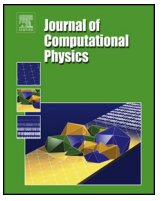
Journal of Computational Physics 491 (2023) 112377



Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp



Sensitivity-enhanced generalized polynomial chaos for efficient uncertainty quantification



Kyriakos D. Kantarakias, George Papadakis *

Department of Aeronautics, Imperial College London, SW7 2AZ, UK

ARTICLE INFO

Article history:
Received 19 June 2022
Received in revised form 2 July 2023
Accepted 12 July 2023
Available online 18 July 2023

Keywords:
Uncertainty quantification
Generalised polynomial chaos
Sensitivity-enhanced ℓ_2 minimisation
Optimal sampling of stochastic space

ABSTRACT

We consider the Least Squares (LSQ) regression method for Uncertainty Quantification (UQ) using generalised polynomial chaos (gPC) and augment the linear system with the gradient of the Quantity of Interest (QoI) with respect to the stochastic variables. The gradient is computed very efficiently for all variables from the adjoint system of equations. To minimise the condition number of the augmented LSQ system, an effective sampling strategy of the stochastic space is required. We compare two strategies. In the first, we apply pivoted QR decomposition to the standard LSQ matrix and evaluate both the QoI and its gradient at the sample points identified. In the second strategy, we apply QR decomposition directly to the augmented matrix. We find that the first strategy is more efficient in terms of accuracy vs number of evaluations. We call the new approach sensitivity-enhanced generalised polynomial chaos, or se-gPC, and apply it to several test cases including an aerodynamic case with 40 stochastic parameters. The method can produce accurate estimations of the statistical moments using a small number of sampling points. The computational cost scales as $\sim m^{p-1}$, instead of $\sim m^p$ of the standard LSQ formulation, where m is the number of stochastic variables and p the chaos order. The solution of the adjoint system of equations is implemented in many computational mechanics packages, thus the infrastructure exists for the application of the method to a wide variety of engineering problems.

© 2023 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

In all practical applications, the performance of an engineering system is affected by stochastic uncertainties in the system parameters, boundary conditions etc. Significant research has been directed towards quantifying the effect of such uncertainties on the performance of engineering systems. The performance is evaluated by a metric, called Quantity of Interest (QoI), that depends on the system under investigation. For example, in the area of aerodynamics it can be the lift or drag coefficient of an airfoil, and the uncertainty may arise due to stochastic variations in the angle of attack, approaching velocity etc. This area of research is known as Uncertainty Quantification (UQ).

The simplest UQ approaches are random sampling techniques, such as the Monte–Carlo (MC) [1,2]. Such approaches however require a large number of samples for the statistics of the QoI to converge and are therefore computationally expensive. On the other hand, methods based on Polynomial Chaos Expansion (PCE) have much reduced computational cost. In these approaches, the uncertain QoI is expanded in terms of multivariate polynomial bases that are orthogonal with

* Corresponding author.
E-mail address: g.papadakis@imperial.ac.uk (G. Papadakis).

https://doi.org/10.1016/j.jcp.2023.112377
0021-9991/© 2023 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

respect to the joint probability density function (PDF) of the uncertain parameters. The method was first introduced in [3] for Gaussian stochastic variables and later extended for other distributions in [4]. These methods have been very successful in predicting the statistics of the QoI in a wide range of applications, such as applied mechanics, fluid dynamics, space, medicine, etc [5–10].

In the standard Galerkin-projection formulation of these methods, the cost of computing the spectral coefficients of the QoI expansion scales exponentially with the dimension of the stochastic space. For the non-intrusive projection methods, this is due to the large number of sampling points when extending one-dimensional quadrature integration schemes to multi-dimensions using tensor products, for details see the book of [11]. At each sampling point, the evaluation of the model describing the system is by far the most time consuming part of the method. In sparse grid approaches such as Smolyak grids, [12–16], the spectral coefficients are computed from evaluations at a significantly smaller number of sample points. In [17] an adaptive method to build a sparse polynomial chaos (PC) expansion basis using least-angle regression (LARS) is introduced that requires a reduced number of evaluations.

Another approach to compute the spectral coefficients is based on least-squares (LSQ) regression, that minimises the expectation of the square of the difference between the QoI evaluation and the PCE prediction at different sampling points, [18,17,11]. For the LSQ regression to provide accurate results, the model has to be evaluated at a number of sampling points, say q, which scales with the number of unknown PCE coefficients. The oversampling ratio is usually between 2-3 [17] or even larger [19]. For m stochastic variables and polynomial chaos order p, the number of PCE coefficients is $\frac{(p+m)!}{m!p!}$, see also equation (4) below, which behaves asymptotically as $\sim \frac{1}{p!}m^p$ when $m \to \infty$, [17]. Thus, the number of evaluations grows rapidly with increasing p and m, and the computational cost quickly becomes intractable. Similarly, for Smolyak quadrature the computational cost scales as $\sim \frac{2^p}{p!}m^p$, [17]. For the smallest chaos order $p = 1$, the cost grows linearly with m in either approach.

In the standard LSQ formulation, the model evaluation at each sampling point provides one equation for the linear system of the PCE coefficients. The system can be augmented with the gradient (or sensitivity) of the QoI with respect to the stochastic variables that can be obtained very efficiently using the adjoint formulation [20–23]. This idea was explored in [24–28] for LSQ (i.e. ℓ_2 norm) minimisation and more recently in [29–31] for ℓ_1 norm minimisation. Sensitivity-enhanced Kriging methods have been also proposed and applied to UQ, see for example [32–36] for fluid mechanics applications.

Harnessing the adjoint formulation to perform UQ offers clear computational benefits. The cost of the solution of the (linear) adjoint system is comparable to that of a single model evaluation. Therefore, at each sampling point we can obtain $1+m$ equations at the cost of two evaluations, one for the direct and one for the adjoint system. Thus for q sampling points, we have a linear system of $q(1+m)$ equations, and if we make this number of equations commensurate to the number of unknown PCE coefficients, we get $q(1+m) \sim \frac{(p+m)!}{m!p!}$, which for large m becomes $q(1+m) \sim m^p$ or $q \sim m^{p-1}$. Thus the computational cost now scales as $\sim m^{p-1}$, as opposed to $\sim m^p$ of the standard formulation. For example, for $p = 1$ the cost becomes independent of m.

However, two major issues however arise when gradients are included in the formulation of the LSQ problem. The first is that the standard PCE basis functions (for example Hermite polynomials) are no longer orthonormal with respect to an inner product that includes their derivatives. The second is that currently there is no guidance at how to select the sampling points, see section V of [28] for more details about these two issues. The question of basis functions was addressed in [27], where new polynomial bases were derived that satisfy the orthonormality with respect to the new inner product. In particular, the authors showed that a tensor product multivariate orthogonal polynomial basis of an arbitrary degree can no longer be constructed.

The answer to the second question however remains open. In the aforementioned papers, the sampling of the stochastic space was random. When the standard bases are used to form the measurement matrix, the condition number (CN) of the LSQ system can be high. For example, for $m = 12$ uncertain variables and chaos order $p = 3$ and using standard Hermite polynomials, values of CN of $\mathcal{O}(10^{17})$ and $\mathcal{O}(10^{13})$ were reported in [27] for 32 and 72 randomly sampled points respectively. When the newly derived bases were used, the CN improved significantly; values of $\mathcal{O}(10^5)$ and $\mathcal{O}(10^3)$ were reported for the two sampling cases respectively (oversampling reduces the CN).

On the other hand, significant progress has been made on optimal sampling for the standard LSQ problem, see [37–40]. A review of the techniques is provided in [19]. Several optimality criteria are discussed, such as the alphabetic D-, A-, E-, K- and I-criteria, and a hybrid sampling method, called alphabetic-coherence-optimal, is proposed. In the present paper, we apply the D-design algorithm (that maximises the determinant of the covariance matrix) to provide guidance on the selection of the sampling points. More specifically, we compare two sampling strategies. In the first, we apply pivoted QR decomposition to the standard LSQ matrix and evaluate both the QoI and its gradient at the sample points identified. In the second strategy, we apply QR decomposition directly to the augmented matrix. We find that the first strategy is more efficient, and proceed to apply the method to several test cases, including an aerodynamic case with 40 stochastic parameters.

The paper is organized as follows. In section 2, an overview of UQ using gPC with LSQ is outlined. Section 3 presents the formulation of sensitivity-enhanced gPC, followed by section 4 that reviews the adjoint method for efficient computation of sensitivities. Section 5 discusses the sampling strategies of the input stochastic space. Results for several test cases are presented in section 6 and we conclude in section 7.

2. Uncertainty quantification with gPC

Let $\mathbb{P} = (\Omega, \Sigma, d\mathcal{P})$ denote a complete probability space, where Ω is the sample set of random events, and $d\mathcal{P}$ is a probability measure defined on the σ-algebra Σ. An uncertain input is modelled as a stochastic vector $\boldsymbol{\xi}$ of m independent stochastic variables ξ_i, i.e. $\boldsymbol{\xi} = [\xi_1, \dots, \xi_m] : \Omega \to \mathbb{R}^m$ with $m \in \mathbb{N}$. We use subscripts to denote random variables, and superscripts for realisations, for example the $i-th$ realisation of vector $\boldsymbol{\xi}$ is $\boldsymbol{\xi}^{(i)} = [\xi_1^{(i)}, \dots, \xi_m^{(i)}]$. The stochastic variable ξ_i follows a probability density function (PDF) $w_i(\xi_i)$, defined in the domain $\mathcal{E}_i$. Since the variables are independent, the m-dimensional vector $\boldsymbol{\xi}$ follows the PDF $W = \prod_{i=1}^{m} w_i(\xi_i)$, defined in the domain $\mathcal{E} = \prod_{i=1}^{m} \mathcal{E}_i$. A polynomial basis with members $\Psi_j(\boldsymbol{\xi})$ that is orthogonal with respect to W is defined as,

$$\langle \Psi_j, \Psi_k \rangle = \int_{\mathcal{E}} \Psi_j \Psi_k W d\boldsymbol{\xi} = \delta_{jk} \langle \Psi_j, \Psi_j \rangle. \tag{1}$$

The polynomials are normalized so that $\langle \Psi_j, \Psi_j \rangle = 1$. Note that unless specified otherwise, repeated indices don't imply summation. When $m > 1$, the multidimensional basis functions Ψ_j are defined as tensor products of the univariate polynomials ψ, for details see [11].

A Quantity of Interest (QoI) $\mathcal{M} : \boldsymbol{\xi} \to \mathcal{M}_\mu \in \mathbb{R}$ with finite variance can be written in terms of Ψ_i as,

$$\mathcal{M}(\boldsymbol{\xi}) = \sum_{i=0}^{\infty} c^{(i)} \Psi_i(\boldsymbol{\xi}) \tag{2}$$

In practise, the above expansion is truncated by limiting the maximum order of the polynomials $\Psi_i(\boldsymbol{\xi})$ to a value p, known as chaos order. In this case,

$$\mathcal{M}(\boldsymbol{\xi}) = \sum_{i=0}^{P} c^{(i)} \Psi_i(\boldsymbol{\xi}) + \epsilon(\boldsymbol{\xi}), \tag{3}$$

where $\epsilon(\boldsymbol{\xi})$ is the truncation error, and the number of retained $P+1$ terms is given by

$$P + 1 = \frac{(p+m)!}{p!m!}. \tag{4}$$

The statistical moments of $\mathcal{M}$ can be computed algebraically from the spectral coefficients $c^{(i)}$,

$$\begin{aligned} &\text{Mean value:} \quad E[\mathcal{M}] = c^{(0)} \qquad \text{Variance:} \quad \sigma^2[\mathcal{M}] = \sum_{i=1}^{P} c^{(i)} \\ &\text{Skewness:} \quad S[\mathcal{M}] = \frac{E[\mathcal{M}^3] - 3E[\mathcal{M}]\sigma^2[\mathcal{M}] - E^3[\mathcal{M}]}{\sigma^3[\mathcal{M}]} \\ &\text{Kurtosis:} \quad K[\mathcal{M}] = \frac{E[\mathcal{M}^4] - 4E[\mathcal{M}]E[\mathcal{M}^3] + 6E^2[\mathcal{M}]\sigma^2[\mathcal{M}] + 3E^4[\mathcal{M}]}{\sigma^4[\mathcal{M}]} \end{aligned} \tag{5}$$

In this work, we use the Weighted Least Squares (WLSQ) approach to compute the coefficients $c^{(i)}$. A set of q samples of $\boldsymbol{\xi}$, $\boldsymbol{\xi}^{(i)}$ $(i = 1, \dots, q)$, are defined and the value of the QoI at each sample is computed, forming the vector $\mathbf{Q} = \left[\mathcal{M}\left(\boldsymbol{\xi}^{(1)}\right), \dots, \mathcal{M}\left(\boldsymbol{\xi}^{(q)}\right)\right]^\top \in \mathbb{R}^q$. The set of the unknown spectral coefficients $c^{(i)}$ forms the vector $\boldsymbol{c} = [c^{(0)}, \dots, c^{(P)}]^\top \in \mathbb{R}^{P+1}$, and using this notation, the truncated expansion (3) can be written as

$$\mathbf{Q} = \boldsymbol{\psi}\boldsymbol{c} + \epsilon, \tag{6}$$

where $\boldsymbol{\psi}$ is the measurement matrix with elements $\boldsymbol{\psi}_{ij} = \Psi_j\left(\boldsymbol{\xi}^{(i)}\right) \in \mathbb{R}^{q \times (P+1)}$, i.e. the i-th row contains the values of the orthogonal polynomial bases Ψ_j evaluated at the i-th sample point, $\boldsymbol{\xi}^{(i)}$ and ϵ is the vector of truncation errors. In expanded form, $\boldsymbol{\psi}$ is written as

$$\boldsymbol{\psi} = \begin{bmatrix} \Psi_0\left(\boldsymbol{\xi}^{(1)}\right) & \dots & \Psi_P\left(\boldsymbol{\xi}^{(1)}\right) \\ \vdots & \ddots & \vdots \\ \Psi_0\left(\boldsymbol{\xi}^{(q)}\right) & \dots & \Psi_P\left(\boldsymbol{\xi}^{(q)}\right) \end{bmatrix}. \tag{7}$$

The coefficient vector $\boldsymbol{c}$ is computed as a solution to the following weighted least-squares minimization problem,

$$\min_{c} \left\| \boldsymbol{W}^{\frac{1}{2}} \left(\mathbf{Q} - \boldsymbol{\psi} \boldsymbol{c} \right) \right\|_2^2 = \min_{c} \left(\mathbf{Q} - \boldsymbol{\psi} \boldsymbol{c} \right)^\top \boldsymbol{W} \left(\mathbf{Q} - \boldsymbol{\psi} \boldsymbol{c} \right), \tag{8}$$

where $\|.\|_2$ is the Euclidean norm, $\boldsymbol{W}^{\frac{1}{2}}$ is a diagonal positive-definite matrix (to be defined later) and $\boldsymbol{W} = (\boldsymbol{W}^{\frac{1}{2}})^\top \boldsymbol{W}^{\frac{1}{2}}$. The solution $\hat{\boldsymbol{c}}$ satisfies the normal equation

$$\left(\boldsymbol{\psi}^\top \boldsymbol{W} \boldsymbol{\psi} \right) \hat{\boldsymbol{c}} = \boldsymbol{\psi}^\top \boldsymbol{W} \mathbf{Q}. \tag{9}$$

The computational cost is determined by the number of samples q, assuming that the cost of evaluating the QoI far exceeds the cost of solving system (9). When random Monte Carlo sampling that follows the PDF is employed to obtain $\boldsymbol{\xi}^{(i)}$, $q \gg P+1$ is required for the system (9) to be well-conditioned. For the close connection between the LSQ approach as $q \to \infty$ and the non-intrusive Galerkin projection with Gaussian quadrature, see chapter 3 of the book [11].

The diagonal of the weighting matrix $\boldsymbol{W}^{\frac{1}{2}}$ is formed by the following weighting functions for Gaussian and uniform input distributions,

$$\begin{aligned} w_H^{\frac{1}{2}}(\boldsymbol{\xi}) &= exp(-\|\boldsymbol{\xi}^{(j)}\|_2^2/4) \text{ (Gaussian distribution)}, \\ w_L^{\frac{1}{2}}(\boldsymbol{\xi}) &= \prod_{k=1}^{m} \left(1 - {\xi_k^{(j)}}^2 \right)^{1/4} \text{ (Uniform distribution)}, \end{aligned} \tag{10}$$

where the subscripts 'H' and 'L' refer to Hermite and Legendre polynomials respectively. These analytic expressions are derived via an approach known as asymptotic sampling [19,18,41].

System (9) will be augmented with information on the sensitivity of $\mathcal{M}(\boldsymbol{\xi}^{(i)})$ with respect to $\boldsymbol{\xi}^{(i)}$; this is detailed in the next section. The sensitivities with respect to all the elements of $\boldsymbol{\xi}^{(i)}$ i.e. all random variables, can be evaluated very efficiently with a single adjoint computation, as explained in section 4.

3. Sensitivity-enhanced gPC

Taking the gradient of the truncated expansion (3) with respect to the k-th random variable at the j-th sample point, we get

$$\frac{d\mathcal{M}}{d\xi_k^{(j)}} = \sum_{i=0}^{P} c^{(i)} \frac{\partial \Psi_i}{\partial \xi_k^{(j)}} + \eta_k(\boldsymbol{\xi}^{(j)}) \quad (j = 1, \ldots, q), \tag{11}$$

where $\eta_k(\boldsymbol{\xi}^{(j)})$ is the truncation error. For each k, we obtain a block of q equations that can be written in matrix form as,

$$\frac{d\mathbf{Q}}{d\xi_k} = \frac{\partial \boldsymbol{\psi}}{\partial \xi_k} \boldsymbol{c} + \boldsymbol{\eta}_k(\boldsymbol{\xi}) \quad (k = 1, \ldots, m), \tag{12}$$

where $\frac{d\mathbf{Q}}{d\xi_k} = \left[\frac{d\mathcal{M}}{d\xi_k^{(1)}}, \ldots, \frac{d\mathcal{M}}{d\xi_k^{(q)}} \right]^\top$ and

$$\frac{\partial \boldsymbol{\psi}}{\partial \xi_k} = \begin{bmatrix} \frac{\partial \Psi_0\left(\boldsymbol{\xi}^{(1)}\right)}{\partial \xi_k^{(1)}} & \cdots & \frac{\partial \Psi_P\left(\boldsymbol{\xi}^{(1)}\right)}{\partial \xi_k^{(1)}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_0\left(\boldsymbol{\xi}^{(q)}\right)}{\partial \xi_k^{(q)}} & \cdots & \frac{\partial \Psi_P\left(\boldsymbol{\xi}^{(q)}\right)}{\partial \xi_k^{(q)}} \end{bmatrix}. \tag{13}$$

We now stack together one block of q equations of the form (6) (arising from the evaluation of $\mathcal{M}\left(\boldsymbol{\xi}^{(j)}\right)$) and m blocks, of q equations each, of the form (12) (arising from the evaluation of sensitivity of $\mathcal{M}\left(\boldsymbol{\xi}^{(j)}\right)$ at all sample points with respect to each uncertain variable). In total, there are $(1+m) \times q$ equations that can be put in matrix form as

$$\mathbf{G} = \boldsymbol{\phi} \boldsymbol{c} + \theta, \tag{14}$$

where θ is the vector of truncation errors, $\mathbf{G}$ is the block vector,

$$\mathbf{G} = \left[\mathbf{Q}, \frac{d\mathbf{Q}}{d\xi_1}, \ldots, \frac{d\mathbf{Q}}{d\xi_m}\right]^\top = \left[\mathcal{M}\left(\boldsymbol{\xi}^{(1)}\right), \ldots, \mathcal{M}\left(\boldsymbol{\xi}^{(q)}\right), \frac{d\mathcal{M}}{d\xi_1^{(1)}}, \ldots, \frac{d\mathcal{M}}{d\xi_1^{(q)}}, \ldots, \frac{d\mathcal{M}}{d\xi_m^{(1)}}, \ldots, \frac{d\mathcal{M}}{d\xi_m^{(q)}}\right]^\top \in \mathbb{R}^{(1+m)q\times 1}, \tag{15}$$

and $\boldsymbol{\phi}$ is the row-block matrix

$$\boldsymbol{\phi} = \begin{bmatrix} \boldsymbol{\psi} \\ \frac{\partial \boldsymbol{\psi}}{\partial \xi_1} \\ \vdots \\ \frac{\partial \boldsymbol{\psi}}{\partial \xi_m} \end{bmatrix} = \begin{bmatrix} \Psi_0\left(\boldsymbol{\xi}^{(1)}\right) & \ldots & \Psi_P\left(\boldsymbol{\xi}^{(1)}\right) \\ \vdots & \ddots & \vdots \\ \Psi_0\left(\boldsymbol{\xi}^{(q)}\right) & \ldots & \Psi_P\left(\boldsymbol{\xi}^{(q)}\right) \\ \frac{\partial \Psi_0\left(\boldsymbol{\xi}^{(1)}\right)}{\partial \xi_1^{(1)}} & \cdots & \frac{\partial \Psi_P\left(\boldsymbol{\xi}^{(1)}\right)}{\partial \xi_1^{(1)}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_0\left(\boldsymbol{\xi}^{(q)}\right)}{\partial \xi_1^{(q)}} & \cdots & \frac{\partial \Psi_P\left(\boldsymbol{\xi}^{(q)}\right)}{\partial \xi_1^{(q)}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_0\left(\boldsymbol{\xi}^{(1)}\right)}{\partial \xi_m^{(1)}} & \cdots & \frac{\partial \Psi_P\left(\boldsymbol{\xi}^{(1)}\right)}{\partial \xi_m^{(1)}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_0\left(\boldsymbol{\xi}^{(q)}\right)}{\partial \xi_m^{(q)}} & \cdots & \frac{\partial \Psi_P\left(\boldsymbol{\xi}^{(q)}\right)}{\partial \xi_m^{(q)}} \end{bmatrix} \in \mathbb{R}^{(1+m)q\times(P+1)}. \tag{16}$$

We now define the following weighted least squares minimisation problem

$$\min_{\boldsymbol{c}} \| \boldsymbol{W'}^{\frac{1}{2}} (\mathbf{G} - \boldsymbol{\phi}\boldsymbol{c}) \|_2^2 = \min_{\boldsymbol{c}} (\mathbf{G} - \boldsymbol{\phi}\boldsymbol{c})^\top \boldsymbol{W'} (\mathbf{G} - \boldsymbol{\phi}\boldsymbol{c}), \tag{17}$$

where $\boldsymbol{W'}$ is a block diagonal weighting matrix, consisting of $1+m$ blocks of $\boldsymbol{W}$ i.e. $\boldsymbol{W'} = diag(\underbrace{\boldsymbol{W}, \ldots, \boldsymbol{W}}_{1+m \text{ blocks}})$. The above problem can be written in expanded form as

$$\min_{\boldsymbol{c}} (\mathbf{Q} - \boldsymbol{\psi}\boldsymbol{c})^\top \boldsymbol{W} (\mathbf{Q} - \boldsymbol{\psi}\boldsymbol{c}) + \sum_{k=1}^{m} \left(\frac{d\mathbf{Q}}{d\xi_k} - \frac{\partial \boldsymbol{\psi}}{\partial \xi_k}\boldsymbol{c}\right)^\top \boldsymbol{W} \left(\frac{d\mathbf{Q}}{d\xi_k} - \frac{\partial \boldsymbol{\psi}}{\partial \xi_k}\boldsymbol{c}\right), \tag{18}$$

which is a generalisation of the minimisation problem (8). For simplicity, we have used the same weights for the equations of the QoI value and its derivatives, (3) and (11) respectively, but this is not necessary. Depending on the case, these quantities can differ by orders of magnitude and different weights may be warranted, in line with standard multi-objective LSQ approaches.

The vector $\hat{\boldsymbol{c}}$ that minimises the above function is obtained from the solution of the normal set of equations

$$\left(\boldsymbol{\phi}^\top \boldsymbol{W'} \boldsymbol{\phi}\right) \hat{\boldsymbol{c}} = \boldsymbol{\phi}^\top \boldsymbol{W'} \mathbf{G}, \tag{19}$$

which is analogous to system (9).

4. Efficient computation of sensitivities with an adjoint formulation

We now consider the computation of the sensitivities $\frac{d\mathcal{M}}{d\xi_k^{(j)}}$ $(k = 1 \ldots m)$ at the j-th sampling point. We assume systems that are described by a set of partial differential equations that after spatial discretisation take the form

$$\boldsymbol{R}(\boldsymbol{u}, \boldsymbol{\xi}) = 0, \tag{20}$$

where $\boldsymbol{u} \in \mathbb{R}^{N_{\boldsymbol{u}}}$ is the state vector and $\boldsymbol{R} : \mathbb{R}^{N_{\boldsymbol{u}}} \times \mathbb{R}^m \to \mathbb{R}^{N_{\boldsymbol{u}}}$ varies smoothly with $\boldsymbol{u}$ and $\boldsymbol{\xi}$. For simplicity, we suppose that (20) describes mathematically a steady problem, i.e. $\boldsymbol{u}$ does not depend on time, but the analysis can be easily extended to unsteady problems as well. The QoI can therefore be written as

$$\mathcal{M} = \mathcal{M}(\boldsymbol{u}(\boldsymbol{\xi}), \boldsymbol{\xi}) : \mathbb{R}^{N_{\boldsymbol{u}}} \times \mathbb{R}^m \to \mathbb{R}, \tag{21}$$

where $\boldsymbol{u}(\boldsymbol{\xi})$ is the solution of (20) for one realisation of $\boldsymbol{\xi}$.

Using the adjoint formulation, the sensitivities $\frac{d\mathcal{M}}{d\xi_k^{(j)}}$ $(k = 1 \dots m)$ at each sampling point j can be computed at a cost which is independent of the dimension m of the vector $\boldsymbol{\xi}$. Below we outline the basic steps of the discrete adjoint approach, more details can be found in [42] and the review paper [21]. It is also possible to derive a continuous adjoint formulation, but this requires to use the continuous form of the equations from which the discrete system (20) is derived. The discrete formulation suffices for the purposes of this section. We employ the continuous formulation in one of the test cases of section 6.

We start by defining the Lagrangian $\mathcal{L} : \mathbb{R}^{N_{\boldsymbol{u}}} \times \mathbb{R}^m \to \mathbb{R}$

$$\mathcal{L}(\boldsymbol{u}(\boldsymbol{\xi}), \boldsymbol{\xi}) = \mathcal{M}(\boldsymbol{u}(\boldsymbol{\xi}), \boldsymbol{\xi}) + \boldsymbol{\lambda}^\top \boldsymbol{R}(\boldsymbol{u}(\boldsymbol{\xi}), \boldsymbol{\xi}), \tag{22}$$

where $\boldsymbol{\lambda} \in \mathbb{R}^{N_{\boldsymbol{u}}}$ are known as Lagrange multipliers (or adjoint variables), and $\boldsymbol{u}(\boldsymbol{\xi})$ is the solution of (20) at the j-th sampling point, $\boldsymbol{\xi}^{(j)} = [\xi_1^{(j)}, \dots, \xi_m^{(j)}]$. Differentiating $\mathcal{L}(\boldsymbol{u}(\boldsymbol{\xi}), \boldsymbol{\xi})$ with respect to $\boldsymbol{\xi}$ yields

$$\frac{d\mathcal{L}}{d\boldsymbol{\xi}} = \left(\frac{\partial \mathcal{M}}{\partial \boldsymbol{u}}\right)^\top \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{\xi}} + \frac{\partial \mathcal{M}}{\partial \boldsymbol{\xi}} + \boldsymbol{\lambda}^\top \left(\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{u}} \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{\xi}} + \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{\xi}}\right), \tag{23}$$

where $\frac{d\mathcal{L}}{d\boldsymbol{\xi}} = \left[\frac{d\mathcal{L}}{d\xi_1^{(j)}} \dots \frac{d\mathcal{L}}{d\xi_m^{(j)}}\right]^\top$ and $\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{u}}$ is the Jacobian matrix. The above expression can be re-arranged as,

$$\frac{d\mathcal{L}}{d\boldsymbol{\xi}} = \left[\left(\frac{\partial \mathcal{M}}{\partial \boldsymbol{u}}\right)^\top + \boldsymbol{\lambda}^\top \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{u}}\right] \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{\xi}} + \boldsymbol{\lambda}^\top \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{\xi}} + \frac{\partial \mathcal{M}}{\partial \boldsymbol{\xi}}. \tag{24}$$

The Lagrange multipliers are chosen so that $\frac{d\mathcal{L}}{d\boldsymbol{\xi}}$ becomes independent of variations of $\boldsymbol{u}$ with respect to $\boldsymbol{\xi}$. This is achieved by setting the term within square brackets equal to $\mathbf{0}$, resulting in the Field Adjoint Equation (FAE)

$$\left(\frac{\partial \mathcal{M}}{\partial \boldsymbol{u}}\right)^\top + \boldsymbol{\lambda}^\top \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{u}} = \mathbf{0} \Longrightarrow \left(\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{u}}\right)^\top \boldsymbol{\lambda} = -\frac{\partial \mathcal{M}}{\partial \boldsymbol{u}}. \tag{25}$$

Solving (25), the Lagrangian multiplies $\boldsymbol{\lambda}$ are obtained and the sensitivities can be computed from

$$\frac{d\mathcal{M}}{d\boldsymbol{\xi}} = \frac{d\mathcal{L}}{d\boldsymbol{\xi}} = \boldsymbol{\lambda}^\top \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{\xi}} + \frac{\partial \mathcal{M}}{\partial \boldsymbol{\xi}}, \tag{26}$$

where $\frac{d\mathcal{M}}{d\boldsymbol{\xi}} = \left[\frac{d\mathcal{M}}{d\xi_1^{(j)}} \dots \frac{d\mathcal{M}}{d\xi_m^{(j)}}\right]^\top$.

The state and adjoint variables, $\boldsymbol{u}$ and $\boldsymbol{\lambda}$ respectively, have the same dimension (equal to N_u) and the solution algorithms for the two sets of equations, (20) and (25), are similar. In the following, we will assume that the computational cost is also the same. Note that in practise the cost of solving (25) is lower compared to that of (20), because the former set is linear, so no outer iterations to deal with the non-linearity are necessary. The cost of (26) is considered negligible.

To summarise, the solution of equation (25) at each sampling point provides a set of m extra conditions at the cost of one additional evaluation. For the lowest chaos order $p = 1$, the number of coefficients, see equation (4), is equal to $P + 1 = \frac{(1+m)!}{1!m!} = 1 + m$. One direct and one adjoint evaluation can yield $1 + m$ equations at a single sampling point, thus making the computational cost of the se-gPC independent from the dimension of the stochastic space m.

For the standard LSQ method, if q sampling points are needed for a well conditioned least-squares system (9), q equations are obtained at a cost of q evaluations. For the same number of sampling points, the se-gPC will yield $(1 + m)q$ equations, at a cost of $2q$ evaluations. If the number of equations is kept constant and equal to q, then se-gPC allows one to use a smaller number of sampling points, equal to $q/(1 + m)$, with cost $2q/(1 + m)$ evaluations. For cases with a large number of uncertain parameters, this is a very significant reduction in computational cost.

If the computational cost is kept constant (for example due to a fixed computational budget) and equal to q evaluations, the number of sampling points with se-gPC will be $q/2$ and the number of equations $(1 + m)q/2$.

The minimum number of equations needed to compute the spectral coefficients is equal to $P + 1$ (this corresponds to oversampling ratio equal to 1), thus $P + 1$ and $(P + 1)/(1 + m)$ sampling points are required for the standard gPC and se-gPC respectively. Of course, one can use larger oversampling ratios. In the present paper we are using the smallest number of sampling points.

In Fig. 1, a comparison of the number of evaluations required by Smolyak as well as weighted Least Squares (WLSQ) and se-PCE (the latter two with the minimum number of sampling points) is made. Results are shown for two values of chaos order $p = 1, 2$. For $p = 1$ (left panel), the number of evaluations required by the se-gPC is constant and equal to 2, i.e. independent of the number of stochastic parameters, as mentioned earlier. For the other two methods, it scales linearly with m; for WLSQ it is equal to $m + 1$, and for Smolyak $2m + 1$. For $p = 2$ (right panel), the number of evaluations scales quadratically with m for WLSQ and Smolyak (equal to $(m + 1)(m + 2)/2$ and $(m + 1)(2m + 1)$ respectively), but linearly for se-PCE (equal to $(m + 2)$).

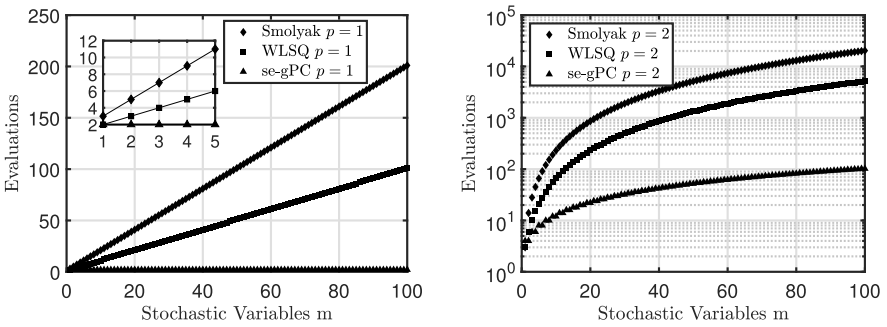


Fig. 1. Comparison of evaluations required by Smolyak grids, WLSQ and se-gPC at the QR points (see section 5), for $p=1$ (left) and $p=2$ (right).

Since in the paper the smallest number of sampling points is employed, their selection needs to be done carefully; this is examined in the next section.

5. Sampling of the stochastic space

There are many different algorithms to sample the stochastic space for least squares PCE, for an overview see [19]. In this paper we apply the pivoted QR decomposition algorithm, which has been used for sparse sensor placement for flow or image reconstruction [43], or to find optimal evaluation points and weights for integration [39]. This is a greedy algorithm, that maximises the determinant of an appropriate matrix; in this sense it is a D-optimal design method. For the application of this algorithm for sampling the stochastic space, see [37,43]. Below we summarise the central idea and implementation.

Consider again the standard least-squares problem, equation (6), repeated below for convenience,

$$\mathbf{Q} = \boldsymbol{\psi}\boldsymbol{c} + \boldsymbol{\epsilon}, \tag{27}$$

and assume that the elements of the residual vector $\boldsymbol{\epsilon}$ are independent, zero mean Gaussian random variables with the same standard deviation, i.e. for the i-th element $\epsilon_i \sim N(0,\eta^2)$, with $i=1,\dots,q$. Assuming for the time being that the weighting matrix is identity, $\boldsymbol{W}=\boldsymbol{I}_q$, the solution $\hat{\boldsymbol{c}}$ of the least squares system satisfies $\left(\boldsymbol{\psi}^\top\boldsymbol{\psi}\right)\hat{\boldsymbol{c}} = \boldsymbol{\psi}^\top\mathbf{Q}$, see (9). The covariance matrix of the difference $\hat{\boldsymbol{c}}-\boldsymbol{c}$ is given by

$$Var(\hat{\boldsymbol{c}}-\boldsymbol{c}) = \eta^2\left(\boldsymbol{\psi}^\top\boldsymbol{\psi}\right)^{-1} \tag{28}$$

This can be easily proved,

$$\begin{aligned}\hat{\boldsymbol{c}}-\boldsymbol{c} &= (\boldsymbol{\psi}^\top\boldsymbol{\psi})^{-1}\boldsymbol{\psi}^\top\mathbf{Q}-\boldsymbol{c} = (\boldsymbol{\psi}^\top\boldsymbol{\psi})^{-1}\boldsymbol{\psi}^\top(\boldsymbol{\psi}\boldsymbol{c}+\boldsymbol{\epsilon})-\boldsymbol{c}\\ &= (\boldsymbol{\psi}^\top\boldsymbol{\psi})^{-1}\boldsymbol{\psi}^\top\boldsymbol{\psi}\boldsymbol{c}+(\boldsymbol{\psi}\boldsymbol{\psi})^{-1}\boldsymbol{\psi}^\top\boldsymbol{\epsilon}-\boldsymbol{c}\\ &= \boldsymbol{c}+(\boldsymbol{\psi}^\top\boldsymbol{\psi})^{-1}\boldsymbol{\psi}^\top\boldsymbol{\epsilon}-\boldsymbol{c} = (\boldsymbol{\psi}^\top\boldsymbol{\psi})^{-1}\boldsymbol{\psi}^\top\boldsymbol{\epsilon}\end{aligned} \tag{29}$$

Setting $\boldsymbol{A}=(\boldsymbol{\psi}^\top\boldsymbol{\psi})^{-1}\boldsymbol{\psi}^\top$, we get

$$\begin{aligned}&Var(\hat{\boldsymbol{c}}-\boldsymbol{c}) = \mathcal{E}\left[(\hat{\boldsymbol{c}}-\boldsymbol{c})(\hat{\boldsymbol{c}}-\boldsymbol{c})^\top\right] = \mathcal{E}\left[(\boldsymbol{A}\boldsymbol{\epsilon})\,(\boldsymbol{A}\boldsymbol{\epsilon})^\top\right] = \mathcal{E}\left[\boldsymbol{A}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top)\boldsymbol{A}^\top\right] =\\ &\boldsymbol{A}\mathcal{E}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top)\boldsymbol{A}^\top = \boldsymbol{A}\eta^2\boldsymbol{I}_q\boldsymbol{A}^\top = \eta^2\boldsymbol{A}\boldsymbol{A}^\top = \eta^2(\boldsymbol{\psi}^\top\boldsymbol{\psi})^{-1},\end{aligned} \tag{30}$$

where $\mathcal{E}(.)$ is the expectation operation. If we include the weighting matrix and follow the same steps, the covariance matrix becomes

$$Var(\hat{\boldsymbol{c}}-\boldsymbol{c}) = \eta^2\left(\boldsymbol{\psi}^\top\boldsymbol{W}\boldsymbol{\psi}\right)^{-1}. \tag{31}$$

The measurement matrix $\boldsymbol{\psi}$ contains information for all q sampling points, see (7). The objective of the D-optimal design is to select the sampling points that minimise the determinant of the matrix $\left(\boldsymbol{\psi}^\top\boldsymbol{\psi}\right)^{-1}$ or $\left(\boldsymbol{\psi}^\top\boldsymbol{W}\boldsymbol{\psi}\right)^{-1}$. Here we follow a greedy approach, i.e. from a large pool of q candidate sample points, we select the required number of points (equal to $P+1$ for the standard least squares method) that maximise the determinant. The process is as follows.

To select only a subset of the q points, we multiply equation (27) with the row selection matrix $\boldsymbol{P}\in\mathbb{R}^{(P+1)\times q}$. At each row of $\boldsymbol{P}$ all elements are 0, except the element at the column that corresponds to the selected sampling point, which takes the value of 1. Thus (27) becomes,

$$\mathbf{Q}_{\boldsymbol{P}} = \boldsymbol{P}\boldsymbol{\psi}\boldsymbol{c}+\boldsymbol{\epsilon}_{\boldsymbol{P}}, \tag{32}$$

where $\mathbf{Q_P} = \boldsymbol{P}\mathbf{Q}$ now contains only the values of the QoI $\mathcal{M}$ at the selected $P+1$ sampling points. The covariance matrix is

$$Var(\hat{\boldsymbol{c}} - \boldsymbol{c}) = \eta^2 \left[(\boldsymbol{P}\boldsymbol{\psi})^\top (\boldsymbol{P}\boldsymbol{\psi}) \right]^{-1}, \tag{33}$$

and if we include the weighting matrix, it becomes

$$Var(\hat{\boldsymbol{c}} - \boldsymbol{c}) = \eta^2 \left[\left(\boldsymbol{P}\boldsymbol{W}^{1/2}\boldsymbol{\psi} \right)^\top \left(\boldsymbol{P}\boldsymbol{W}^{1/2}\boldsymbol{\psi} \right) \right]^{-1} \tag{34}$$

The aim now is to choose the row-selection matrix $\boldsymbol{P}$ that minimizes the determinant of $\left[\left(\boldsymbol{P}\boldsymbol{W}^{1/2}\boldsymbol{\psi} \right)^\top \left(\boldsymbol{P}\boldsymbol{W}^{1/2}\boldsymbol{\psi} \right) \right]^{-1}$ or equivalently maximise the determinant of the inverse, i.e.

$$\mathcal{P} = \underset{\boldsymbol{P}}{\operatorname{argmax}}\, det \left[\left(\boldsymbol{P}\boldsymbol{W}^{1/2}\boldsymbol{\psi} \right)^\top \left(\boldsymbol{P}\boldsymbol{W}^{1/2}\boldsymbol{\psi} \right) \right] \tag{35}$$

The matrix $\boldsymbol{P}\boldsymbol{W}^{1/2}\boldsymbol{\psi} \in \mathbb{R}^{(P+1)\times(P+1)}$ is square, thus $det \left[\left(\boldsymbol{P}\boldsymbol{W}^{1/2}\boldsymbol{\psi} \right)^\top \left(\boldsymbol{P}\boldsymbol{W}^{1/2}\boldsymbol{\psi} \right) \right] = \left[det \left(\boldsymbol{P}\boldsymbol{W}^{1/2}\boldsymbol{\psi} \right) \right]^2$, and the maximisation problem (35) becomes

$$\mathcal{P} = \underset{\boldsymbol{P}}{\operatorname{argmax}} \left| det \left(\boldsymbol{P}\boldsymbol{W}^{1/2}\boldsymbol{\psi} \right) \right| \tag{36}$$

The solution to the above maximisation problem is given by the pivoted QR decomposition

$$\left(\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\psi} \right)^\top \mathcal{P} = \boldsymbol{Q}\,\boldsymbol{R}. \tag{37}$$

There are very robust subroutines to perform this decomposition; in this paper we have used the 'qr' command of MATLAB. The permutation matrix $\mathcal{P}$ is chosen so that the diagonal elements R_{ii} of the upper diagonal matrix $\boldsymbol{R}$ are ordered in descending order, i.e.

$$|R_{11}| \geq |R_{22}| \geq \cdots \geq |R_{(P+1)(P+1)}| \tag{38}$$

The absolute value of the determinant is the product of the diagonal entries [43],

$$\left| det \left(\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\psi} \right) \right| = \prod_{i=1}^{P+1} |R_{ii}| \tag{39}$$

We apply pivoted QR decomposition to the matrix $\left(\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\psi} \right)^\top$ of the standard weighted LSQ approach, as well as the augmented matrix $\left(\boldsymbol{W'}^{\frac{1}{2}}\boldsymbol{\phi} \right)^\top$ of the se-gPC approach; see (16) for the expanded form of $\boldsymbol{\phi}$. In the former strategy, we identify the $P+1$ points required by the standard LSQ approach, and then select the most important $(P+1)/(m+1)$ points (i.e. the ones with the highest diagonal elements $|R_{ii}|$) and apply on those the se-gPC. The second strategy will provide directly the $P+1$ points where either the QoI or one of its derivatives will be computed. Below we compare the two strategies for the case of two independent random variables.

5.1. Sampling for two stochastic variables

We consider the case of $m = 2$ uncertain input variables, ξ_1, ξ_2, both following standard normal distribution, $N(0,1)$. We selected chaos order $p = 4$, that results in $P+1 = 15$ unknowns. A pool of 10000 random sample points were generated with Monte-Carlo from the input PDF. Note that this is insufficient when employing high degree PCE polynomials; in this case it's better to sample from an asymptotic measure, see [19,18,41] for the theoretical explanation.

We first apply pivoted QR decomposition to matrix $\left(\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\psi} \right)^\top$. The absolute values of the diagonal elements of the matrix $\boldsymbol{R}$, $|R_{ii}|$, are shown in Fig. 2a. The elements are tugged with their index number, $i = 1 \ldots 15$. On the right panel, Fig. 2b, the locations of the selected 15 sampling points are shown, also tagged with the corresponding index i. The points are arranged in three concentric circles. The inner circle contains only one point; this point is located at the mean input, i.e. $\xi_1^{(1)} = 0$, $\xi_2^{(1)} = 0$, and corresponds to the highest ranked element $|R_{11}|$. The next circle contains 5 points, corresponding to elements $|R_{22}| \ldots |R_{66}|$; the points are located at an average radius $\sim 1.47\sigma$. The outer circle contains 8 points, corresponding to elements $|R_{77}| \ldots |R_{15}|$, and has a radius of $\sim 2.77\sigma$.

We repeated the process for chaos order $p = 2$ and the same input variables; now there are $P+1 = 6$ unknowns. The joint PDF and the selected QR sampling points are shown in Fig. 3. Again the first point is placed at the peak of the PDF, i.e. $\xi_1^{(1)} = 0$, $\xi_2^{(1)} = 0$, and 5 points are symmetrically placed at a radius $\sim 1.75\sigma$.

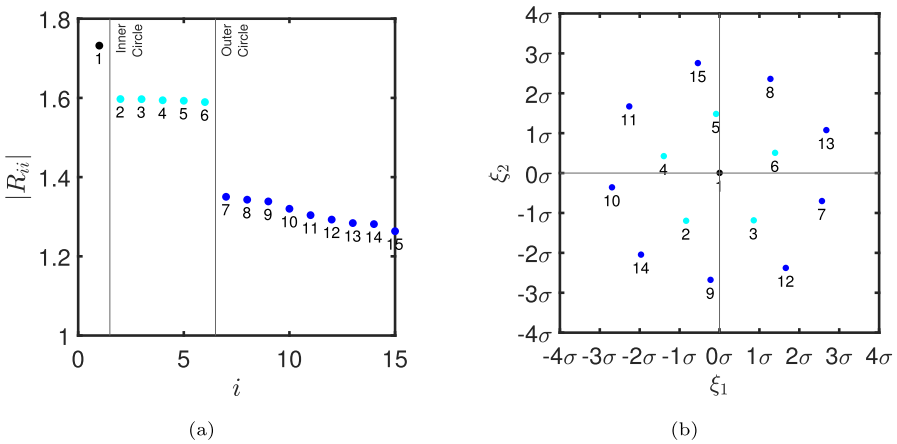


Fig. 2. (a) Absolute values of the diagonal elements $|R_{ii}|$. (b) Selected QR sample points of matrix $\left(\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\psi}\right)^{\top}$. Results for chaos order $p=4$. The stochastic input is $\boldsymbol{\xi}\in\mathbb{R}^2$ with $\xi_1,\xi_2\sim N(0,1)$ sampled at 10000 random points.

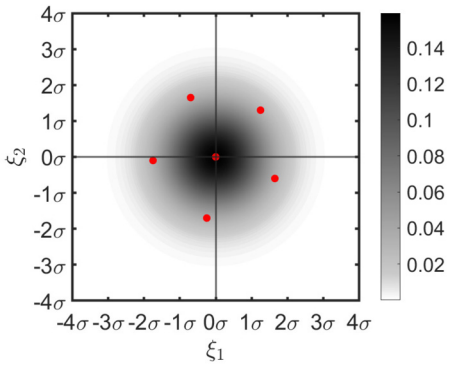


Fig. 3. Locations of QR sampling points of matrix $\left(\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\psi}\right)^{\top}$ for chaos order $p=2$, and two independent variables following standard normal distribution, $\xi_1,\xi_2\sim N(0,1)$. The contour of the joint PDF is superimposed.

Table 1
Parameters of a numerical experiment to compute the histogram of CN of matrix $\left(\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\psi}\right)^{\top}$. The experiment was repeated 100 times and the average CN evaluated at the QR sampling points is reported in the last column. The percentage of samples from all permutations that have lower CN is also reported within brackets.

p	$P+1$	q	Number of permutations	$CN\left[\text{QR of}\left(W^{\frac{1}{2}}\psi\right)^{\top}\right]$
1	3	1000	166167000	2.2 (0.24%)
2	6	50	15890700	16.5 (0.05%)

We now proceed to compute the condition number (CN) of the matrix $\left(\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\psi}\right)^{\top}$. Note that the matrix appearing in (9) has CN which is the square of that of $\left(\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\psi}\right)^{\top}$. For $m=2$, and two chaos orders $p=1$ and $p=2$, all possible permutations of $P+1$ samples out of a pool of q candidate samples (equal to $\frac{q!}{(q-P-1)!(P+1)!}$), were found and the $CN\left(\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\psi}\right)^{\top}$ was computed. We performed each numerical experiment 100 times; the averaged results are shown in Table 1. For example, for $p=1$, $P+1=3$ and $q=1000$, there are 166167000 permutations. The QR sampling points result in a system with condition number of 2.2; only 0.24% of all permutations have smaller CN. For $p=2$, the percentage is even smaller, 0.05%.

The histogram of the CN for all permutations is shown in Fig. 4. Note that the horizontal axis is in logarithmic scale due to the very large spread of the CN values. For $p=1$ (left panel), the median of the PDF is ~ 43, i.e. 20 times larger. For $p=2$ (right panel) the PDF has similar shape, and the median CN is 2977 i.e. 180 times higher than the one for the QR points.

We now proceed to apply pivoted QR decomposition to the augmented matrix $\left(\boldsymbol{W'}^{\frac{1}{2}}\boldsymbol{\phi}\right)^{\top}$. The QR sampling points are shown in Fig. 5b. We have used different symbols to denote the locations where the QoI or its derivatives in the ξ_1 or ξ_2

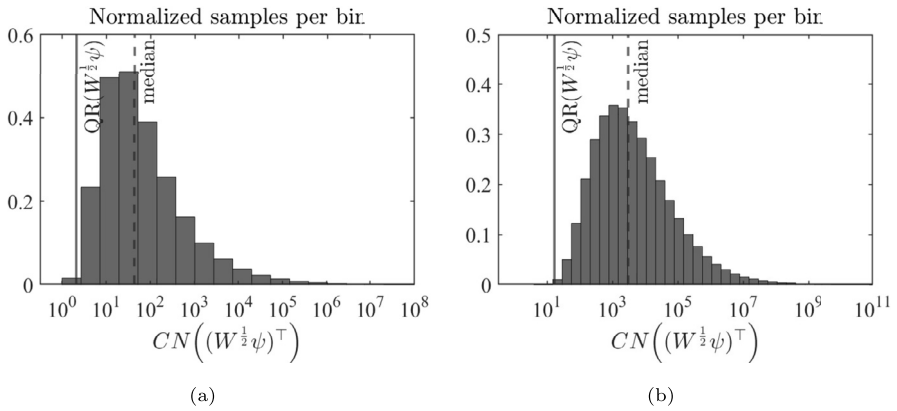


Fig. 4. Histogram of the CN (normalised samples per bin) of matrix $\left(\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\psi}\right)^\top$ for all permutations for $m=2$ and (a) $p=1$ and (b) $p=2$. The vertical solid and dashed lines mark the CN at the QR sampling points and the median CN respectively.

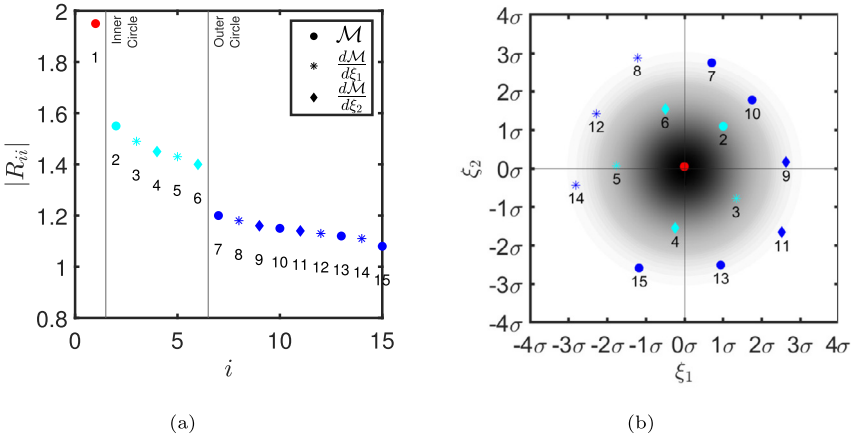


Fig. 5. (a) Absolute values of the diagonal elements $|R_{ii}|$. Circles, asterisks and rhombi denote the points where $\mathcal{M}$, $\frac{d\mathcal{M}}{d\xi_1}$ and $\frac{d\mathcal{M}}{d\xi_2}$ are to be evaluated respectively, see figure legend. (b) Locations of QR sampling points of matrix $\left(\boldsymbol{W'}^{\frac{1}{2}}\boldsymbol{\phi}\right)^\top$. The contour of the joint PDF is superimposed. Results for chaos order $p=4$, and two independent variables following standard normal distribution, $\xi_1, \xi_2 \sim N(0,1)$. The stochastic space is sampled at 10000 random points.

directions must be computed. It is interesting to notice that again the evaluation points form co-centric circles which are similar, but not identical, to those shown in Fig. 2b. In Fig. 5a, the distribution of the diagonal elements $|R_{ii}|$ is shown. Again the values are clustered in groups, with each group corresponding to one circle in panel (b). Compared to Fig. 2a, there is larger variation in the values within each group, most probably because they refer to different quantities (some values refer to the QoI and some to its derivative along ξ_1 or ξ_2).

The CN histograms of matrix $\left(\boldsymbol{W'}^{\frac{1}{2}}\boldsymbol{\phi}\right)^\top$ for $p=1,2$ are shown in Fig. 6. The relevant parameters are provided in Table 2. Now at every location there are the 3 random samples, $\mathcal{M}$, $\frac{d\mathcal{M}}{d\xi_1}$ and $\frac{d\mathcal{M}}{d\xi_2}$. Thus, in order to keep the pool sample size q similar to Table 1, we reduce the number of physical locations to 333 and 16 for $p=1$ and $p=2$ respectively (the corresponding values are $q=999$ and $q=48$, see 3rd column of Table 2). From this pool, we randomly select $P+1$ samples (can be the QoI or any of its derivatives) and compute the histograms of CN of $\left(\boldsymbol{W'}^{\frac{1}{2}}\boldsymbol{\phi}\right)^\top$ from all permutations. The CN evaluated at the QR points of $\left(\boldsymbol{W'}^{\frac{1}{2}}\boldsymbol{\phi}\right)^\top$ is 1.4 for $p=1$ and 11.6 for $p=2$ (only 0.31% and 0.12% of the samples have lower CN, see Table 2). Interestingly, the range of CN values is similar to Fig. 4, most likely because of low chaos order and number of variables.

Most importantly, we also compute the value of the CN of matrix $\left(\boldsymbol{W'}^{\frac{1}{2}}\boldsymbol{\phi}\right)^\top$ by ranking the $P+1$ QR points of $\left(\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\psi}\right)^\top$, selecting the most important $(P+1)/(m+1)$ ones, and evaluating at these points $\mathcal{M}$, $\frac{d\mathcal{M}}{d\xi_1}$ and $\frac{d\mathcal{M}}{d\xi_2}$. As can be seen, the resulting CN is slightly larger as expected, but still only 1.42% and 0.96% of the samples have lower CN. This is a very favourable outcome because it allows for the maximum possible exploitation of the adjoint information, in the sense that

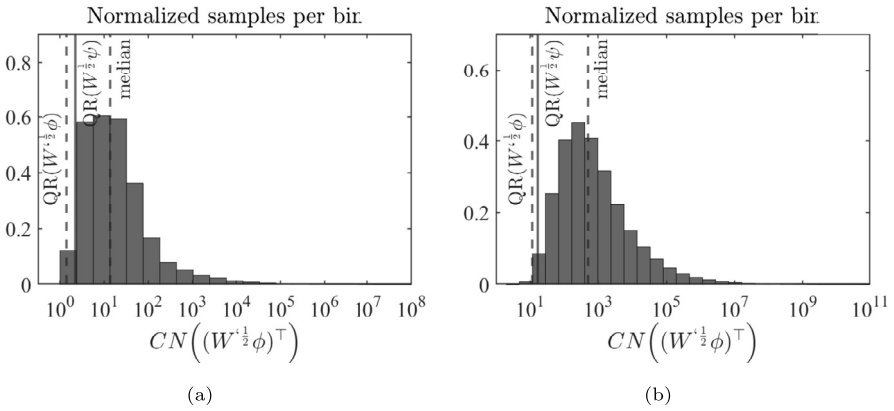


Fig. 6. Histogram of CN (normalised samples per bin) of matrix $\left(\boldsymbol{W'}^{\frac{1}{2}}\boldsymbol{\phi}\right)^\top$ for all permutations for $m=2$ and (a) $p=1$ and (b) $p=2$. Vertical lines denote the median CN and the CN from the QR sampling points of $\left(\boldsymbol{W'}^{\frac{1}{2}}\boldsymbol{\phi}\right)^\top$ and $\left(\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\psi}\right)^\top$ (each line is marked accordingly).

Table 2
Parameters of a numerical experiment to compute the histogram of CN of matrix $\left(\boldsymbol{W'}^{\frac{1}{2}}\boldsymbol{\phi}\right)^\top$. The experiment was repeated 100 times and the average CN evaluated at the QR sampling points of $\left(\boldsymbol{W'}^{\frac{1}{2}}\boldsymbol{\phi}\right)^\top$ and $\left(\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{\psi}\right)^\top$ are reported in the penultimate and last columns respectively (see text for more details). The percentage of samples that have lower CN are reported within brackets.

p	$P+1$	q	Number of permutations	CN $\left[\text{QR of}\left(W'^{\frac{1}{2}}\phi\right)^\top\right]$	CN $\left[\text{QR of}\left(W^{\frac{1}{2}}\psi\right)^\top\right]$
1	3	999 (=3×333)	165668499	1.4 (0.31%)	2.28 (1.42%)
2	6	48 (=3×16)	12271512	11.6 (0.12%)	21.7 (0.96%)

all derivatives obtained at one sample point are employed in the augmented LSQ problem. This is the sampling strategy that we follow in the rest of the paper. When QR is applied directly to matrix $\left(\boldsymbol{W'}^{\frac{1}{2}}\boldsymbol{\phi}\right)^\top$, only very limited information is utilised at each sampling point, either the QoI value or one of its derivatives in a particular direction.

In the following section, we apply standard WLSQ, se-PCE and Smolyak quadrature in several test cases and compare the three methods in terms of accuracy and computational cost. For the Smolyak quadrature, we have used Gauss-Legendre 1D rules. The index set was determined by the multi-index algorithm of [16] and the growth rule was linear, $2l+1$, where l is the 1D level. We also report the values of CN of the augmented matrix $\boldsymbol{\phi}^\top\boldsymbol{W'}\boldsymbol{\phi}$; this is another performance metric that becomes important especially for problems with large m. As mentioned in the introduction, very large CN values were reported in [27] for a case with $m=12$, $p=3$ and random selection of samples.

6. Application to test cases

6.1. Ishigami function

The Ishigami function [44],

$$Y = sinX_1 + \alpha sin^2 X_2 + \beta X_3^4 sinX_1, \tag{40}$$

is a commonly used benchmark case for UQ and Global Sensitivity Analysis [45–48]. It has three stochastic variables X_1, X_2 and X_3 (so $m=3$) that follow a uniform distribution in the interval $[-\pi,\pi]$. This is a challenging case because the QoI, Y, is a strongly nonlinear and non-monotonic function, thus requires a high spectral order for accurate results; it is therefore an excellent case to assess the performance of se-gPC.

The statistical moments and sensitivity measures of Y can be computed analytically. For $\alpha=7$ and $\beta=0.1$, the reference values for the first four moments are $\widehat{\mu_Y}=3.5000$, $\widehat{\sigma_Y}=3.7208$, $\widehat{\delta_Y}=0$ and $\widehat{\kappa_Y}=3.5072$ [49]. We computed the moments with Smolyak Quadrature as well as WLSQ and se-gPC at the minimum number of QR points and evaluated the error with respect to the reference values. For se-gPC, the sensitivities of QoI with respect to the 3 stochastic variables at the selected sampling points were computed analytically.

The results are plotted against the number of evaluations in Fig. 7. Each symbol in the plot represents a different chaos order $p=1,\ldots,10$ (no symbols are shown for skewness and kurtosis for $p=1$ and $p=1,2$ respectively, because these moments cannot be computed for these values of p). As can be seen, all methods converge to very small errors. An interesting observation can be made from these plots. For small values of p, Smolyak quadrature can give slightly better

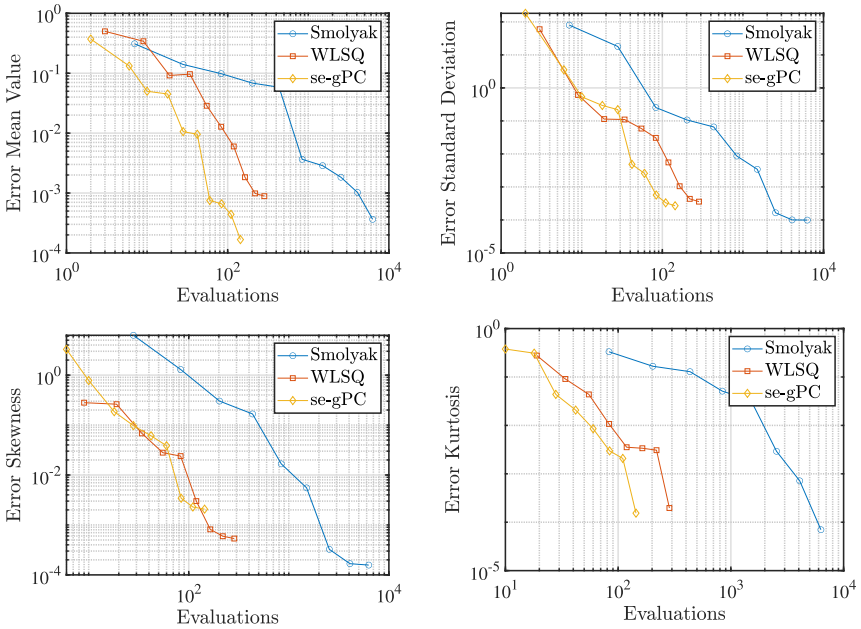


Fig. 7. Convergence of the first four moments the Ishigami function (for $\alpha = 7$ and $\beta = 0.1$) against the number of evaluations, as computed with Smolyak Quadrature, WLSQ and se-GPC. Each symbol represents a chaos order $p = 1, \ldots, 10$. Symbols are shown only for values of p that allow the computation of the moments.

Table 3
Sobol indices comparison as computed by the gPC of the Ishigami function for $\alpha = 7$ and $\beta = 0.1$. The analytical values are obtained from [49].

Sobol index	Analytic value	se-gPC ($p = 6$)	se-gPC ($p = 10$)
SU_1	0.5574	0.5558	0.5567
SU_2	0.4424	0.4442	0.4433
SU_3	0.2436	0.2448	0.2449

results than se-gPC, see for example mean and standard deviation plots (top row). This is because Smolyak samples the QoI at more points, while for $p = 1$ se-gPC samples the QoI and the sensitivities at just the first QR point. If high chaos order is needed for a challenging QoI like Ishigami, information at a single point is not sufficient to produce accurate results. However, at larger p values se-gPC samples the QoI and its sensitivities from more points, and this leads to rapidly improved convergence, as seen from the mean value and standard deviation plots. For the skewness and kurtosis, the superiority of se-gPC is not always evident compared to WLSQ, but both methods are more accurate compared to Smolyak quadrature for the same number of evaluations. Note that in this case, even though the derivatives of the Ishigami function are computed analytically at negligible computational cost, the evaluations of the se-gPC are still counted as if an adjoint evaluation was performed (i.e. at each sampling point the cost is two evaluations).

The Ishigami function is also commonly used for the evaluation of global sensitivity indices, like the Sobol indices [49], which quantify the effect of each input random variable onto the variance of the QoI. This information is useful for reducing the dimensionality of the input stochastic space by eliminating variables that have a small effect of the variance. The computed Sobol indices are shown in Table 3 for two chaos orders; the se-gPC predictions match with the analytic values for $p = 6$ to two decimal points. The CN values of $\boldsymbol{\phi}^\top \boldsymbol{W}' \boldsymbol{\phi}$ are 1.2×10^3 and 3.5×10^5 for $p = 6$ and $p = 10$ respectively.

In the Ishigami function, where the number of uncertain variables is relatively small ($m = 3$) and the required chaos order is high, the computational advantages of the se-gPC are not maximized, even though the method still outperforms the sparse quadrature. In the following, we consider cases with larger values of m.

6.2. Two-dimensional viscous Burgers' equations

We consider the two-dimensional viscous Burgers' equations,

$$\begin{aligned} u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} &= \frac{1}{Re}\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right), \\ u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} &= \frac{1}{Re}\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right). \end{aligned} \tag{41}$$

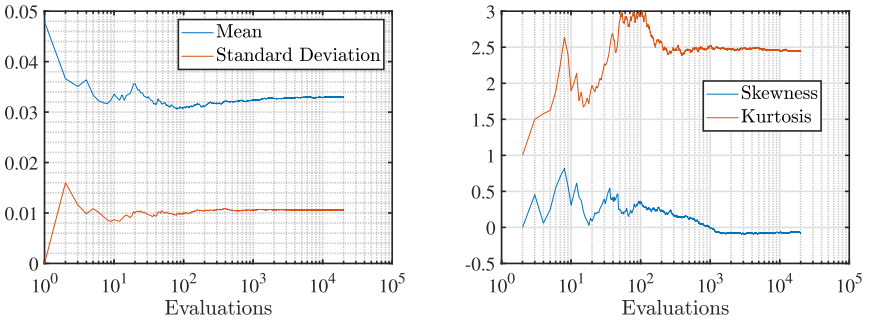


Fig. 8. Variation of the first four moments of k_ϵ, see (42), against the number of Monte-Carlo samples for $Re = 250$ and polynomial coefficients with mean $s_1 = -0.5$, $s_2 = -0.1$, $s_3 = 0.1$, $s_4 = 0.01$, $s_5 = -0.25$, $s_6 = 0.15$, $s_7 = 0.15$, $s_8 = -0.1$, $s_9 = 0.01$, $s_{10} = -0.25$, and standard deviation $\sigma_{s_i} = |s_i|/5$ $(i = 1, \ldots, 10)$.

This set is discretised in a square domain $\Omega = [0, 1] \times [0, 1]$ with Dirichlet boundary conditions at the top and bottom boundaries $u(x, 0) = u(x, 1) = v(x, 0) = v(x, 1) = 0$ and Neumann boundary conditions, $\frac{\partial u}{\partial x}(1, y) = \frac{\partial v}{\partial x}(1, y) = 0$ at the exit. At the inlet, the u component of the velocity is a polynomial of degree $m + 1$, i.e. $u(0, y) = \sum_{i=0}^{m+1} s_i y^i$, where the $m + 2$ polynomial coefficients, s_i, are uncertain variables. Coefficients $s_1, s_2, \ldots, s_m$ are free, while s_0 and s_{m+1} are computed so that the boundary conditions at the top and bottom corners $u(0, 0) = u(0, 1) = 0$ are satisfied. The v velocity at the inlet is given by $v(0, y) = -y^3 + y^2$, which satisfies the Dirichlet conditions $v(0, 0) = v(0, 1) = 0$.

The QoI is the integral of the kinetic energy at the exit of the domain,

$$k_e = \frac{1}{2} \int_0^1 \left(u^2 + v^2\right)\Big|_{x=1} dy. \tag{42}$$

For this test case we use the continuous adjoint approach to compute the gradients of the QoI. The equations for the adjoint variables u^+ and v^+ can be easily derived and take the form,

$$\begin{aligned} u^+ \frac{\partial v}{\partial y} + u \frac{\partial u^+}{\partial x} + v \frac{\partial u^+}{\partial y} + \frac{1}{Re}\left(\frac{\partial^2 u^+}{\partial x^2} + \frac{\partial^2 u^+}{\partial y^2}\right) = v^+ \frac{\partial v}{\partial x}, \\ v^+ \frac{\partial u}{\partial x} + u \frac{\partial v^+}{\partial x} + v \frac{\partial v^+}{\partial y} + \frac{1}{Re}\left(\frac{\partial^2 v^+}{\partial x^2} + \frac{\partial^2 v^+}{\partial y^2}\right) = u^+ \frac{\partial u}{\partial y}, \end{aligned} \tag{43}$$

The boundary conditions for u^+ are $u^+(0, y) = 0$, $u^+(x, 0) = 0$ and $u^+(x, 1) = 0$. Similarly for v^+, we get $v^+(0, y) = 0$, $v^+(x, 0) = 0$ and $v^+(x, 1) = 0$. The Neumann conditions at the exit, together with the form of the QoI, lead to Robin boundary conditions for the adjoint variables,

$$u^+ u + \frac{1}{Re}\frac{\partial u^+}{\partial x} + u = 0 \text{ and } v^+ u + \frac{1}{Re}\frac{\partial v^+}{\partial x} + v = 0, \text{ for } x = 1 \tag{44}$$

The sensitivities of the QoI with respect to s_i $(i = 1, \ldots, m)$ are given by

$$\frac{dk_e}{ds_i} = -\int_0^1 \left(u^+ + v^+\right) u\Big|_{x=0} y^i dy - \frac{1}{Re}\int_0^1 \frac{\partial u^+}{\partial x}\bigg|_{x=0} y^i dy \tag{45}$$

We consider a case with $m = 10$ uncertain coefficients. The Reynolds number was set to $Re = 250$ and Ω was discretized using second-order finite differences in a uniform 31×31 grid. To generate reference data for comparison, we performed $20,000$ Monte-Carlo simulations. The variation of the four moments of k_e against the number of samples is shown in Fig. 8. The 3^{rd} and 4^{th} order moments, skewness and kurtosis respectively, have converged to one significant digit.

In Fig. 9 the percentage errors of the first four moments (with respect to the aforementioned Monte-Carlo statistics) as computed by Smolyak Quadrature, WLSQ and se-gPC with minimum QR points, are presented. Again each point represents a different chaos order. The number of evaluations indicates the number of flow solutions, either direct or adjoint (they are assumed to have the same cost), that are required. Symbols are shown only for values of p that allow the computation of the moments. All methods converge, but the se-gPC achieves higher accuracy for the same computational cost compared to the other methods. For the same percentage error, the other two methods require one or two orders of magnitude more evaluations (note the logarithmic scale of the horizontal axis). Again it is possible that for low p, Smolyak may yield a lower error, but se-gPC converges faster as p increases. The benefits of se-gPC are significant for this case that has a relatively large number of stochastic variables. The CN of $\boldsymbol{\phi}^\top \boldsymbol{W}' \boldsymbol{\phi}$ for $p = 3$ was 7.8×10^3.

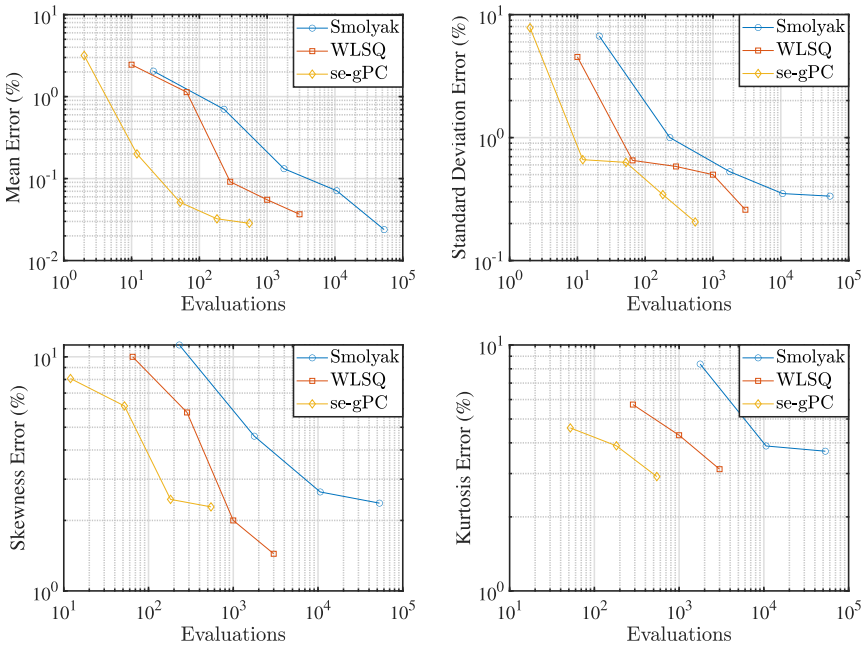


Fig. 9. Convergence of the first four moments of QoI k_ϵ, equation (42), computed using Smolyak quadrature, WLSQ and se-gPC at the QR points, against number of evaluations. The points on each curve refer to a different chaos order. Symbols are shown only for values of p that allow the computation of the moments.

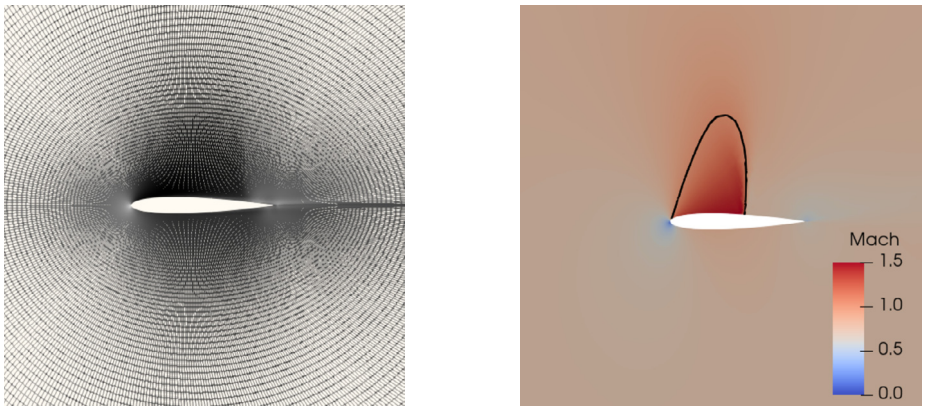


Fig. 10. Left: Unstructured mesh around the NACA0012 airfoil with 85000 nodes (zoomed-in view close to the airfoil). Right: Contour plots of Mach number at $M_\infty = 0.75$ and $a_\infty = 4.0°$. The black solid line indicates $M = 1$.

6.3. Transonic inviscid 2D flow around a NACA0012 airfoil

This case considers the inviscid transonic $2D$ flow around a NACA0012 airfoil. The Euler solver of the SU2 package [50] was used to perform the numerical simulations for the direct and adjoint problems. The mesh consists of approximately $85,000$ nodes, it is refined close to the wall, and extends to 20 chord lengths in the far field. A zoomed-in view is shown in Fig. 10a. At nominal free-stream conditions, the angle of attack is $a_\infty = 4.0°$ and the Mach number is $M_\infty = 0.75$. For these conditions, the predicted lift coefficient is $C_L = 0.84$, in good agreement with experimental data [50]. Contour plots of the Mach number are shown in Fig. 10b. The flow is in the transonic regime, and a shock appears at the suction side of the airfoil.

The QoI for this case is the lift coefficient, C_L. Two scenarios are considered; uncertainty in the free-stream conditions, with or without uncertainty in the geometry.

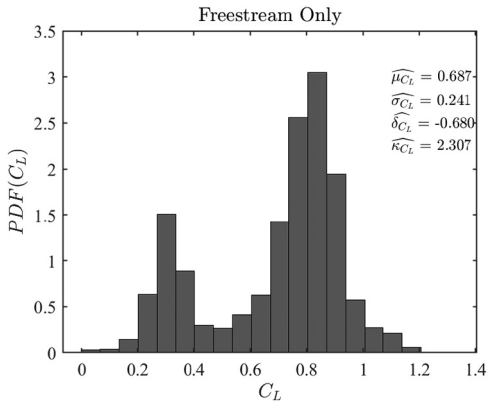


Fig. 11. PDF of C_L obtained from 5000 Monte Carlo simulations for $a_\infty \sim N(4.00°, 0.40°)$ and $M_\infty \sim N(0.75, 0.16)$.

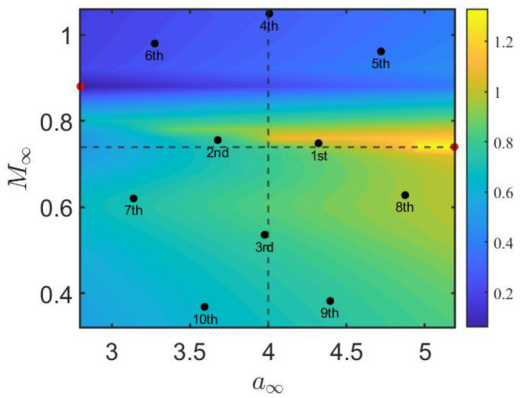


Fig. 12. Contour plot of C_L in the (a_∞, M_∞) plane. The QR points for $p = 3$ are also shown and are labelled with their ranking order. The two red dots mark the conditions of minimum and maximum C_L (contours of Mach number are shown in Fig. 13). The vertical and horizontal dashed lines denote the nominal free-stream conditions. (For interpretation of the colours in the figure(s), the reader is referred to the web version of this article.)

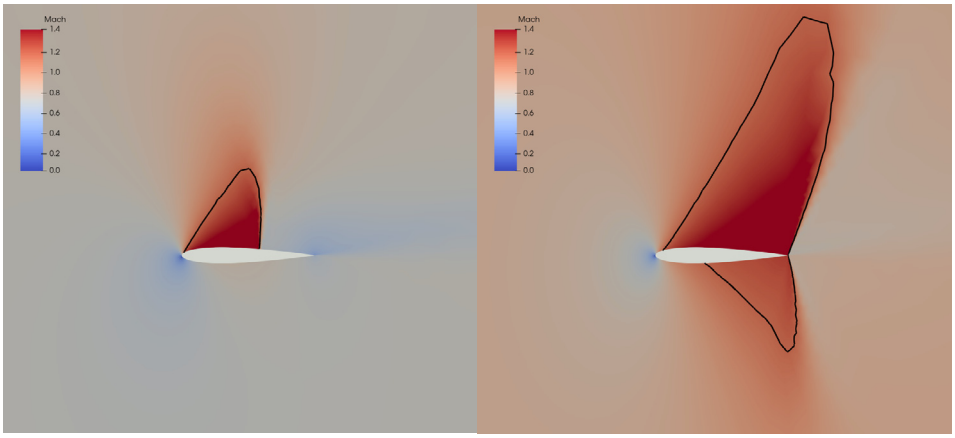
6.3.1. Uncertainty in the free-stream conditions

In the first scenario, uncertainty is introduced only in the free-stream conditions, that follow a normal distribution with $a_\infty \sim N(4.0°, 0.40°)$ and $M_\infty \sim N(0.75, 0.16)$. We run Monte-Carlo simulations with 5000 samples to quantify the effect of these uncertainties on the QoI. The PDF of C_L is shown in Fig. 11; it has a bimodal distribution with two modes, one with low C_L (in the range $[0.2, 0.45]$) and one with high C_L (in the range $[0.6, 1.1]$). At the nominal operating conditions, the lift coefficient belongs to the high C_L mode. The strong deviation from Gaussian distribution makes this a very interesting case.

To shed more light into this behaviour, a contour plot of C_L in the (a_∞, M_∞) plane is shown in Fig. 12; the examined rage is $\pm 3\sigma$ around the nominal values for both stochastic variables. In the plot, we also superimpose the QR sampling points for $p = 3$, with the labels indicating the ranking order. As expected, the highest ranked points are closer to the mean. However, contrary to Figs. 2b and 3, the mean value is not one of the QR points. We have empirically observed that for Gaussian uncertainties the QR points include the mean value only for even p, while for odd p the points are placed on either side of the mean.

As can be seen, the lift coefficient depends strongly on both parameters. It is well known that the variation of C_L with M_∞ in the transonic regime is highly non-linear [51]. Visualisation of the flow fields has shown that with increasing the Mach number, while keeping a_∞ constant, a shock starts to appear on the suction side of the airfoil at around $M_\infty = 0.6$, with area and strength increasing as M_∞ increases. This corresponds to the region of high C_L (green/yellow area on the plot). However, for higher M_∞, the lift coefficient transitions from the high C_L mode to the low C_L mode (blue area). This is associated with a qualitative change in the flow field.

In Fig. 13a contours of Mach number at $M_\infty = 0.74$ and $a_\infty = 5.2°$ are shown; these conditions (marked with a red dot on the right edge of Fig. 12) correspond to the maximum $C_L = 1.35$. It can be seen that only one shock appears on the suction surface of the airfoil. On the other hand, for free-stream conditions $M_\infty = 0.88$ and $a_\infty = 2.8°$ (marked with a red dot on the left edge of Fig. 12), a shock appears on both the suction and pressure sides close to the trailing edge, see Fig. 13b. In the largest portion of the bottom side the flow is supersonic and the pressure is low, causing the lift to drop to the smallest value of $C_L = 0.06$. More generally, free-stream conditions that lead to the presence of shocks on both sides induce lower C_L coefficients, and are responsible for the low-lift mode of the PDF. On the other hand, conditions that lead



(a) (b)

Fig. 13. Contour plot of M for (a) $M_\infty = 0.74$ and $a_\infty = 5.2°$ and (b) $M_\infty = 0.88$ and $a_\infty = 2.8°$. The two conditions are marked with red dots in Fig. 12. The black solid line indicates $M = 1$.

Table 4
First four moments of C_L coefficient as computed by Monte Carlo (with 5000 samples) and se-gPC (with $p = 3$). The Monte Carlo values are considered as reference for the evaluation of percentage error (the variance of the Monte-Carlo computed C_L is 0.0058).

Method	$\widehat{\mu}_{C_L}$	$\widehat{\sigma}_{C_L}$	$\widehat{\delta}_{C_L}$	$\widehat{\kappa}_{C_L}$
Monte-Carlo	0.687	0.241	−0.680	2.307
se-gPC ($p = 3$)	0.687	0.240	−0.659	2.511
Error (%)	< 0.01	0.41	3.08	8.84

to a single shock on the suction side only have a higher lift coefficient and belong to the high-lift mode. This change of the flow pattern explains the bi-modal shape of the PDF. For more details regarding the observed minimum and maximum C_L values and the change of the flow patterns with increasing M_∞ in the transonic flow regime, refer to chapter 9 of [51].

Comparison of the first 4 moments of the C_L coefficient as computed by Monte Carlo and se-gPC with $p = 3$ are shown in Table 4. For this case, there are $P + 1 = 10$ unknown coefficients. At each QR point, the value of C_L and the sensitivities dC_L/dM_∞, dC_L/da_∞ are computed with one direct and one adjoint evaluation, thus we need only 4 QR points (that will provide 12 equations). As can be seen from the table, the mean and standard deviation are computed with less than 1% error, and even for the kurtosis the error is less than 10%. The CN for the se-gPC system for $p = 3$ was equal to 8.5.

6.3.2. Geometric and free-stream uncertainties

In the second scenario, we consider uncertainties in both the free-stream conditions as well as the geometry. Uncertainties in the geometric shape are present in physical settings and can affect the airfoil performance. To model them, we employ Hicks-Henne bump functions, which have been used to describe shape deformation in aerodynamic optimization and wing design problems [52,53]. The parameterized geometry is given by,

$$y(x) = y_{base}(x) + \sum_{i=1}^{m_G} b_i(x), \qquad b_i(x) = a_i \left[\sin\left(\pi x^{\frac{log 0.5}{log h_i}} \right) \right]^{t_i}, \tag{46}$$

where m_G is the number of bump functions, a_i is the bump amplitude, t_i controls the width of the bump, and h_i determines the location of the bump maximum. We consider $m_G = 38$ stochastic coefficients a_i that follow normal distribution, $a_i \sim N(1, 0.01)$. The standard deviation is selected to be small; larger values can result in non-physical airfoil shapes. In total, we have $m = m_G + 2 = 40$ stochastic parameters, with large variance in the free-stream conditions and small variance for the parameters that describe the geometry. This case mirrors some of the features that are found in robust aerodynamic shape optimization. Note that the number of uncertain parameters is not an exaggeration, similar cases in aerodynamic shape optimization use dozens of shape parameters [54–56]. The open source CFD software $SU2$ [50] was used again to simulate the flow and calculate the sensitivities of C_L with respect to the geometric parameters using the adjoint approach.

The PDF of the lift coefficient C_L, obtained from 5000 Monte Carlo simulations, is shown in Fig. 14. The PDF remains bi-modal as in the first scenario, but has different statistical properties. The performance of the airfoil is significantly affected by the geometric uncertainties. For example, the mean value of the lift is reduced by 8.3% from 0.687 to 0.630. The higher order moments are also affected.

The computational cost (number of evaluations) of the three methods considered (Smolyak Quadrature, WLSQ, and se-gPC at QR sampling points) is seen in Table 5 for different values of the chaos order p. The CPU hours required for each

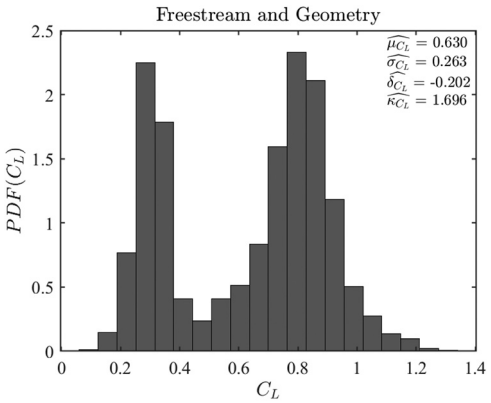


Fig. 14. PDF of C_L for the NACA12 airfoil produced by 5000 Monte-Carlo simulations. The $m = 40$ stochastic inputs are $M_\infty = N(0.75, 0.16)$, $a_\infty = N(4°, 0.4°)$ and $a_i \sim N(1, 0.01)$ with $i = 1, \ldots, 38$.

Table 5
Number of evaluations (top value in each cell) and CPU hours (bottom value in each cell) required for Smolyak Quadrature, WLSQ and se-gPC at the QR points for different chaos orders p. Stars indicate computational cost more expensive than the Monte Carlo simulations.

Method	$p = 1$	$p = 2$	$p = 3$
Smolyak	81	3321	91881
	1.98h	80.5h	$*$
WLSQ (at QR)	41	861	12341
	0.97h	21.2h	$*$
se-gPC (at QR)	2	42	602
	0.05h	1.04h	15.4h

Table 6
Comparison of the first four moments of C_L for uncertain free-stream and geometry conditions, as computed by Smolyak Quadrature, WLSQ and se-gPC at the WTS points for different values of the chaos order p. Errors are computed with respect to MC with 5000 sample (reference values are, $\widehat{\mu} = 0.630$, $\widehat{\sigma} = 0.263$, $\widehat{\delta} = -0.202$ and $\widehat{\kappa} = 1.696$. A dash denotes that the moment cannot be computed.

Method	Smolyak $p = 1$	Smolyak $p = 2$	WLSQ $p = 1$	WLSQ $p = 2$	se-gPC $p = 1$	se-gPC $p = 2$	se-gPC $p = 3$
$\widehat{\mu}$	0.638	0.633	0.626	0.627	0.635	0.633	0.631
$\epsilon_{\widehat{\mu}}$	1.26%	0.48%	0.64%	0.48%	0.79%	0.48%	0.15%
$\widehat{\sigma}$	0.259	0.261	0.266	0.264	0.266	0.265	0.263
$\epsilon_{\widehat{\sigma}}$	1.52%	0.76%	1.14%	0.38%	1.14%	0.76%	0.08%
$\widehat{\delta}$	–	−0.182	–	−0.234	–	−0.219	−0.209
$\epsilon_{\widehat{\delta}}$	–	9.9%	–	15.84%	–	8.41%	3.46%
$\widehat{\kappa}$	–	–	–	–	–	–	1.799
$\epsilon_{\widehat{\kappa}}$	–	–	–	–	–	–	6.07%

method are also presented (a star indicates that the computational time is longer than the Monte-Carlo simulations). The application of se-gPC results in a very drastic reduction in the number of evaluations and CPU time (one to two orders of magnitude). The difference in performance is much more prominent for this case compared to the previous ones because of the large number of stochastic variables. The CN of the se-gPC system for $p = 3$ was 3.8×10^5.

Comparison with the Monte Carlo results for the four first moments is shown in Table 6, with the top value in each cell representing the prediction and the bottom value the % error with respect to MC with $N = 5000$ evaluations. Se-gPC provides similar accuracy compared to the WLSQ and Smolyak Quadrature, but at a much reduced cost. For example, for $p = 1$, se-gPC provides a very good approximation to the first two statistical moments that is ~ 20 times less expensive than WLSQ and ~ 40 times less expensive than Smolyak (see Table 5). As mentioned earlier, for $p = 1$ the cost of the se-gPC is independent of the number of uncertain variables. In the context of aerodynamic design the number of control variables is usually large, and it is useful to be able to get fast estimations of the mean and standard deviation. Larger values of p allow accurate approximations of high order moments also at a much reduced computational cost. For example, the se-gPC produced the most accurate prediction of the skewness and it was the only method that was able to produce an estimation of the kurtosis.

7. Conclusions

An LSQ approach enriched with the sensitivity (or gradient) of the QoI is considered for uncertainty quantification. The gradient of QoI with respect to the stochastic variables is computed very efficiently (at the cost equal to a single direct evaluation) with the aid of adjoint variables. The sampling points to apply the direct and adjoint evaluations are selected from a large pool of candidate points so as to maximise the determinant of the appropriate least squares system matrix. Two sampling strategies based on pivoted QR decomposition of the standard LSQ matrix and the augmented LSQ matrix were considered. It was found that selecting the QR points of the standard LSQ matrix and computing at these points the gradient fully exploits all the sensitivity information provided by the adjoint formulation. The resulting augmented LSQ matrix has condition number that reaches moderate values even for cases with large number of stochastic variables and chaos orders. If the sampling points are selected randomly, as previously done in the literature, the CN can be orders of magnitude larger. The method is called sensitivity-enhanced generalised polynomial chaos, or se-gPC.

For the lowest chaos order $p = 1$, se-gPC applied at the minimum number of QR points requires only two evaluations, i.e. the cost is independent of the dimension of the stochastic space, m. For higher values of p, the number of the required evaluations is smaller compared to other methods, such as Smolyak quadrature and the standard LSQ, by a factor equal to m. Thus, the savings increase as the dimension of the stochastic space grows. The accuracy of se-gPC was assessed against the standard weighted LSQ as well as Smolyak quadrature and the results were found to be of comparable accuracy for the same chaos order, but were obtained at a much reduced computational cost, between one to two orders of magnitude smaller for the cases examined.

The se-gPC aims to speed up non-intrusive UQ approaches when adjoint sensitivity information is available. The latter is also useful for optimisation purposes and many open source and commercial packages provide solvers for the adjoint equations. This functionality can be directly exploited for efficient UQ calculations for complex engineering problems with many stochastic variables.

CRediT authorship contribution statement

Kyriakos D. Kantarakias: Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **George Papadakis:** Conceptualization, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We would like to thank the anonymous reviewers for bringing to our attention pertinent references, and for their detailed and constructive comments that have significantly improved the paper. K.D.K. wishes to acknowledge the financial support of the President's Scholarship Award from Imperial College London. GP is supported by EPSRC, grants EP/X017273/1, EP/W001748/1 and EP/W00481X/1.

References

[1] S. Asmussen, P. Glyn, Stochastic Simulation: Algorithms and Analysis. Stochastic Modelling and Applied Probability, Springer-Verlag, New York, 2007.
[2] W.J. Morokoff, R.E. Caflisch, Quasi-Monte Carlo integration, J. Comput. Phys. 122 (2) (1995) 218–230.
[3] N. Wiener, The homogeneous chaos, Am. J. Math. 60 (1938) 897–936.
[4] D. Xiu, G.E. Karniadakis, The Wiener–Askey polynomial chaos for stochastic differential equations, SIAM J. Sci. Comput. 24 (2) (2002) 619–644.
[5] R. Ghanem, P.D. Spanos, Polynomial chaos in stochastic finite elements, J. Appl. Mech. 57 (1) (1990) 197–202.
[6] J.-C. Chassaing, D. Lucor, J. Trégon, Stochastic nonlinear aeroelastic analysis of a supersonic lifting surface using an adaptive spectral method, J. Sound Vib. 331 (2) (2012) 394–411.
[7] D. Schiavazzi, A. Doostan, G. Iaccarino, A. Marsden, A generalized multi-resolution expansion for uncertainty propagation with application to cardiovascular modeling, Comput. Methods Appl. Mech. Eng. 314 (2017) 196–221.
[8] M. Chatzimanolakis, K.-D. Kantarakias, V. Asouti, K. Giannakoglou, A painless intrusive polynomial chaos method with RANS-based applications, Comput. Methods Appl. Mech. Eng. 348 (2019) 207–221.
[9] K.D. Kantarakias, K. Shawki, G. Papadakis, Uncertainty quantification of sensitivities of time-average quantities in chaotic systems, Phys. Rev. E 101 (2020) 022223.
[10] B.A. Jones, A. Doostan, Satellite collision probability estimation using polynomial chaos expansions, Adv. Space Res. 52 (11) (2013) 1860–1875.
[11] O.P.L. Maitre, O.M. Knio, Spectral Methods for Uncertainty Quantification, Springer, Dordrecht, 2010.
[12] S. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, Sov. Math. Dokl. 4 (1963) 240–243.

[13] B. Ganapathysubramanian, N. Zabaras, Sparse grid collocation schemes for stochastic natural convection problems, J. Comput. Phys. 225 (1) (2007) 652–685.
[14] P.G. Constantine, M.S. Eldred, E.T. Phipps, Sparse pseudospectral approximation method, Comput. Methods Appl. Mech. Eng. 229–232 (2012) 1–12.
[15] J. Winokur, P. Conrad, I. Sraj, O. Knio, A. Srinivasan, W.C. Thacker, Y. Marzouk, M. Iskandarani, A priori testing of sparse adaptive polynomial chaos expansions using an ocean general circulation model database, Comput. Geosci. 899 (911) (2013) 17.
[16] T. Gerstner, M. Griebel, Numerical integration using sparse grids, Numer. Algorithms 18 (1998) 209–232.
[17] G. Blatman, B. Sudret, Adaptive sparse polynomial chaos expansion based on least angle regression, J. Comput. Phys. 230 (6) (2011) 2345–2367.
[18] J. Hampton, A. Doostan, Compressive sampling of polynomial chaos expansions: convergence analysis and sampling strategies, J. Comput. Phys. 280 (2015) 363–386.
[19] M. Hadigol, A. Doostan, Least squares polynomial chaos expansion: a review of sampling strategies, Comput. Methods Appl. Mech. Eng. 332 (2018) 382–407.
[20] A.D. Belegundu, Lagrangian approach to design sensitivity analysis, J. Eng. Mech. 111 (5) (1985) 680–695.
[21] P. Luchini, A. Bottaro, Adjoint equations in stability analysis, Annu. Rev. Fluid Mech. 46 (1) (2014) 493–517.
[22] J.J. Reuther, A. Jameson, J.J. Alonso, M.J. Rimlinger, D. Saunders, Constrained multipoint aerodynamic shape optimization using an adjoint formulation and parallel computers, Part 1, J. Aircr. 36 (1) (1999) 51–60.
[23] J. Reuther, J. Alonso, M. Rimlinger, A. Jameson, Aerodynamic shape optimization of supersonic aircraft configurations via an adjoint formulation on distributed memory parallel computers, Comput. Fluids 28 (4) (1999) 675–700.
[24] A.K. Alekseev, I.M. Navon, M.E. Zelentsov, The estimation of functional uncertainty using polynomial chaos and adjoint equations, Int. J. Numer. Methods Fluids 67 (3) (2011) 328–341.
[25] S.S. Isukapalli, A. Roy, P.G. Georgopoulos, Efficient sensitivity/uncertainty analysis using the combined stochastic response surface method and automated differentiation: application to environmental and biological systems, Risk Anal. 20 (5) (2000) 591–602.
[26] N.H. Kim, H. Wang, N.V. Queipo, Adaptive reduction of random variables using global sensitivity in reliability-based optimisation, Int. J. Reliab. Saf. 1 (1–2) (2006) 102–119, https://doi.org/10.1504/IJRS.2006.010692.
[27] Y. Li, M. Anitescu, O. Roderick, F. Hickernell, Orthogonal bases for polynomial regression with derivative information in uncertainty quantification, Int. J. Uncertain. Quantificat. 1 (4) (2011) 297–320.
[28] O. Roderick, M. Anitescu, P. Fischer, Polynomial regression approaches using derivative information for uncertainty quantification, Nucl. Sci. Eng. 164 (2) (2010) 122–139.
[29] L. Guo, A. Narayan, T. Zhou, A gradient enhanced ℓ_1-minimization for sparse approximation of polynomial chaos expansions, J. Comput. Phys. 367 (2018) 49–64.
[30] J. Jakeman, M. Eldred, K. Sargsyan, Enhancing ℓ_1-minimization estimates of polynomial chaos expansions using basis selection, J. Comput. Phys. 289 (2015) 18–34.
[31] J. Peng, J. Hampton, A. Doostan, On polynomial chaos expansion via gradient-enhanced ℓ_1-minimization, J. Comput. Phys. 310 (2016) 440–458.
[32] H.-S. Chung, J. Alonso, Using gradients to construct co-Kriging approximation models for high-dimensional design optimization problems, in: 40th AIAA Aerospace Sciences Meeting & Exhibit, 2002.
[33] R. Dwight, Z.-H. Han, Efficient uncertainty quantification using gradient-enhanced Kriging, in: 50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, 2009.
[34] J. Laurenceau, P. Sagaut, Building efficient response surfaces of aerodynamic functions with kriging and cokriging, AIAA J. 46 (2) (2008) 498–507.
[35] B.A. Lockwood, M. Anitescu, Gradient-enhanced universal kriging for uncertainty propagation, Nucl. Sci. Eng. 170 (2) (2012) 168–195.
[36] B. Lockwood, D. Mavriplis, Gradient-based methods for uncertainty quantification in hypersonic flows, in: International Workshop on Future of CFD and Aerospace Sciences, Comput. Fluids 85 (2013) 27–38.
[37] P. Diaz, A. Doostan, J. Hampton, Sparse polynomial chaos expansions via compressed sensing and D-optimal design, Comput. Methods Appl. Mech. Eng. 336 (2018) 640–666.
[38] K. Manohar, B.W. Brunton, J.N. Kutz, S.L. Brunton, Data-driven sparse sensor placement for reconstruction: demonstrating the benefits of exploiting known patterns, IEEE Control Syst. Mag. 38 (3) (2018) 63–86, https://doi.org/10.1109/MCS.2018.2810460.
[39] A. Sommariva, M. Vianello, Computing approximate Fekete points by QR factorizations of Vandermonde matrices, Comput. Math. Appl. 57 (8) (2009) 1324–1336.
[40] L. Guo, A. Narayan, L. Yan, T. Zhou, Weighted approximate Fekete points: sampling for least-squares polynomial approximation, SIAM J. Sci. Comput. 40 (1) (2018) A366–A387.
[41] J. Hampton, A. Doostan, Coherence motivated sampling and convergence analysis of least squares polynomial chaos regression, Comput. Methods Appl. Mech. Eng. 290 (2015) 73–97.
[42] A. Jameson, Aerodynamic design via control theory, J. Sci. Comput. 3 (3) (1988) 233–260.
[43] K. Manohar, B.W. Brunton, J.N. Kutz, S.L. Brunton, Data-driven sparse sensor placement for reconstruction: demonstrating the benefits of exploiting known patterns, IEEE Control Syst. Mag. 38 (3) (2018) 63–86.
[44] T. Ishigami, T. Homma, An importance quantification technique in uncertainty analysis for computer models, in: [1990] Proceedings. First International Symposium on Uncertainty Modeling and Analysis, 1990, pp. 398–403.
[45] B. Sudret, Global sensitivity analysis using polynomial chaos expansions, Reliab. Eng. Syst. Saf. 93 (7) (2008) 964–979.
[46] S. Abraham, M. Raisee, G. Ghorbaniasl, F. Contino, C. Lacor, A robust and efficient stepwise regression method for building sparse polynomial chaos expansions, J. Comput. Phys. 332 (2017) 461–474.
[47] D.L. Allaire, K.E. Willcox, A variance-based sensitivity index function for factor prioritization, Reliab. Eng. Syst. Saf. 107 (2012) 107–114.
[48] P. Kersaudy, B. Sudret, N. Varsier, O. Picon, J. Wiart, A new surrogate modeling technique combining kriging and polynomial chaos expansions – application to uncertainty analysis in computational dosimetry, J. Comput. Phys. 286 (2015) 103–117.
[49] G. Blatman, B. Sudret, An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis, Probab. Eng. Mech. 25 (2) (2010) 183–197.
[50] T.D. Economon, F. Palacios, S.R. Copeland, T.W. Lukaczyk, J.J. Alonso, SU2: an open-source suite for multiphysics simulation and design, AIAA J. 54 (3) (2016) 828–846.
[51] J.J. Bertin, R.M. Cummings, Aerodynamics for Engineers, 6th edition, Pearson Education Limited, 2014.
[52] R.M. Hicks, P.A. Henne, Wing design by numerical optimization, J. Aircr. 15 (7) (1978) 407–412.
[53] G. Yang, A. Da Ronch, J. Drofelnik, Z.-T. Xie, Sensitivity assessment of optimal solution in aerodynamic design optimisation using SU2, Aerosp. Sci. Technol. 81 (2018) 362–374.
[54] A. Shahrokhi, A. Jahangirian, Airfoil shape parameterization for optimum Navier–Stokes design with genetic algorithm, Aerosp. Sci. Technol. 11 (6) (2007) 443–450.
[55] S. Schmidt, C. Ilic, V. Schulz, N.R. Gauger, Three-dimensional large-scale aerodynamic shape optimization based on shape calculus, AIAA J. 51 (11) (2013) 2615–2627.
[56] T. Lukaczyk, F. Palacios, J.J. Alonso, Active subspaces for shape optimization, in: 10th AIAA Multidisciplinary Design Optimization Conference, American Institute of Aeronautics and Astronautics, National Harbor, Maryland, 2014, pp. 1–18.