Quantification of airfoil geometry-induced aerodynamic uncertainties - comparison of approaches

Dishi Liu[1]
Lilienthalplatz 7, 38108 Braunschweig, Germany

Alexander Litvinenko[2]
4700 KAUST, M.B. 4436, Thuwal, Jeddah 23955-6900, Saudi Arabia

Claudia Schillings[3]
ETH, 8092 Zurich, Switzerland

Volker Schulz[4]
Universitätsring 15, 54296 Trier, Germany

Uncertainty quantification in aerodynamic simulations calls for efficient numerical methods since it is computationally expensive, especially for the uncertainties caused by random geometry variations which involve a large number of variables. This paper compares five methods, including quasi-Monte Carlo quadrature, polynomial chaos with coefficients determined by sparse quadrature and gradient-enhanced version of Kriging, radial basis functions and point collocation polynomial chaos, in their efficiency in estimating statistics of aerodynamic performance upon random perturbation to the airfoil geometry which is parameterized by 9 independent Gaussian variables. The results show that gradient-enhanced surrogate methods achieve better accuracy than direct integration methods with the same computational cost.

---

[1] Research Scientist, Institute of Aerodynamics and Flow Technology, German Aerospace Center (DLR), Germany. Email: dishi.liu@dlr.de
[2] Research Scientist, SRI-UQ center, King Abdullah University of Science and Technology, Saudi Arabia.
[3] Postdoctoral Associate, Seminar for Applied Mathematics ETH, Switzerland.
[4] Professor, Universität Trier, Germany.

I. Introduction

Among many uncertainties aerodynamic simulations are subject to, uncertainties caused by unpredictable variation in aircraft geometry is the focus of this paper. These geometry variations could occur by various reasons, e.g. manufacturing tolerances, ice accumulated on aircraft surface, hail damage, fatigue of material etc. They have the potential to dramatically lower the aerodynamic performance and need to be included in the numerical simulation to ensure reliability of the designs. Enabling uncertainty quantification in this framework is critical for the robust design of aircrafts. Uncertainty quantification is a vivid, fast growing research field in the CFD community. In the context of aerodynamic design, most papers focus on the treatment of uncertainties in the flow conditions, e.g. the Mach number or the angle of attack, see [1, 2], and only very little attention has been paid to uncertainties in the geometry itself. In [3, 4], first experiments based on first or second order Taylor approximations to quantify the influence of shape uncertainties can be found. But the method strongly relies on a small perturbations assumptions and becomes invisible for larger variations and nonlinear behavior of the forward problem. In [5], an analysis for geometry uncertainties is presented, which relies on a parametrization of the uncertainties using only three parameters. This approach allows to work in the low dimensional setting, but significantly restricts the space of possible variations. A general framework based on random fields can be found in [6, 7]. However this quantification is usually computationally challenging due to two reasons: the large number of variables and the high computational cost of computational fluid dynamics (CFD) models, which call for efficient numerical methods. In this paper, we will focus on smooth deviations from the original shape as they may arise for instance during the manufacturing process. We will set up a test case with uncertain airfoil geometry, and compare the efficiency of five numerical methods on quantifying the uncertainties in the aerodynamic performance.

Table 1 shows a categorization of some typical methods used in uncertainty quantification to obtain some target statistics of aerodynamic performance. There are basically two types: direct integration and surrogate-based integration. The former obtains the statistics by a direct integration

	Direct integration	Surrogate-based integration
Scattered sampling	Monte Carlo Quasi-Monte Carlo	Non-intrusive polynomial chaos Kriging Radial basis functions
Regular grid sampling	Tensor product quadrature Sparse grid quadrature	Stochastic collocation

Table 1 Some typical uncertainty quantification methods

with the samples, while the latter uses the samples to set up a surrogate of the CFD model and obtains the statistics by sampling the much cheaper surrogate model. These methods adopt either scattered sampling or regular grid sampling scheme, the former scheme admits samples anywhere in the variable domain, while the latter only samples on a prescribed grid (either tensorial or sparse).

In uncertainty quantification (UQ) problems with a large number of variables, like those with geometry uncertainties, the methods based on tensorial grid are usually too expensive. In [**?**] and [8] sparse grid quadratures were employed to alleviate the "curse of dimensionality". But with these quadratures one has limited choices of sample numbers. In UQ methods with scattered sampling scheme, Monte-Carlo quadrature and its variance-reduced kin (e.g. Latin Hypercube) have dimension-independent error convergence rate $O(N^{-\frac{1}{2}})$ with N the sample number, while the quasi-Monte Carlo quadrature has the worst case rate $O(\log^d(N)N^{-1})$ with d the number of variables. Due to the higher degree of sample uniformity the latter is usually more accurate than the former even in the cases that d is large.

UQ methods based on surrogate models are gaining more use recently, e.g. using radial basis functions or Kriging in [6, 9–11] and using polynomial chaos expansion in [12–14], [15] proposes a non-intrusive Galerkin way of constructing surrogate. [11] shows that a Kriging surrogate is better than plain Monte Carlo and Latin Hypercube methods in estimating mean value of a bivariate Rosenbrock function. Surrogate-based methods have another advantage when the gradients can be sampled at a relatively lower cost. This is because the direct integrations cannot effectively utilize the gradient information (augmenting samples generated by finite differences are not statistically independent hence bring little benefit to the accuracy of the statistics). Cheaper gradients can be

obtained by using an adjoint CFD solver [16] if the number of response quantities is less than the number of variables.

In this work five methods (see description in Section IV) are applied to an aerodynamic UQ problem caused by airfoil geometry uncertainties to compare their efficiency. They include one direct integration method, namely quasi-Monte Carlo (QMC) quadrature, and four surrogate-based methods: polynomial chaos (PC) with the coefficients computed by a sparse Gauss-Hermite (SGH) quadrature (as introduced in [17]), gradient-enhanced Kriging (GEK), gradient-enhanced radial basis functions (GERBF) and gradient-enhanced polynomial chaos (GEPC) method. The efficiency of these methods are then compared in estimating some statistics and probability distribution of the uncertain lift and drag coefficients. To make the comparison more meaningful to industrial applications the number of CFD model evaluations is kept small (≤ 500).

The plan for the rest of the paper is as follows: Section II sets up a test case with uncertainty in airfoil geometry and a parameterization of the uncertainties. Section III introduces the target and criteria of the efficiency comparison. Section IV is devoted to introductions of methods in the comparison. Numerical result and discussion are presented in Section V, and conclusion in Section VI.

II. A test case of uncertain airfoil geometry

The test case in this work is a CFD model of inviscid flow around a RAE2822 airfoil solved by TAU [18] at a Mach number of 0.73 and an angle of attack of 2.0 degree. The space is discretized by a 193×33 grid in which the airfoil is discretized by 128 surface nodes, as shown in Figure 1.

The source of uncertainty is a random perturbation to the original airfoil geometry, which is modeled by a random field in the simulation. Representing the original geometry (see Figure 4) by a matrix $\mathbf{G} \in \mathbb{R}^{128\times 2}$ which contains the x, y coordinates of the 128 surface nodes, the random field of perturbation can be written as $\boldsymbol{R}(\boldsymbol{g}, \omega) \in \mathbb{R}^{128}$ with $\boldsymbol{g} = \{x, y\} \in \mathbf{G}$ and $\omega \in \Omega$, Ω is the sample space of the random perturbation.

To define $\boldsymbol{R}(\boldsymbol{g}, \omega)$ one first defines a Gaussian random field $\boldsymbol{\psi}(\boldsymbol{g}, \omega)$ and then make $\boldsymbol{R}$ a transformation of $\boldsymbol{\psi}$. The reason for this is that Gaussian fields can be easily parameterized by independent

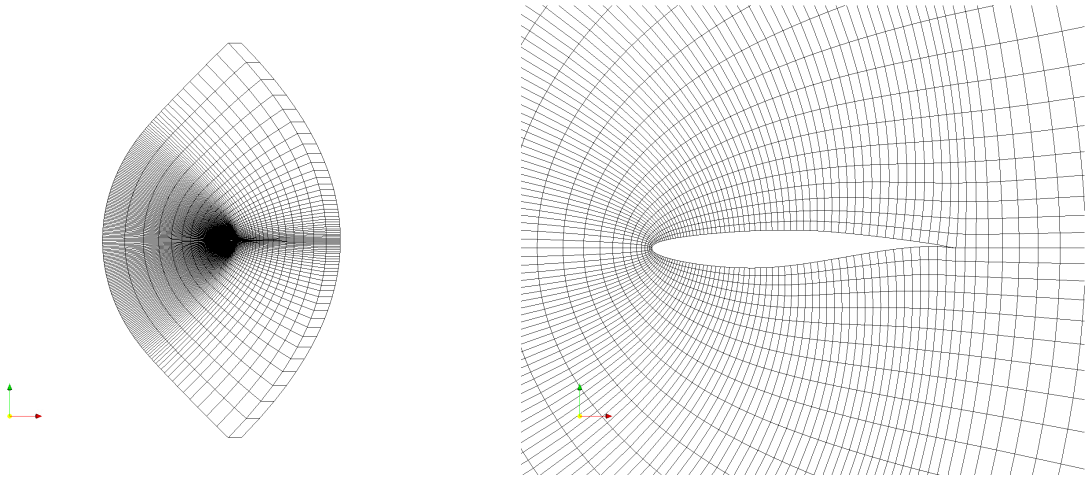

Fig. 1 Grid for the RAE2822 airfoil : the total grid (left) and zoom around the airfoil (right).

random variables by using Karhunen-Loève expansions.

The Gaussian field $\boldsymbol{\psi}(\boldsymbol{g},\omega)$ is defined by a zero mean value and a Gaussian-type correlation function:

$$\operatorname{cov}(\boldsymbol{g}_i,\boldsymbol{g}_j) = \sigma(\boldsymbol{g}_i)\ \sigma(\boldsymbol{g}_j) \exp\left(-\frac{\|\boldsymbol{g}_i-\boldsymbol{g}_j\|^2}{\ell^2}\right) \tag{1}$$

with correlation length $\ell = 0.005$, the standard deviation σ is assumed to be:

$$\sigma(\boldsymbol{g}) = \begin{cases} (0.8-x)^{0.75} & \text{if } x \le 0.8 \\ 0 & \text{otherwise} \end{cases}$$

This is to avoid perturbation to the trailing end of the airfoil to guarantee the convergence of the CFD solver.

To ensures the boundedness of the perturbations, the field of random perturbation $\boldsymbol{R}$ is made to have a sine-shaped probability distribution by such a transformation of the Gaussian field $\boldsymbol{\psi}$:

$$\boldsymbol{R}(\boldsymbol{g},\omega) = s(x)\cdot\arccos\left(1-2\Phi\left(\boldsymbol{\psi}\left(\boldsymbol{g},\omega\right)\right)\right), \tag{2}$$

where Φ is the Gaussian cumulative distribution function and $s(x) = (0.8-x)^{0.75}\cdot\sqrt{0.00002}$.

The given model leads to smooth, small-scale variations of the original shape as they arise for instance during manufacturing processes. Note that the methods described below to quantify the influence of these uncertainties do not rely on the specific transformation chosen in this paper. However, the treatment of sharp or even discontinuous perturbations requires additional effort in

the modeling process as well as in the forward propagation of these uncertainties and is not subject of this study.

Perturbed geometries are generated by imposing the perturbation $\boldsymbol{R}$ to the original geometry, in the direction normal to the airfoil surface:

$$\widetilde{\mathbf{G}}(\omega) = \mathbf{G} + \boldsymbol{R}(\boldsymbol{g}, \omega) \cdot \boldsymbol{n}(\boldsymbol{g})$$

$\boldsymbol{n}(\boldsymbol{g})$ is the normal vector at $\boldsymbol{g}$.

However the above representation has 128 correlated variables while one prefers a representation with independent variables in a fewer number. A model reduction can be made by an approximation of $\boldsymbol{\psi}$ through Karhunen-Loève expansions (KLE) [19], i.e.

$$\boldsymbol{\psi}(\boldsymbol{g}, \omega) \approx \sum_{i=1}^{k} \sqrt{\lambda_i}\, \boldsymbol{V}_i(\boldsymbol{g})\, \xi_i(\omega) =: \widehat{\boldsymbol{\psi}}(\boldsymbol{g}, \omega) \tag{3}$$

where λ_i and $\boldsymbol{V}_i$ denote the eigenvalues and eigenvectors of the covariance matrix generated by Equation (1) and ξ_i's are independent standard Gaussian random variables. $k = 9$ in this test case since the first 9 eigenvalues correspond to 99.98% of the characteristics. This KLE representation minimizes the loss in the variance of perturbation due to the truncation of terms, so is optimal in all linear form representations with the same number of variables [19]. In numerical implementation $\widehat{\boldsymbol{\psi}}(\boldsymbol{g}, \omega)$ is used instead of $\boldsymbol{\psi}(\boldsymbol{g}, \omega)$ in Equation (2), so that the random field of perturbation $\boldsymbol{R}$ is parameterized by 9 independent Gaussian variables. In [20] it is shown how to compute only k largest eigenvalues and eigenvectors for large covariance matrices.

Figure 2 and 3 display three realizations of the $\boldsymbol{\psi}$ and $\boldsymbol{R}$ respectively. Examples of perturbed geometries $\widetilde{\mathbf{G}}$ are shown in Figure 4. Figure 5 shows the fast decline of the eigenvalues λ_i.

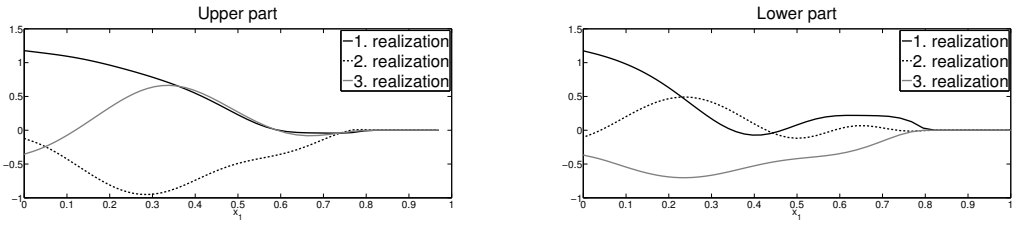

Fig. 2 Realizations of the Gaussian random field $\boldsymbol{\psi}$: upper (left) and lower (right) part of airfoil.

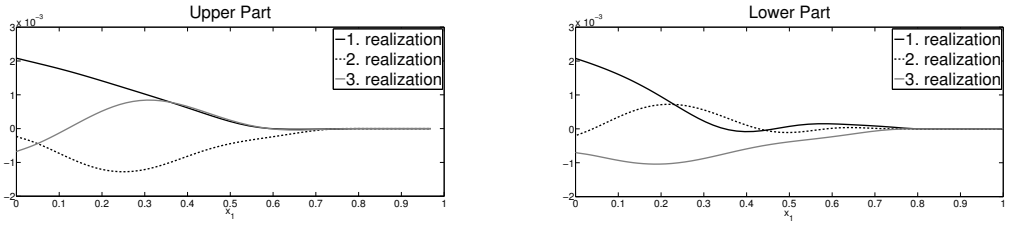

Fig. 3 Realizations of the transformed random field R: upper (left) and lower (right) part of airfoil.

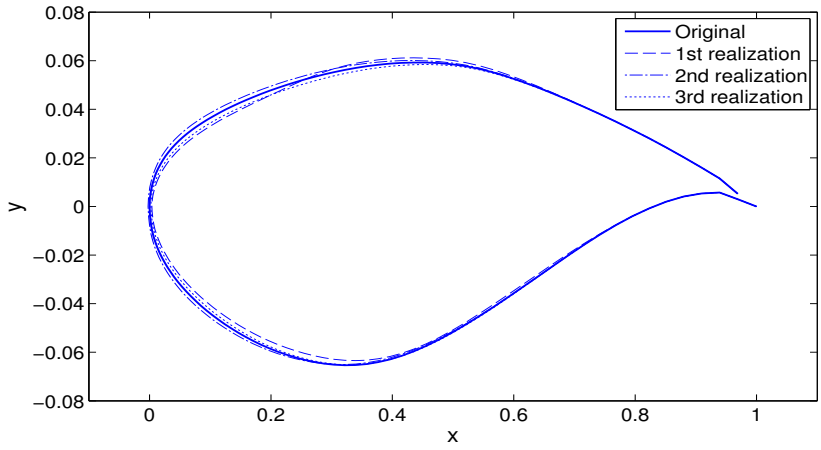

Fig. 4 Three realizations of perturbed geometry

III. Settings of the comparison

With the above parameterization of geometric uncertainties, the five UQ methods are applied to the test case and compare their efficiency in estimating some target statistics of two aerodynamic

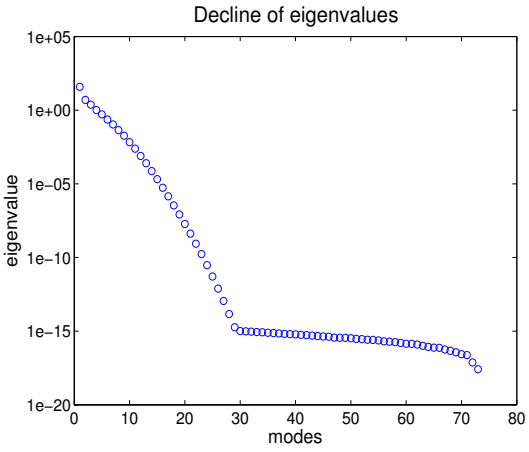

Fig. 5 Decline of eigenvalues

system response quantities (SRQ), namely the coefficients of lift and drag (C_L and C_D), as well as the probability distribution functions (pdf).

A. Target statistics and their reference values

The following statistics of C_L and C_D are to be estimated:

- means μ_L and μ_D
- standard deviations σ_L and σ_D
- exceedance probabilities $P_{L,\kappa} = \text{Pro}\{C_L \leq \mu_L - \kappa \cdot \sigma_L\}$ and $P_{D,\kappa} = \text{Pro}\{C_D \geq \mu_D + \kappa \cdot \sigma_D\}$ with $\kappa = 2, 3$

The reference values of these statistics are obtained from a relatively large number ($N = 4 \times 10^6$) of quasi-Monte Carlo (QMC) samples of the CFD model. To justify the validity of using these reference values in the efficiency comparison, their accuracy are estimated using Snyder's multipartition method [21] since the theoretical error bound of QMC integration is not a practical accuracy indicator. Suppose $\boldsymbol{S}$ denote any of the above statistics, by this method one makes an equal-size m-partition of the 4×10^6 samples, and obtained m estimates $\{\widetilde{\boldsymbol{S}}_i\}_{i=1}^m$ by integrating on each partition only, then compute the sample standard deviation ς_m of $\widetilde{\boldsymbol{S}}_i$. ς_m is an estimate of the standard deviation of QMC integral of $\boldsymbol{S}$ with sample number N/m. To extrapolate ς_m to ς_1 one computes ς_m for four values of m, i.e. $m = \{128, 64, 32, 16\}$, and fit a line across the four $(\log(N/m), \log(\varsigma_m))$ points using weighted linear least square method, m values are used as weights to account for the increasing variability for smaller values of m. With the fitted slope a and intercept b the ς_1 is extrapolated as $\varsigma_1 = \log^{-1}(a \log(N) + b)$ which is an estimate of the standard deviation of the QMC integral of $\boldsymbol{S}$ with N samples.

ς_1 values for each reference statistics are listed in Table 2. In the results of the efficiency comparison of the UQ methods (given in section V) the smallest absolute errors measured by using these reference values in estimating mean and standard deviation (stdv) are at least by 10 times larger than $3 \times \varsigma_1$, which means, by taking the assumption that the reference values in use are Gaussian distributed around the *true* values of the statistics, the measured errors have a 99.73%

confidence interval of at widest ±10%. For exceedance probabilities this confidence interval is still valid except for a few measured errors, as shown in Figure 6 and 7 where the values of the corresponding $3 \times \varsigma_1$ are depicted by dash line.

$\varsigma_1(\mu_L)$	$\varsigma_1(\sigma_L)$	$\varsigma_1(P_{L,2})$	$\varsigma_1(P_{L,3})$
9.1e-9	5.0e-8	1.8e-5	9.3e-6

$\varsigma_1(\mu_D)$	$\varsigma_1(\sigma_D)$	$\varsigma_1(P_{D,2})$	$\varsigma_1(P_{D,3})$
3.0e-9	1.4e-8	1.2e-5	6.8e-6

Table 2 Estimated stdv of reference statistics for C_L (left) and C_D (right)

B. Measure of accuracy and cost

The accuracy is measured by the absolute error in the estimates of the target statistics which is obtained by comparing with the reference values.

The computational cost is measured in term of "compensated evaluation number" M. With N denoting the number of CFD model evaluations, M is defined as $M = 3N$ for the three gradient-employing methods, since the computational cost of the gradients of each system response quantity (SRQ) using adjoint solver equals approximately to the cost of one CFD evaluation and one needs the gradients for two SRQs (C_L and C_D). For QMC and PC-SGH methods M is set to be equal to N .

C. Range of sample numbers

The number of CFD evaluations is kept smaller than 500 in this comparison, as shown in the Table 3. Most of the methods have the same range of sample number except for PC-SGH for which only quadrature level 2 and 3 are used since the sample number of level 4 is well beyond 500.

IV. Methods

In this section the methods in the comparison are introduced. More weights are given to two novel implementations, the Gradient-enhanced radial basis functions (GERBF) method and Gradient-enhanced polynomial chaos (GEPC) method. Only brief introductions or references are given to other methods.

For the Quasi-Monte Carlo (QMC) quadrature [22], Sobol low discrepancy sequence generated

Methods	N	M
QMC	$30 \sim 420$	$30 \sim 420$
PC-SGH (order 2, 3)	19 and 181	19 and 181
GEK	$10 \sim 140$	$30 \sim 420$
GERBF	$10 \sim 140$	$30 \sim 420$
GEPC (order 2,3 and 4)	11, 44 and 143	33, 132 and 429

Table 3 Sample number in the comparison

by the algorithm given in [23] is used since the sequence is shown comparable or slightly better than other sequences in a numerical comparison [24].

The gradient-enhanced Kriging (GEK) [25] is implemented by using *Surrogate-Modeling for Aero-Data Toolbox (SMART)* [26] developed at DLR, opting for ordinary Kriging and a correlation model of cubic spline type which is considered the most efficient in similar situations in [27]. The internal parameters of the correlation model are fine-tuned to fit the sampled data by a maximum likelihood estimation [28].

The gradient-enhanced radial basis function method is introduced in Section IV A, polynomial chaos method based on a sparse Gauss-Hermite quadrature (PC-SGH) and gradient-enhanced polynomial chaos method (GEPC) are introduced in Section IV B.

To quantify uncertainty via surrogate methods, one first establishes the surrogate based on a series of small number of QMC samples (as tabulated in Figure 3) of the CFD model, and integrates for the target statistics and probability density function (pdf) by a large number (1×10^6) of QMC samples on the surrogate model, only except for the polynomial chaos methods where the mean and variance can be directly obtained from the PC coefficients.

A. Gradient-enhanced radial basis functions

The radial basis function (RBF) method [29] approximates by a weighted linear combination of *radial basis functions* each being radially symmetric about a center. Denoting the CFD model as

f, its RBF approximate takes the form

$$\hat{f}(\boldsymbol{\xi}) = \sum_{i=1}^{N} w_i \, \phi_i(\|\boldsymbol{\xi} - \boldsymbol{\xi}^{\langle i \rangle}\|),$$

where ϕ_i are radial basis functions, $\|\cdot\|$ denotes the Euclidean norm, and $\boldsymbol{\xi}^{\langle i \rangle}$ is a sample point where ϕ_i is radial about. The coefficients w_i are determined by fitting $\hat{f}(\boldsymbol{\xi})$ to N samples.

Denoting the Euclidean distance from the center as r, popular types of $\phi(r)$ include $\sqrt{r^2 + a^2}$ (multiquadric), $1/\sqrt{r^2 + a^2}$ (inverse multiquadric), $\exp(-a^2 r^2)$ (Gaussian) and $r^2 \ln(ar)$ (thin plate spline), in which a is a parameter to be fine-tuned for a particular set of samples. Gradient-assisted RBF were proposed in [30, 31] where the gradients are exploited in the same way as gradient-enhanced Kriging (GEK). This method is numerically equal to a GEK that uses RBF as the correlation function.

A different gradient-employing RBF method is implemented in this work, which only requires RBF to be first-order differentiable, in contrast to the second-order differentiability required by the gradient-assisted RBF, thus is called gradient-enhanced RBF (GERBF) in this paper. To accommodate the gradient information, this method introduces additional RBF that are centered at *non-sampled* points, i.e. an GERBF approximation is simply

$$\hat{f}(\boldsymbol{\xi}) = \sum_{i=1}^{N(1+d)} w_i \, \phi_i(\|\boldsymbol{\xi} - \boldsymbol{\xi}^{\langle i \rangle}\|), \qquad \text{with } d \text{ the number of variables}$$

The $\boldsymbol{\xi}^{\langle i \rangle}$ with $i \leq N$ are sampled points, while those with $i > N$ are non-sampled points which can be chosen arbitrarily as long as none of them duplicates (or too close to) the sampled ones. The coefficients $\boldsymbol{w} = \{w_1, \cdots, w_{N(1+d)}\}^T$ are determined by solving the following system,

$$\mathbf{A}\boldsymbol{w} = \boldsymbol{f}$$

in which

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_0 \\ \vdots \\ \mathbf{A}_d \end{bmatrix}, \mathbf{A}_0 = \begin{bmatrix} \phi_1(\boldsymbol{\xi}^{\langle 1 \rangle}) & \cdots & \phi_{N(1+d)}(\boldsymbol{\xi}^{\langle 1 \rangle}) \\ \vdots & \ddots & \vdots \\ \phi_1(\boldsymbol{\xi}^{\langle N \rangle}) & \cdots & \phi_{N(1+d)}(\boldsymbol{\xi}^{\langle N \rangle}) \end{bmatrix}, \underset{d \geq i > 0}{\mathbf{A}_i} = \begin{bmatrix} \phi_1^{(i)}(\boldsymbol{\xi}^{\langle 1 \rangle}) & \cdots & \phi_{N(1+d)}^{(i)}(\boldsymbol{\xi}^{\langle 1 \rangle}) \\ \vdots & \ddots & \vdots \\ \phi_1^{(i)}(\boldsymbol{\xi}^{\langle N \rangle}) & \cdots & \phi_{N(1+d)}^{(i)}(\boldsymbol{\xi}^{\langle N \rangle}) \end{bmatrix}$$

and

$$\boldsymbol{f} = \begin{bmatrix} \boldsymbol{f}_0 \\ \vdots \\ \boldsymbol{f}_d \end{bmatrix}, \quad \boldsymbol{f}_0 = \begin{bmatrix} f(\boldsymbol{\xi}^{\langle 1 \rangle}) \\ \vdots \\ f(\boldsymbol{\xi}^{\langle N \rangle}) \end{bmatrix}, \quad \underset{d \geq i > 0}{\boldsymbol{f}_i} = \begin{bmatrix} f^{(i)}(\boldsymbol{\xi}^{\langle 1 \rangle}) \\ \vdots \\ f^{(i)}(\boldsymbol{\xi}^{\langle N \rangle}) \end{bmatrix}$$

where $\boldsymbol{\xi}^{\langle i \rangle} = \{\xi_1, \xi_2, \cdots, \xi_d\}_i$ with $i = 1, \cdots, N$ denotes a sample point, $\phi_i^{(j)} = \partial \phi_i / \partial \xi_j$, $f^{(j)} = \partial f / \partial \xi_j$.

Since $\mathbf{A}$ is often ill-conditioned the system is solved through a truncated singular value decomposition (SVD), i.e. $\boldsymbol{w} = (\mathbf{V}\mathbf{D}\mathbf{U}^\top)\boldsymbol{f}$, where $\mathbf{V}$ and $\mathbf{U}$ comes from a SVD $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top$, and the diagonal matrix $\mathbf{D}$ is related to $\boldsymbol{\Sigma}$ by:

$$\mathbf{D}_{ii} = \begin{cases} 0 & \text{if } |\boldsymbol{\Sigma}_{ii}|/\max_i(|\boldsymbol{\Sigma}_{ii}|) \leq N(1+d)\times 10^{-13} \\ 1/\boldsymbol{\Sigma}_{ii} & \text{otherwise} \end{cases}, \quad i = 1, \cdots, N(1+d)$$

This gives a solution with the smallest L_2 norm which is less spoiled by the ill-conditionedness [32, Chapter 15].

Inverse multiquadric RBF is used in this comparison. The internal parameter a is fine-tuned by a *leave-one-out* error minimizing procedure as in [9].

B. Polynomial chaos methods

According to Wiener [33], a surrogate could also be made in the form of a truncated *polynomial chaos expansion* (PCE) :

$$\hat{f}(\boldsymbol{\xi}) = \sum_{i=0}^{K} c_i H_i(\boldsymbol{\xi}) \tag{4}$$

where H_i is a multivariate Hermite polynomial (to which an introduction can be found in, e.g. [34]) with unbounded domain which is a natural choice in this test case with Gaussian variables. It is called "truncated" since the total number of terms K is taken finite for practical computation rather than infinite as in the theoretical formulation for an exact representation of a function $f \in L^2$. K is related to the maximal order of polynomial (p) and number of variables (d) by $K = (p+d)!/(p!d!)$.

The coefficients c_i in Equation (4) could be determined by various ways, in this work the following two of them are implemented.

1. *Projection based on sparse Gauss-Hermite quadrature*

Due to the mutual orthogonality of H_i's with respect to their inner product, the coefficients c_i can be expressed by

$$c_i = \frac{\int f(\boldsymbol{\xi}) H_i(\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi}}{\int H_i^2(\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi}},$$

where the integration can be made by using Gauss-Hermite quadrature if $\boldsymbol{\xi}$ is Gaussian. Using sparse Gauss-Hermite (SGH) quadrature alleviates the "curse of dimension" in this 9-variate integration, the disadvantage is the far-apart sample numbers of different levels, as shown in Table 3.

2. Gradient-enhanced point-collocation polynomial chaos method

Point-collocation polynomial chaos method [35] determines c_i by fitting equation (4) to N collocation points. [36] suggests to use $N = 2K$ for the best performance which leads to an over-determined system to be solved by least square approach. A gradient-employing version of this method is implemented in this work, in the name gradient-enhanced polynomial chaos (GEPC) method, where accordingly one chooses $N(1+d) = 2K$. $\boldsymbol{c} = \{c_0, c_1, \cdots, c_K\}^T$ is determined by solving the following system,

$$\boldsymbol{\Psi c} = \boldsymbol{f}$$

with

$$\boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{\Psi}_0 \\ \vdots \\ \boldsymbol{\Psi}_d \end{bmatrix}, \quad \boldsymbol{\Psi}_0 = \begin{bmatrix} H_1(\boldsymbol{\xi}^{\langle 1\rangle}) & \cdots & H_K(\boldsymbol{\xi}^{\langle 1\rangle}) \\ \vdots & \ddots & \vdots \\ H_1(\boldsymbol{\xi}^{\langle N\rangle}) & \cdots & H_K(\boldsymbol{\xi}^{\langle N\rangle}) \end{bmatrix}, \quad \underset{d\geq j>0}{\boldsymbol{\Psi}_j} = \begin{bmatrix} H_1^{(j)}(\boldsymbol{\xi}^{\langle 1\rangle}) & \cdots & H_K^{(j)}(\boldsymbol{\xi}^{\langle 1\rangle}) \\ \vdots & \ddots & \vdots \\ H_1^{(j)}(\boldsymbol{\xi}^{\langle N\rangle}) & \cdots & H_K^{(j)}(\boldsymbol{\xi}^{\langle N\rangle}) \end{bmatrix}$$

with $H_i^{(j)} = \partial H_i/\partial \xi_j$. $\boldsymbol{\Psi}$ consists of the values of the H_i and $\partial H_i/\partial \xi_j$ evaluated at N sample points, sized $N(1+d) \times K$. $\boldsymbol{f}$ is as defined in the GERBF method. Polynomial chaos of order $p = 2, 3$ and 4 are used in this test case, the corresponding values of N and "compensated sample number" M are tabulated in Table 3.

Once a surrogate based on polynomial chaos is established, the estimates of the mean and the variance of $f(\boldsymbol{\xi})$ can be directly obtained from the coefficients:

$$\mu = c_0\ , \qquad \sigma^2 \approx \sum_{i=1}^{K} \left[c_i^2 \int H_i^2(\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi} \right], \tag{5}$$

where the integration can be made analytically. The exceedance probabilities and pdf are integrated by a large number (1×10^6) of QMC samples on the surrogate model.

V. Numerical results

The results of the efficiency comparison are shown in Figures 6-9. Figures 6 and 7 show the errors of the five methods in estimating the target statistics of C_L and C_D. The numbers of CFD

evaluations (horizontal coordinates in these pictures) are given in Table 3. It is observed there that generally the gradient-employing surrogate methods perform better than direct integration methods. This can be ascribed to that the former utilize more information with the same computational cost M, i.e. they use $(1+d)\frac{M}{3}$ conditions (SRQs and their gradients) while a direct integration method uses M conditions (SRQs only). This advantage comes from the cheaper cost of the gradients computed by an adjoint solver in the case that the number of SRQ's (in our case, 2) is smaller than d (in our case, 9), and the advantage would increase with an increasing d. Of course, surrogate methods could have outperformed even without the gradient information but simply by a good approximation of the CFD model which provides cheap and accurate surrogate samples.

The PC-SGH method has only two data points due to the very limited choices of sample number. It is hard to evaluate its error convergence property on only two data points. This also shows a shortcoming on using sparse grid quadratures on problems with a relative high number of variables. The GEPC method shows constant convergence in error, but not as efficient as the other two gradient-employing methods. The reason could be that the increase of admissible order of polynomial chaos along increasing M is slow in this 9-variate problem, and that the polynomial surrogate tend to "overshoot" in the outskirt of the domain. But the PC methods have a merit that they do not need a parameter optimization procedure.

GEK and GERBF are the most efficient methods as far as seen from these results. This can be attributed to properties of the kernel functions they use, namely cubic spline and inverse multiquadric functions, and also to the effort of tuning parameters. Convergence rate of inverse multiquadric RBF was estimated $\mathcal{O}(e^{-c/h})$ with h the fill distance and c a constant which translated to a rate in N at $\mathcal{O}(e^{-cN^{1/d}})$ [37]. We assume the error in the statistics is proportional to that in surrogate samples, and find in this 9-variate test case the observed convergence rate is much better than $\mathcal{O}(e^{-cN^{1/9}})$, this hints that the "effective" fill distance h reduces faster than $\mathcal{O}(N^{-1/d})$ due to that some variables are less important than the others (typical for a KLE parameterization).

GEK seems slightly better than GERBF according to these results, especially at smaller M values. But this cannot be generalized to other implementations of the two methods. The advantage of GEK could come from the possible advantage of cubic spline over inverse multiquadric in this

particular case. Different ways of utilizing gradient information by GEK (involving 1st and 2nd-order derivatives of basis functions and generating a symmetric kernel matrix) and GERBF (involving only 1st order derivative and generating non-symmetric matrix) could also contribute to the difference in their performance.

At larger M values a "rebound" of error can be observed in GERBF and GEK results, more obvious in the former. This is caused by the stabilizing treatment of these methods. Specifically, in the truncated SVD in the GERBF and matrix regularization in the GEK implemented by SMART (where a 10^{-13} "nugget" is added to the diagonal of the system matrix) the hard-coded 10^{-13} makes the matrix over- and under-regularized in the two methods respectively as the matrix size ($N(1+d)$) getting larger. Though there is an *uncertain relation* (or trade-off principle) exist stating that accuracy and stability cannot be good at the same time [38], more advanced stabilizing techniques are expected for these methods with larger sample numbers.

For mean and stdv the reliability of the measure errors are justified by that they are at least by 10 times larger than the corresponding $3 \times \varsigma_1$. For exceedance probabilities the $3 \times \varsigma_1$ values are not that small (as indicates by thick dash lines in the pictures) but are still small enough for most of the measured errors.

Figures 8 and 9 show the pdf of C_L and C_D estimated by QMC and the three gradient-employing surrogate methods with $M = 33$ (the smallest possible M value for GEPC method), comparing with the reference pdf (computed through 4×10^6 QMC samples). There one observes that for the same computational cost, the surrogate methods yield much more accurate pdf's. This is consistent with their relative performance in estimating the statistics.

VI. Conclusion

To compare efficiency of various methods in quantifying aerodynamic uncertainties, this paper sets up a test case where geometry of an RAE2822 airfoil is perturbed by a Gaussian random field which is parameterized by 9 independent Gaussian variables through a Karhunen-Loève expansion. Quasi-Monte Carlo (QMC) quadrature and four surrogate methods, polynomial chaos with coefficients determined by sparse grid quadrature, gradient-enhanced radial basis functions,

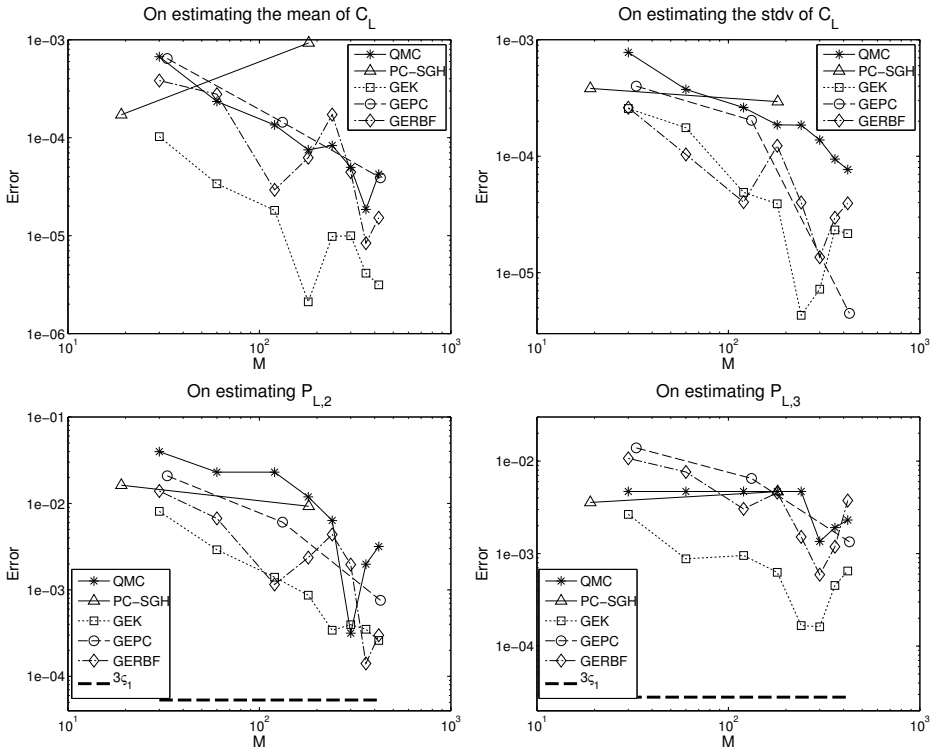

Fig. 6 Absolute error in estimating mean, standard deviation (upper row) and exceedance probabilities (lower row) of C_L

gradient-enhanced polynomial chaos and gradient-enhanced Kriging, are applied to the test case and compared in their efficiency in estimating some statistics and probability distribution of the uncertain lift and drag coefficients. The results show that gradient-employing surrogate methods achieve better accuracy than direct integration methods with the same computational cost.

Gradient information obtained at relatively lower cost by using adjoint solver gives an edge to surrogate methods since direct integration methods cannot utilized gradients effectively. This advantage of surrogate methods holds valid as long as the number of system response quantities (SRQ) is less than the number of variables d, and is expected to increase with an increasing d.

Acknowledgments

This research has been conducted within the project MUNA under the framework of the German Luftfahrtforschungsprogramm funded by the Ministry of Economics (BMWi).

Alexander Litvinenko did most of his work during his PostDoc research at Technische Univer-

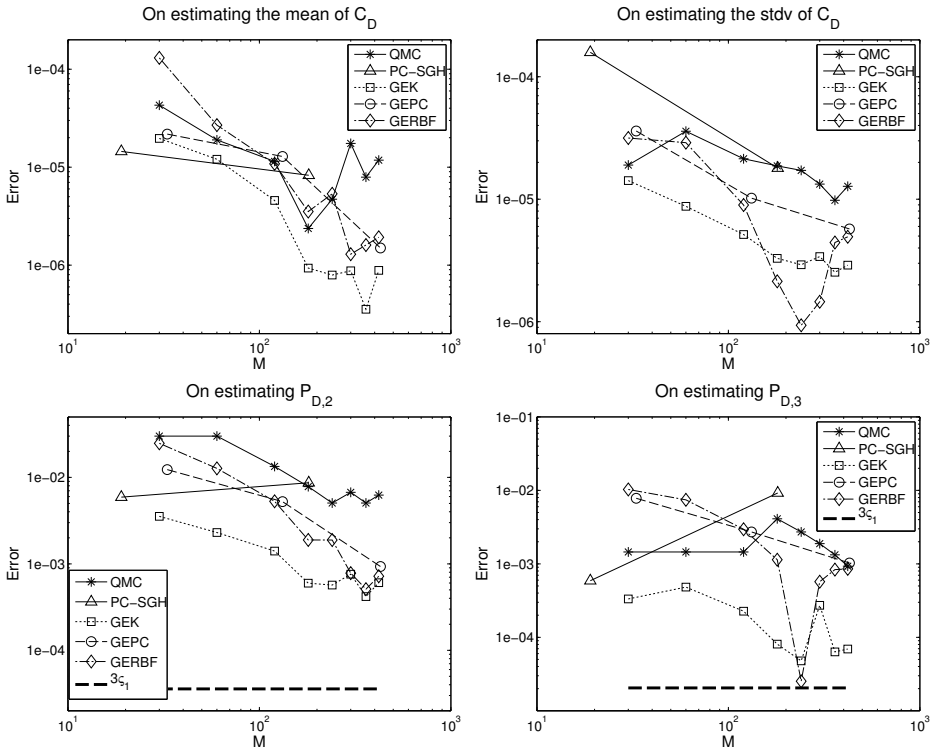

Fig. 7 Absolute error in estimating mean, standard deviation (upper row) and exceedance probabilities (lower row) of C_D

sität Braunschweig, Germany, the rest of his work was done at KAUST SRI-UQ center.

References

[1] H. Bijl, D. Lucor, S. Mishra, and C. Schwab, *Uncertainty quantification in computational fluid dynamics*, Lecture Notes in Computational Science and Engineering (Springer, 2013).

[2] V. Schulz and C. Schillings, AIAA Journal **47**, 646 (2009).

[3] S. Padula, C. Gumbert, and W. Li, Optimization and Engineering **7**, 317 (2006), ISSN 1389-4420.

[4] C. R. Gumbert, P. A. Newman, and G. J.-W. Hou, in *20th AIAA Applied Aerodynamics Conference* (2002).

[5] G. Loeven and H. Bijl, in *48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference* (AIAA-2008-2070, 2008).

[6] D. Liu and S. Görtz, in *New Results in Numerical and Experimental Fluid Mechanics IX*, edited by A. Dillmann et al. (Springer International Publishing, 2014), pp. 65–73.

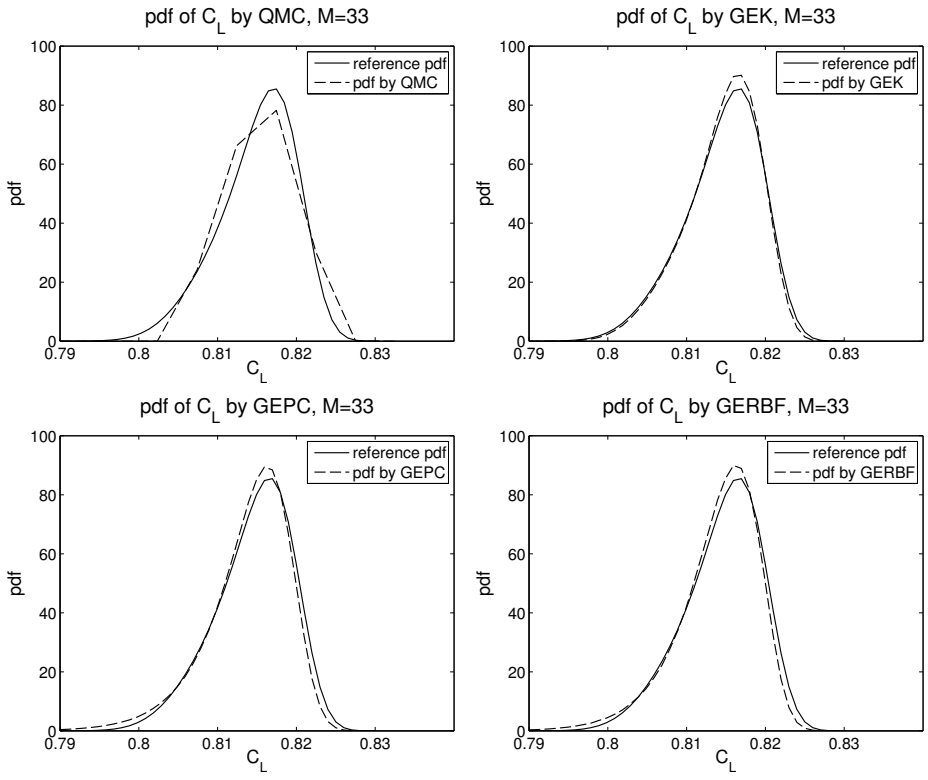

Fig. 8 Estimated pdf (in dash line) of C_L by QMC, GEK, GEPC and GERBF at $M = 33$.

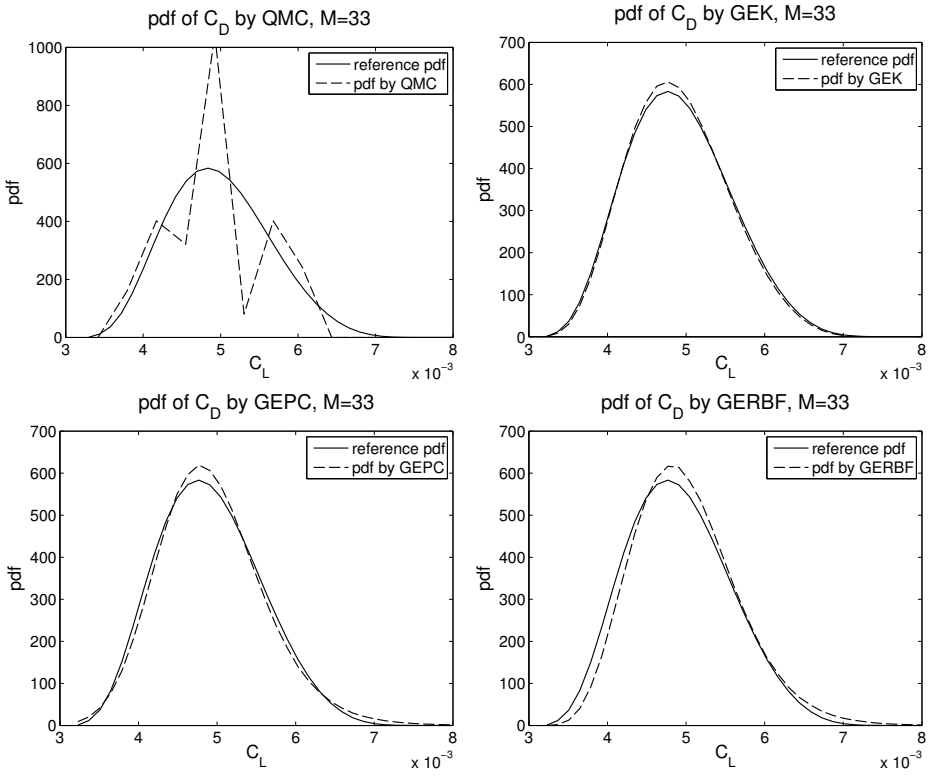

Fig. 9 Estimated pdf (in dash line) of C_D by QMC, GEK, GEPC and GERBF at $M = 33$.

[7] C. Schillings, S. Schmidt, and V. Schulz, Computers and Fluids **46**, 78 (2011), ISSN 0045-7930, 10th ICFD Conference Series on Numerical Methods for Fluid Dynamics (ICFD 2010).

[8] A. Litvinenko and H. G. Matthies, in *International Conference on Stochastic Modeling Techniques and Data Analysis Proceedings. Editor: Chr. H. Skiadas* (Chania, Greece, 2010), pp. 477–484.

[9] M. Bompard, J. Peter, and J.-A. Désidéri, in *V European Conference on Computational Fluid Dynamics, ECCOMAS CFD 2010* (Lisbon, Portugal, 2010).

[10] G. Loeven, J. Witteveen, and H. Bijl, in *Proceedings of the NATO RTO-MP-AVT-147 Computational Uncertainty in Military Vehicle design symposium* (2007).

[11] A. A. Giunta, M. S. Eldred, and J. P. Castro, in *9th ASCE Specialty Conference on Probabolistic Mechanics and Structural Reliability* (Albuquerque, New Mexico, USA, 2004).

[12] S. Dolgov, B. N. Khoromskij, A. Litvinenko, and H. G. Matthies, arXiv preprint arXiv:1406.2816 (2014).

[13] A. Litvinenko, H. Matthies, and T. A. El-Moselhy, in *Monte Carlo and Quasi-Monte Carlo Methods 2012*, edited by J. Dick, F. Y. Kuo, G. W. Peters, and I. H. Sloan (Springer Berlin Heidelberg, 2013), vol. 65 of *Springer Proceedings in Mathematics & Statistics*, pp. 535–551, ISBN 978-3-642-41094-9.

[14] A. Litvinenko and H. G. Matthies, PAMM **11**, 877 (2011), ISSN 1617-7061.

[15] L. Giraldi, A. Litvinenko, D. Liu, H. G. Matthies, and A. Nouy, arXiv preprint arXiv:1309.1617 (2013).

[16] J. Brezillon and R. Dwight, in *EUROGEN 2005 - Sixth Conference on Evolutionary and Deterministic Methods for Design, Optimization and Control with Applications to Industrial and Societal Problems* (Munich, Germany, 2005).

[17] B. Eisfeld, H. Barnewitz, W. Fritz, and F. Thiele, *Management and Minimisation of Uncertainties and Errors in Numerical Aerodynamics - Results of the German collaborative project MUNA* (Springer, 2013), ISBN 10: 3642361846.

[18] R. Heinrich, R. Dwight, M. Widhalm, and A. Raichle, in *MEGAFLOW - Numerical Flow Simulation for Aircraft Design*, edited by N. Kroll and J. Fassbender (Springer, 2005), vol. 89, pp. 93–108, ISBN 978-3-540-24383-0.

[19] R. J. Adler and J. E. Taylor, *Random Fields and Geometry* (Springer-Verlag, Berlin, 2007).

[20] B. N. Khoromskij, A. Litvinenko, and H. G. Matthies, Computing **84**, 49 (2009).

[21] W. C. Snyder, Mathematics and Computers in Simulation **54**, 131 (2000), ISSN 0378-4754.

[22] R. E. Caflisch, Acta Numerica **7**, 1 (1998).

[23] S. Joe and F. Kuo, SIAM Journal on Scientific Computing **30**, 2635 (2008).

[24] I. Radović, I. M. Sobol, and R. F. Tichy, Monte Carlo Methods and Applications **2**, 1 (1996).

[25] H.-S. Chung and J. J. Alonso, AIAA paper **317**, 14 (2002).

[26] Z. H. Han, S. Görtz, and R. Zimmermann, Journal of Aerospace Science and Technology (2012).

[27] D. Liu, Tech. Rep., Center for Computer Applications in AeroSpace Science and Engineering, Deutsches Zentrum für Luft- und Raumfahrt (DLR) (2012), iB 124-2012/2.

[28] R. Zimmermann, J. App. Math. (2010), article ID 494070.

[29] M. D. Buhmann, Acta Numerica **9**, 1 (2000).

[30] K. C. Giannakoglou, D. I. Papadimitriou, and I. C. Kampolis, Computer Methods in Applied Mechanics and Engineering **195**, 6312 (2006).

[31] Y. Ong, K. Lum, and P. Nair, Computational Optimization and Applications **39**, 97 (2008), ISSN 0926-6003.

[32] W. H. Press, *Numerical recipes 3rd edition: The art of scientific computing* (Cambridge university press, 2007).

[33] N. Wiener, Amer. J. Math. **60**, 897 (1938).

[34] H. G. Matthies, in *Encyclopedia of Computational Mechanics*, edited by E. Stein, R. de Borst, and T. R. J. Hughes (John Wiley & Sons, Chichester, 2007).

[35] S. Hosder, R. W. Walters, and R. Perez, in *Proceedings of the 44th AIAA Aerospace Sciences Meeting* (2006), vol. 14, pp. 10649–10667.

[36] S. Hosder, R. W. Walters, and M. Balch, in *Proceedings of the 48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, number AIAA-2007-1939, Honolulu, HI* (2007), vol. 125.

[37] H. Wendland, *Scattered data approximation* (Cambridge University Press, 2005).

[38] R. Schaback, Advances in Computational Mathematics **3**, 251 (1995).