} \right], \quad (7)$$

$$\sigma^2_{y_{GQ2}} = -[\overline{y}_{GQ2} - f(\boldsymbol{x}_0)]^2 + \sum_{p=1}^{n} W_p \left\{ \frac{\left[f(\boldsymbol{x}_{p+}) - f(\boldsymbol{x}_0) \right]^2}{h_p^+} - \frac{\left[f(\boldsymbol{x}_{p-}) - f(\boldsymbol{x}_0) \right]^2}{h_p^-} \right\} +$$
$$+ \sum_{p=1}^{n} \sum_{\substack{q=1 \\ q \neq p}}^{n} W_{pq} \left[\frac{\left[f(\boldsymbol{x}_{p+q+}) - f(\boldsymbol{x}_0) \right]^2}{2h_p^+ h_q^+} - \frac{\left[f(\boldsymbol{x}_{p+q-}) - f(\boldsymbol{x}_0) \right]^2}{h_p^- h_q^+} + \frac{\left[f(\boldsymbol{x}_{p-q-}) - f(\boldsymbol{x}_0) \right]^2}{2h_p^- h_q^-} \right], \quad (8)$$

where:

$\boldsymbol{x}_0 = \overline{\boldsymbol{x}}$,

$\boldsymbol{x}_{p\pm} = \overline{\boldsymbol{x}} + h_p^{\pm} \sigma_{x_p} \boldsymbol{e}_p$,

$\boldsymbol{x}_{p\pm q\pm} = \overline{\boldsymbol{x}} + h_p^{\pm} \sigma_{x_p} \boldsymbol{e}_p + h_q^{\pm} \sigma_{x_q} \boldsymbol{e}_q$,

and $\boldsymbol{e}_p$ is the p^{th} column of the identity matrix of size *n*. The weights and the nodes are instead:

$$W_p = \frac{1 + \sum_{i=1}^{n} (h_i^+ h_i^-)^{-1} - (h_p^+ h_p^-)^{-1}}{h_p^+ - h_p^-},$$

$$W_{pq} = \left[(h_p^+ - h_p^-)(h_q^+ - h_q^-) \right]^{-1},$$

$$h_p^{\pm} = \gamma_p / 2 \pm \sqrt{\Gamma_p - 3\gamma_p^2 / 4},$$

where $\gamma_p = E(x_p - \overline{x}_p)^3 / \sigma_{x_p}^3$ and $\Gamma_p = E(x_p - \overline{x}_p)^4 / \sigma_{x_p}^4$.

Evans showed that the error of this formula is $O(\sigma_x{}^5)$ in the general case and $O(\sigma_x{}^6)$ in the case where all the input distributions are symmetric. This method requires $2n^2 + 1$ function evaluations, and captures the skewness and the kurtosis of the output distribution as well as its mean and variance.

3.3 Taylor-based moment propagation

In the case of continuous variables and functions, an alternative method for propagating uncertainty is to approximate the statistical moments of the system response by means of a truncated Taylor series expansion of *y*. The function of interest is expanded about the mean of the input variables, and one then calculates the moments of the truncated series (typically mean and variance). This technique is also called method of moments, and classified according to the order of truncated series (denoted by a Roman number in this paper, e.g., I MM for first-order series) and the number of moments considered. From the general Taylor series expansion truncated to the fourth order:

$$y = f(\overline{\boldsymbol{x}}) + \sum_{p=1}^{n} \left(\frac{\partial f}{\partial x_p} \right) dx_p + \frac{1}{2} \sum_{p=1}^{n} \sum_{q=1}^{n} \left(\frac{\partial^2 f}{\partial x_p \partial x_q} \right) dx_p dx_q + \frac{1}{6} \sum_{p=1}^{n} \sum_{q=1}^{n} \sum_{r=1}^{n} \left(\frac{\partial^3 f}{\partial x_p \partial x_q \partial x_r} \right) dx_p dx_q dx_r +$$
$$+ \frac{1}{24} \sum_{p=1}^{n} \sum_{q=1}^{n} \sum_{r=1}^{n} \sum_{s=1}^{n} \left(\frac{\partial^4 f}{\partial x_p \partial x_q \partial x_r \partial x_s} \right) dx_p dx_q dx_r dx_s + O(d\boldsymbol{x}^5), \quad (9)$$

and assuming independent input variables, the mean and variance of *y* are given by, respectively:

$$\overline{y}_{MM} = \overbrace{f(\overline{\boldsymbol{x}})}^{M_1} + \overbrace{\frac{1}{2} \sum_{p=1}^{n} \left(\frac{\partial^2 f}{\partial x_p^2} \right) \sigma_{x_p}^2}^{M_2} + \overbrace{\frac{1}{6} \sum_{p=1}^{n} \left(\frac{\partial^3 f}{\partial x_p^3} \right) \gamma_p \sigma_{x_p}^3}^{M_3} + \overbrace{\frac{1}{24} \sum_{p=1}^{n} \left(\frac{\partial^4 f}{\partial x_p^4} \right) \Gamma_p \sigma_{x_p}^4}^{M_4} +$$

$$+\overbrace{\frac{1}{8}\sum_{p=1}^{n}\sum_{\substack{q=1\\p\neq q}}^{n}\left(\frac{\partial^4 f}{\partial x_p^2 \partial x_q^2}\right)\sigma_{x_p}^2\sigma_{x_q}^2}^{M_5} + \text{terms of order} > \sigma_x^4, \qquad (10)$$

$$\sigma_{y_{MM}}^2 = \overbrace{\sum_{p=1}^{n}\left(\frac{\partial f}{\partial x_p}\right)^2\sigma_{x_p}^2}^{V_1} + \overbrace{\sum_{p=1}^{n}\left(\frac{\partial^2 f}{\partial x_p^2}\right)\left(\frac{\partial f}{\partial x_p}\right)\gamma_p\sigma_{x_p}^3}^{V_2} + \overbrace{\sum_{p=1}^{n}\sum_{\substack{q=1\\q\neq p}}^{n}\left(\frac{\partial^3 f}{\partial x_p^2 \partial x_q}\right)\left(\frac{\partial f}{\partial x_q}\right)\sigma_{x_p}^2\sigma_{x_q}^2}^{V_3} +$$

$$+\overbrace{\frac{1}{2}\sum_{p=1}^{n}\sum_{\substack{q=1\\q\neq p}}^{n}\left(\frac{\partial^2 f}{\partial x_p \partial x_q}\right)^2\sigma_{x_p}^2\sigma_{x_q}^2}^{V_4} + \overbrace{\frac{1}{3}\sum_{p=1}^{n}\left(\frac{\partial^3 f}{\partial x_p^3}\right)\left(\frac{\partial f}{\partial x_p}\right)\Gamma_p\sigma_{x_p}^4}^{V_5} + \overbrace{\frac{1}{4}\sum_{p=1}^{n}\left(\frac{\partial^2 f}{\partial x_p^2}\right)^2\left(\Gamma_p - 1\right)\sigma_{x_p}^4}^{V_6} + \qquad (11)$$

$$+ \text{terms of order } > \sigma_x^4.$$

I MM has been largely applied to uncertainty problems [7, 11, 12, 18, 19]. However, unless one restricts the analysis to very small values of the input variance its accuracy is often unsatisfactory in the frequent case of highly nonlinear functions. A seemingly obvious improvement could be achieved by retaining second-order terms in the truncated Taylor series. Indeed this guarantees a better accuracy of the mean estimate. The term M2 in Eq. (10), in fact, increases the accuracy of the mean to $O(\sigma_x^3)$, which is above the $O(\sigma_x^2)$ error associated with the linearization. As for the output variance, however, retaining only second-order terms of the Taylor series does not necessarily improve the accuracy with respect to that of the linearization. In order to explain this, let us consider the case of symmetrically distributed input variables. In this circumstance, the term V2 in Eq. (11) is zero, and a variance approximation which correctly captures terms of $O(\sigma_x^4)$ can be obtained by retaining not only the terms V4 and V6, but also the terms V3 and V5. These latter, however, contain contributions of the third derivatives, and their magnitude relative to that of the other two $O(\sigma_x^4)$ terms is not known a priori. When V3 and V5 are omitted (i.e. only derivatives up to second order are considered), one cannot define rigorously the order of the variance approximation. In practice, if the third derivatives appearing in V3 and V5 are small, one may increase the range of input variances above that of the linearization and enhance the accuracy of the variance by doing so. However, if the third derivatives are not sufficiently small, the variance estimate will be worse than the result based on the first-order approximation. To emphasize this point, let us consider again the function $y = \sin(x)$, with x normally distributed about $\overline{x} = \pi/4$. The left and right plots of Figure 2 report $\overline{y}$ and σ_y^2 against σ_x^2, respectively, using I MM, II MM and III MM. Note that retaining the second-order derivatives increases the accuracy of $\overline{y}$ compared to the linearization, while it causes σ_y^2 to worsen fairly rapidly.

The main advantage of MM is its computational efficiency, which depends on the available methods for calculating derivatives. A reasonably accurate prediction of the output variance of highly nonlinear functions depending on significantly spread inputs requires retention of higher-order derivatives. These latter could be accurately and efficiently determined by using adjoint and/or forward code differentiation. Derivative codes can be developed by hand and/or using AD. In most applications of practical interest, however, this can be done only if the source code of the analysis program is available. Besides it should be noted that both manual and automatic code differentiation for derivatives beyond the second is a fairly complex task.

4 SIGMA-POINT METHODS

The polynomial approximation obtained by truncating the Taylor series after a certain number of terms may be regarded as an extrapolation process. The fact that the MM estimates of mean and variance worsen as the input variance increases can be viewed as a consequence of this feature. Moreover, obtaining higher-order derivatives for adequately approximating highly nonlinear functions may be a problematic issue. Indeed neither AD nor manual code differentiation may be sufficient to obtain the codes that calculate such derivatives. On the other hand, the use of Finite Difference approximations

would be computationally costly, and would raise concerns about the undesired effects of truncation errors. An alternative to increasing the resolution of the approximation in a single point, typically the input mean, is to use the information collected at multiple points in the input space. In this Section, we present two methods based on this principle, known as Sigma-point (SP) approach: the Unscented Transformation (UT) [15] and the Divided Difference Filter (DDF) [17]. Both methods were developed in Control Theory to overcome the limitations of the Evolved Kalman Filter, which uses linearization for state estimation of nonlinear systems. The UT and DDF methods share notable similarities with the GQ techniques. They require less function evaluations, and this is traded off with the capability of capturing only the first two moments of the output distribution. These latter, however, are the only ones of interest in most RDO problems.

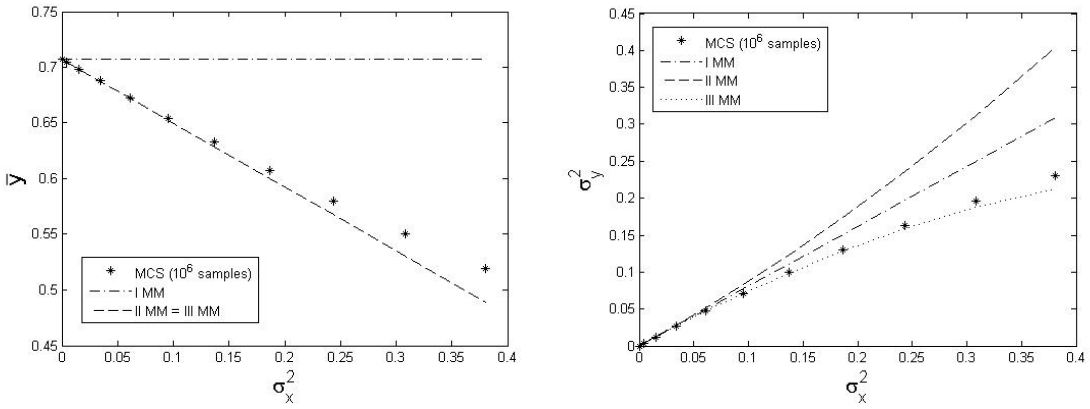

Figure 2. Mean and variance estimate for MM compared with MCS

The UT, which has been considered for robust design applications in [5] and [23] as an alternative to MCS, proceeds from an intuitive consideration [15]: the probabilistic description of *y* is more conveniently obtained by approximating the input distribution rather than the function ***f(x)***. Approximating the input distribution means choosing a sufficient number of key points (Sigma points) in the input space, and evaluating here the function *y*. Assuming symmetric input distributions, the Sigma points are also symmetrically spread about the mean $\bar{\boldsymbol{x}}$, at a distance which depends on the input covariance matrix. For the case of independent input variables, the Sigma points are:

$$\boldsymbol{x}_0 = \bar{\boldsymbol{x}}, \tag{12}$$

$$\boldsymbol{x}_{p\pm} = \bar{\boldsymbol{x}} \pm h_{SP}\,\sigma_{x_p}\boldsymbol{e}_p, \tag{13}$$

where $\boldsymbol{e}_p$ is the p^{th} column of the identity matrix of size *n* and h_{SP} is a suitable scalar; the subscript *SP* is used because these considerations hold for all the SP approaches. The function values $y_p = f(\boldsymbol{x}_p)$ are then used to calculate the mean and variance of the output by means of following weighted sums:

$$\bar{y}_{UT} = W_0 f(\boldsymbol{x}_0) + \sum_{p=1}^{n} W_p \left[f(\boldsymbol{x}_{p+}) + f(\boldsymbol{x}_{p-}) \right], \tag{14}$$

$$\sigma^2_{y_{UT}} = W_0 \left[f(\boldsymbol{x}_0) - \bar{y}_{UT} \right]^2 + \sum_{p=1}^{n} W_p \left\{ \left[f(\boldsymbol{x}_{p+}) - \bar{y}_{UT} \right]^2 + \left[f(\boldsymbol{x}_{p-}) - \bar{y}_{UT} \right]^2 \right\}. \tag{15}$$

The weights are chosen as follows:

$$W_0 = \frac{h_{SP}^2 - n}{h_{SP}^2}, \tag{16}$$

$$W_p = \frac{1}{2h_{SP}^2}, \text{ for } 1 \le p \le n. \tag{17}$$

The DDF, on the other hand, builds a local second-order polynomial approximation of the function of interest by using Stirling interpolation formula. The Sigma points and the weights are chosen exactly as in the UT; the resulting mean estimate $\bar{y}_{DDF}$ is calculated as $\bar{y}_{UT}$ in Eq. (14), while the following formula is proposed for variance estimation [17]:

$$\sigma^2_{y_{DDF}} = \frac{1}{2}\sum_{p=1}^{n}\left\{W_p\left[f(\boldsymbol{x}_{p+}) - f(\boldsymbol{x}_{p-})\right]^2 + \left(W_p - 2W_p^2\right)\left[f(\boldsymbol{x}_{p+}) + f(\boldsymbol{x}_{p-}) - 2f(\boldsymbol{x}_0)\right]^2\right\}. \quad (18)$$

It can be shown that most higher-order terms in the expressions (10) and (11) for mean and variance can be captured by assigning to h_{SP}^2 the kurtosis of the input distribution [15, 16]. For the Normal distribution, it turns out that $h_{SP} = \sqrt{3}$, as in the quadrature rules discussed in subsection 3.2.

The SP approaches require $2n+1$ function evaluations for each analysis, and are derivative-free, which means that, in principle, they can handle discontinuous functions and discrete variables. When the considered function is differentiable, these techniques can be efficiently coupled with gradient-based optimizers: since the SP robust objective is not built using derivatives, its gradient is simply a combination of gradients of the deterministic objective. On the other hand, MM gradient-based optimization would require at least one more level of differentiation, in case of first-order approximations, and even more levels with higher-order representations. Thus, the computational cost of a single step of gradient-based RDO, in terms of function evaluations, is not altered significantly by using the SP methods instead of linearization for the propagation phase. However, as explained in the following subsection, the SP methods can give a more accurate robust objective with respect to I MM.

4.1 Accuracy of the Sigma-point methods

The accuracy of the SP methods can be assessed by comparing the expressions of mean and variance yielded by Taylor-based MM of Eqs. (10) and (11), and those obtained by expanding each $y_p = f(\boldsymbol{x}_p)$ in Eqs. (14), (15) and (18) into fourth-order Taylor series. Choosing h_{SP}^2 as the kurtosis of the input variables, the error terms in the case of independent symmetrically distributed inputs are given by:

$$\varepsilon_{\bar{y}_{UT}} = \varepsilon_{\bar{y}_{DDF}} = \frac{1}{8}\sum_{p=1}^{n}\sum_{\substack{q=1\\p\neq q}}^{n}\left(\frac{\partial^4 f}{\partial x_p^2 \partial x_q^2}\right)\sigma_{x_p}^2\sigma_{x_q}^2 + \text{terms of order} > \sigma_x^4; \quad (19)$$

$$\varepsilon_{\sigma^2_{UT}} = \sum_{p=1}^{n}\sum_{\substack{q=1\\p\neq q}}^{n}\left[\frac{1}{2}\left(\frac{\partial^2 f}{\partial x_p \partial x_q}\right)^2 + \left(\frac{\partial^3 f}{\partial x_p^2 \partial x_q}\right)\left(\frac{\partial f}{\partial x_q}\right) + \frac{1}{4}\left(\frac{\partial^2 f}{\partial x_p^2}\right)\left(\frac{\partial^2 f}{\partial x_q^2}\right)\right]\sigma_{x_p}^2\sigma_{x_q}^2 + \quad (20)$$

$$+ \text{terms of order} > \sigma_x^4;$$

$$\varepsilon_{\sigma^2_{DDF}} = \sum_{p=1}^{n}\sum_{\substack{q=1\\p\neq q}}^{n}\left[\frac{1}{2}\left(\frac{\partial^2 f}{\partial x_p \partial x_q}\right)^2 + \left(\frac{\partial^3 f}{\partial x_p^2 \partial x_q}\right)\left(\frac{\partial f}{\partial x_q}\right)\right]\sigma_{x_p}^2\sigma_{x_q}^2 + \text{terms of order} > \sigma_x^4. \quad (21)$$

The SP methods lead to a better mean estimate compared to I MM, while the variance estimate has the same order of accuracy of I MM. Besides, the SP methods miss some $O(\sigma_x^4)$ contributions in the variance, similarly to II MM. Nevertheless, in the case of functions with weak interaction between input variables, i.e. functions for which the cross-derivatives in Eqs. (20) and (21) are negligible with respect to the other terms, the SP methods may approximate the variance better than I MM.

5 RESULTS

5.1 Uncertainty Propagation Test Cases

The SP methods have been tested with a number of simple, but representative test cases, and compared with results obtained by MCS and MM. The following functions have been considered:

1. $f(\boldsymbol{x}) = \sin(x_1 - 0.21)\sin(x_2 - 0.21)$;
2. $f(\boldsymbol{x}) = 0.5x_1^2 - 1.5x_1 + 0.7x_2^2 - 1.2x_2 + 1.05$;

3. $f(\boldsymbol{x}) = 1.7x_1^3 + 1.3x_1^2 - 2.4x_1 - 0.5x_2^3 + 3.2x_2^2 - 1.6x_2 + 0.1x_1^2x_2 - 0.2x_1x_2^2 + 1.6x_1x_2 - 12.8$;

4. $f(\boldsymbol{x}) = 0.4x_1^2 + 0.7x_1 + 0.5x_2^3 - 1.1x_2^2 - 0.9x_2 - 1.3x_2^2x_1$.

Independent normally distributed input variables with unit mean have been used in all cases, and the analysis has been performed for several input variances. The output mean estimates based on I MM, III MM and the two SP methods are reported in the last three columns of Table 1. Note that a) these values are given as percentage errors with respect to the result of MCS with 10^6 samples, b) the column labelled 'III MM' is obtained using Eq. (10) and omitting the fourth derivative, c) assuming symmetric input distribution, the mean based on III MM and II MM do coincide. These results highlight the anticipated accuracy improvement of the mean estimate obtained by using the SP methods rather than linearization. The variance estimates are reported in Table 2, wherein all results are again provided as percentage errors with respect to MCS with 10^6 samples and the column labelled 'III MM' is based on Eq. (11).

Table 1. Uncertainty Propagation Results: Mean Estimation

Test Number	σ_x	Percentage error with respect to MCS (10^6 samples)		
		I MM	III MM	UT and DDF
1	0.05	0.250	-0.000	-0.000
	0.2	4.082	-0.081	-0.039
	0.5	28.395	-3.703	-1.746
2	0.05	0.671	0.000	0.000
	0.2	11.940	0.000	0.000
	0.5	200.000	0.000	0.000
3	0.05	0.044	-0.000	-0.000
	0.2	0.720	-0.000	-0.000
	0.5	4.679	-0.000	-0.000
4	0.05	-0.051	0.000	0.000
	0.2	-0.824	-0.000	-0.000
	0.5	-4.931	0.004	0.004

Table 2. Uncertainty Propagation Results: Variance Estimation

Test Number	σ_x	Percentage error with respect to MCS (10^6 samples)			
		I MM	III MM	UT	DDF
1	0.05	0.220	-0.030	0.033	0.096
	0.2	4.054	-0.107	0.977	1.998
	0.5	27.031	-4.727	5.285	12.286
2	0.05	-1.224	0.036	-0.560	0.036
	0.2	-16.938	0.017	-8.002	0.017
	0.5	-56.068	-0.017	-26.527	-0.017
3	0.05	-0.119	-0.049	-0.190	-0.192
	0.2	-1.111	-0.003	-2.226	-2.257
	0.5	-7.058	-0.551	-13.254	-13.439
4	0.05	-0.079	0.000	-0.024	-0.026
	0.2	-1.270	-0.019	-0.407	-0.433
	0.5	-7.354	-0.017	-2.276	-2.430

Table 2 reveals a more variegated landscape than Table 1. In fact, there are cases for which the partial inclusion of the higher-order terms associated with the SP methods yields better accuracy than the linearization over the whole range of considered variances (cases 1 and 2). Conversely, in test cases 3 and 4 the cross-terms of Eqs. (20) and (21) are not negligible, and this greatly spoils the accuracy of the SP methods for large input variances. Hence, as for the Taylor-based MM, the presence of strong nonlinearities restricts the range of input variances one may use. To overcome this limitation, one ought to add more points (following the quadrature approach), or higher-order derivatives. Unfortunately, an accurate calculation of higher-order derivatives is often not possible in real world applications, because the function of interest may be defined by sophisticated computer programs that may be very hard to differentiate. Besides, the source code may even not be accessible. Finite differencing is a potential alternative. Unfortunately, it has two significant drawbacks: a) it is

computationally expensive as it requires numerous executions of the analysis program, and b) it may be insufficiently accurate because of truncation errors. These are the reasons for seeking alternatives to MM for propagating uncertainty in robust Engineering Design. In some cases, however, it is possible to use the AD technology either to calculate higher-order derivatives for MM and/or low-order derivatives to enhance the performance of SP methods in robust optimization. An example of both features is provided in next subsection.

5.2 Robust Optimization Test Case

In aeronautical conceptual design, the aircraft layout is determined on the basis of fundamental specifications such as range, number of passengers, take-off and landing field length, and all analyses are performed with simplified models. All these analyses also involve other parameters such as approach speed, maximum take-off weight, noise emission, acquisition and operating cost, which can be treated either as inequality constraints or objectives to be optimized. Subsequent preliminary and detail design may yield variations of the original layout. Robustness of conceptual design optimization is thus required to minimize possible specification and constraint violations induced by such changes.
The aircraft conceptual design test case that we have selected to demonstrate the strength of the Sigma-Point approach for realistic problems has 96 models and 126 variables. The robust optimization problem is set as follows:

- Objective: optimize *MTOW (Maximum take-off weight)*;
- Fixed parameters: *Number of passengers* N_{pax}=150, *Number of engines* N_e=*2*, *Cruise Mach* M_c=0.75 *and altitude* h=31000 ft;
- Independent variables $\boldsymbol{x}$: *Wing area* S, *Wing span* b, *Wing sweep* Λ, *Thickness to chord ratio t/c, Engine sea level thrust* F_e. The respective ranges of variation are: $150\ \mathrm{m}^2 \le S \le 160\ \mathrm{m}^2$, $31\ \mathrm{m} \le b \le 35\ \mathrm{m}$, $19\ \mathrm{deg} \le \Lambda \le 26\ \mathrm{deg}$, $0.08 \le t/c \le 0.12$, $135\ \mathrm{kN} \le F_e \le 150\ \mathrm{kN}$;
- Constraints considered: *Range R* > 4600 Km, *Take-off field length TOFL* < 2000 m, *Wing fuel/Fuselage fuel* K_{fn} < 0.75, *Cruise thrust coefficient* K_{fc} < 1 and *Climb speed* V_z < 500 ft/min.

The input variables are independent Gaussian variables with $\sigma_x / \mu_x = 0.1$. The objective function has been defined as $MTOW_{rob} = \overline{MTOW} + \sigma_{MTOW}$ and the constraints take the following form: $g_c = \bar{g}(\boldsymbol{x}) + 2\sigma_g(\boldsymbol{x}) \le 0$. Mathematically, two optimization problems differing only for the calculation of mean and variance of objective *MTOW* and constraints have been set up. The considered methods are: I MM and DDF. Numerically the two optimization problems have been solved in MATLAB using the gradient-based optimizer *fmincon*. The latest version of the MATLAB overloading AD software MAD developed by Forth [9, 10] has been used to support the solution of the robust optimizations. MAD computes first-order derivatives using the adjoint differentiation mode, and higher-order derivatives using the forward mode. This AD tool can also propagate derivatives through most built-in MATLAB functions such as *fsolve*. We have used MAD in two different ways. In the I MM optimization, MAD yields the first derivatives of the deterministic objective function *MTOW* required to build its variance, and its second derivatives to build the gradient of the robust objective $MTOW_{rob,opt}$ for the optimizer. Using MAD, the calculation of $MTOW_{rob,opt}$ and its gradient is approximately (*n+2*) times that of a single evaluation of *MTOW* due to the use of the adjoint mode for first-order differentiation. Finally, in the SP-based optimizations, MAD is used to calculate the gradient of $MTOW_{rob,opt}$. The cost of computing $MTOW_{rob,opt}$ and its gradient is therefore (4*n*+2) times that of a single *MTOW* calculation. Table 3 presents the results of the three optimizations. In order to partially validate these results, three MCS's with 10^6 samples centred at $\boldsymbol{x}_{opt}$ have also been performed, and the last two rows report the values of $\overline{MTOW}_{opt}$ and $\sigma_{MTOW,opt}$ thus obtained. Overall Table 3 shows that the SP solutions are in good agreement with those of MCS and I MM. This supports the assumption that SP methods can be effectively used to improve computational RDO. In fact, the computational cost required by an SP-based optimization step is comparable to that of the linearization, and the mean estimate of SP is accurate to $O(\sigma_x{}^3)$ whereas that of the linearization is accurate only to $O(\sigma_x{}^2)$. This

feature is likely to make SP-based RDO applicable over a wider range of input variances compared to the linearization-based methods.

Table 3. Robust Optimization Results

Method	I MM	SP
$MTOW_{rob,opt}$ [Kg]	91804.976	92026.726
$\overline{MTOW}_{opt}$ [Kg]	90491.711	90710.430
$\sigma_{MTOW,opt}$ [Kg]	1313.264	1316.295
S [m^2]	159.968	160.000
b [m]	31.000	31.000
Λ [deg]	19.000	19.000
t/c	0.080	0.080
F_e [kN]	135.000	138.457
$\overline{MTOW}_{opt,MCS}$	90489.216	90710.442
$\sigma_{MTOW,opt,MCS}$	1327.358	1322.498

6 CONCLUSIONS

Uncertainty propagation is crucial to both accuracy and efficiency of computational robust design. I MM is largely used to propagate uncertainty due to its simplicity and convenience. The applicability of this method, however, is limited to fairly small input variances when significant nonlinearities exist. This paper has presented the SP approach for uncertainty propagation along with a comparison with MCS, GQ and higher-order MM. The presented error analysis and the solution of some simple but representative numerical examples show that SP methods can lead to better mean estimates compared to I MM. The error analysis also shows that the accuracy of the SP variance is at least as good as the one based on the linearization, and even higher for functions with small interaction between variables. The SP approach is derivative-free and needs $2n+1$ function evaluations for each analysis. Since the SP robust objective is not built using derivatives, its gradient is simply a combination of the gradients of the deterministic objective. Thus, when the objective function is differentiable and the SP method is used with gradient-based optimization, the cost of each step of the robust optimization in terms of function evaluations is comparable to that based on the linearization. An interesting special case is that in which the source code is available and code differentiation is used. This feature has been highlighted by solving a robust optimization problem for the conceptual design of a civil aircraft, with a gradient-based method and computing derivatives with the AD tool MAD, which uses adjoint and forward differentiation.

ACKNOWLEDGMENTS

The research reported in this paper has been carried out within the VIVACE Integrated Project (AIP3 CT-2003-502917) which is partly sponsored by the Sixth Framework Programme of the European Community (2002-2006) under priority 4 “Aeronautics and Space”.

The authors also wish to thank Dr. Shaun Forth from the Engineering Systems Department, Defence College of Management & Technology, Cranfield University, Defence Academy, Shrivenham for providing the software MAD and supporting its deployment.

REFERENCES

[1] Barthelemy J.F. and Hall L.E. Automatic Differentiation as a Tool in Engineering Design. August 1992, TM 107661 (NASA).

[2] Bischof C.H. and Griewank A. Computational Differentiation and Multidisciplinary Design. In *Inverse Problems and Optimal Design in Industry*, H. Engl and J. McLaughlin, eds., 1994, pp. 187–211 (Teubner Verlag, Stuttgart).

[3] Brass H., Fischer J.W. and Petras K. The Gaussian quadrature method. *Abh. Braunschweig. Wiss. Ges.*, 1996, 47, 115–150 (J. Cramer Verlag, Braunschweig).

[4] Buranathiti T., Cao J. and Chen W. A weighted three-point-based strategy for variance estimation. In *Proceedings of DETC’2004*, Salt Lake City, September 28–October 2 2004 (ASME).

[5] Cioffi M., Formisano A., Martone R., Steiner G. and Watzenig G. A Fast Method for Statistical Robust Optimization. *IEEE Transactions of Magnetics*, 2006, 42(4), 1099–1102.
[6] D'Errico J.R. and Zaino N.A.Jr. Statistical tolerancing using a modification of Taguchi's method. *Technometrics*, 1988, 30(4), 397-405.
[7] Du X. and Chen W. Efficient Uncertainty Analysis Methods for Multidisciplinary Robust Design. *AIAA Journal*, 2002, 40(3), 545–552.
[8] Evans D.H. Statistical Tolerancing: the state of the art, part II. *J. Qual. Tech.*, 1975, 7(1), 1–12.
[9] Forth S.A. An Efficient Overloaded Implementation of Forward Mode Automatic Differentiation in MATLAB. *ACM Transactions on Mathematical Software (TOMS)*, 2006, 32(2), 195–222.
[10] Forth S.A. and Edvall M.M. User Guide for Mad - a Matlab Automatic Differentiation Toolbox Version 1.4, 2006 (http://tomopt.com/docs/TOMLAB_MAD.pdf, accessed on 15/01/07).
[11] Green L.L., Lin H.Z. and Khalessi M.R. Probabilistic methods for uncertainty propagation applied to aircraft design. In *Proceedings of 20th AIAA Applied Aerodynamics Conference*, 24-26 June 2002, St. Louis (AIAA).
[12] Gumbert C.R., Newman P.A., and Hou G.J.-W. Effect of random geometric uncertainty on the computational design of a 3-D flexible wing. In *Proceedings of 20th AIAA Applied Aerodynamics Conference*, 24-26 June 2002, St. Louis (AIAA).
[13] Halton J.H. A Retrospective and Prospective Survey of the Monte Carlo Method. *SIAM Review*, 1970, 12(1), 1–63.
[14] Helton J.C. and Davis F.J. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Eng. System Safety*, 2003, 81(1), 23–69.
[15] Julier S.J. and Uhlmann J.K. Unscented Filtering and Nonlinear Estimation. *Proc. IEEE*, 2004, 92(3), 401–422.
[16] Nørgaard M., Poulsen N.K. and Ravn O. Advances in derivative-free state estimation for nonlinear systems. Tech. Rep. IMM–REP 1998–15, Technical Univ. of Denmark, April 2000.
[17] Nørgaard M., Poulsen N.K. and Ravn O. New developments in state estimation of nonlinear systems. *Automatica* 2000, 36, 1627–1638.
[18] Park G.J., Lee T.H. and Hwang K.H. Robust Design: An Overview. *AIAA Journal*, 2006, 44(1), 181–191.
[19] Parkinson A., Sorensen C., Pourhassan N. A General Approach for Robust Optimal Design. *J. mech. Des*, 1993, 115(1), 74–80.
[20] Phadke M.S. Quality Engineering Using Robust Design, 1989 (Prentice Hall, Englewood Cliffs).
[21] Seo H.S. and Kwak B.M. Efficient statistical tolerance analysis for general distributions using three-point information. *Int. J. Prod. Res.*, 40(4):931–944, 2002.
[22] Simpson T.W., Poplinski J.D., Koch P.N. and Allen, J.K., Metamodels for Computer-based Engineering Design: Survey and recommendations. *Eng. Comput. (Lond.)*, 2001, 17(2), 129-150.
[23] Steiner G., Watzenig G., Magele C. and Baumgartner U. Statistical robust design using the unscented transformation. *Compel*, 2005, 24(2):606–619.
[24] Su J. and Renaud J.E. Automatic Differentiation in Robust Optimization. *AIAA Journal*, 1996, 5(6), 1072–1079.
[25] Taguchi G. Performance Analysis Design. *Int. J. Prod. Res.*, 1978, 16, 521–530.
[26] Tari M. and Dahmani A. Refined descriptive sampling: A better approach to Monte Carlo simulation. *Simulation Modelling Practice and Theory*, 2006, 14(2), 143–160.
[27] Xiu D. and Karniadakis E.M. Modeling uncertainty in flow simulations via generalized polynomial chaos. *Journal of Computational Physics*, 2003, 187, 137–167.
[28] Youn B.D. and Choi K.K. Selecting Probabilistic Approaches for Reliability-Based Design Optimization. *AIAA Journal*, 2004, 42 (1), 124–131.

Contact: M. Padulo
Cranfield University, SoE
Dept. of Aerospace Engineering
MK43 0AL Cranfield,
UK
Phone: +44 (0) 1234 750111 ext. 5715
Fax: +44 (0) 1234 758203
e-mail: m.padulo@cranfield.ac.uk
URL: http://www.cranfield.ac.uk/soe/advanced_engineering_design/