AIAA AVIATION Forum
June 25-29, 2018, Atlanta, Georgia
2018 Multidisciplinary Analysis and Optimization Conference

10.2514/6.2018-3102

Hessian-based Dimension Reduction for Optimization Under Uncertainty

Kinshuk Panda* and Jason E. Hicken†
Rensselaer Polytechnic Institute, Troy, New York, 12180

We present an uncertainty propagation method for computing the expected value and variance of a quantity of interest (QoI), which can then be used in a robust design optimization. To avoid intractable costs due to high-dimensional integrals, we use the Hessian of the QoI to identify the dominant nonlinear directions. Specifically, the dominant Hessian eigenmodes provide the dimensions along which the QoI is integrated in stochastic space. Explicit computation of the Hessian is avoided by using Arnoldi's method to estimate the eigenmodes. The method is applied to multi-dimensional quadratic functions and its accuracy is examined for synthetic eigenmodes.

I. Introduction

Deterministic multidisciplinary design optimization (MDO) of aircraft has matured significantly over the last three decades, and state-of-the-art methods can produce designs that perform optimally under the prescribed conditions. However, these deterministically optimized designs can exhibit major degradation in performance when the prescribed operating conditions or design parameters are perturbed [1]. This motivates multidisciplinary design optimization under uncertainty (OUU), where the final design is resilient to variations in the aforementioned operating conditions and design parameters. Such a design philosophy falls within the field of robust design optimization [2].

The methodology for robust design optimization was pioneered by Taguchi [3, 4] and consists of three main steps. First, the uncertain parameters are identified and modeled using probability theory. Second, the uncertainty is propagated through the system being analyzed, which permits the evaluation of appropriate objective functions and constraints that account for uncertainties. Finally, the optimal solution is sought based on the statistical metrics developed in the second step. The second step in this process is typically the most expensive because accounting for all the uncertainties using traditional propagation methods, e.g. Monte Carlo simulation, requires many thousands, if not millions of deterministic analyses. This cost motivates the following research question: is it possible to accurately account for all the uncertainties in design optimization within a computationally tractable framework?

Uncertainties, depending on their nature, can be classified into two categories: aleatoric uncertainty and epistemic uncertainty [5]. Aleatoric uncertainties are inherent to a problem, and cannot be reduced by additional experiments. They are unbiased and often defined within a probabilistic framework. Examples of aleatoric uncertainties in the aeronautical context include manufacturing defects, cruise Mach number, and aircraft trajectory. Epistemic uncertainty arises from simplified model assumptions, or basic lack of knowledge. These uncertainties are biased and, in general, cannot be defined within a probabilistic framework. Examples of epistemic uncertainties include the inviscid flow assumption, simplified boundary conditions, and numerical errors. The focus of this work is on propagating aleatoric uncertainties.

Conventional methods of non-intrusive uncertainty propagation, such as Monte Carlo simulation, method of moments, and stochastic expansion, are prohibitively expensive for gradient-based optimization with many (>20) random variables. They either require large sample sizes and/or suffer from the curse of dimensionality. This has motivated the development of surrogate models that try to balance accuracy with computational costs [6–11]. In the context of aerodynamic design optimization, regression-based surrogate models, particularly kriging and cokriging models, have been used to propagate uncertainties [12–18]. However, these surrogate models also suffer from the curse of dimensionality. While gradient and Hessian information can be used to improve the kriging surrogate for high dimensional problems, the construction of a surrogate in itself requires several evaluations of a quantity of interest (QoI) and its gradient. This makes surrogate modeling potentially computationally expensive. Furthermore, using gradients for generating kriging surfaces is also known to result in ill-conditioning [19].

*Graduate Student, Department of Mechanical, Aerospace, and Nuclear Engineering, Student Member AIAA
†Assistant Professor, Department of Mechanical, Aerospace, and Nuclear Engineering, Member AIAA

In this paper we propose a dimension reduction strategy based on the eigenmodes of the Hessian of a QoI. The dominant eigenmodes of the QoI Hessian in isoprobabilistic space are used as proxies for the directions with the most nonlinearity. These dominant eigenmodes can be approximated using a few iterations of a Krylov method and the QoI gradient, so the Hessian does not need to be formed explicitly. Stochastic collocation, for example, is then applied along these dominant, nonlinear directions to compute statistics.

The remainder of the paper is organized as follows. Section II gives a brief summary of two uncertainty propagation methods. Our proposed dimension-reduction method is described in Section III and the corresponding results are presented in Section IV. Finally, Section V provides a summary of our work.

II. Uncertainty Propagation

We rely on standard uncertainty propagation methods as a kernel operation in our algorithm, so it is prudent that we briefly review the subject. Readers familiar with uncertainty propagation can safely proceed to Section III.

Uncertainty propagation is the process of modeling the impact of uncertainties in input variables on an output QoI. Statistical moments of the QoI are typically used to quantify this impact. Let $\xi = [\xi_1, ..., \xi_n]$ be a realization of an independent random variable vector Ξ that is normally distributed with mean μ_ξ and standard deviation σ. This is usually denoted by $\Xi \sim \mathcal{N}(\mu_\xi, \sigma^2)$. We denote a QoI as $J(\xi)$, whose expected value is given by

$$\mu_J = \int_{\Gamma_\xi} J(\xi) P_\Xi(\xi) d\xi, \tag{1}$$

where $P_\Xi(\xi)$ is the probability density of realization ξ, and Γ_ξ is the domain of the random variables. In general, practical problems do not have a closed form expression for Eq. (1) which necessitates its numerical approximation. We briefly review two popular uncertainty propagation methods for approximating Eq. (1) below.

A. Monte Carlo Simulation

The most common numerical approximation method for Eq. (1) is Monte Carlo simulation. For the simplest case of ν random samples, μ_J is approximated as

$$\mu_J \approx \frac{1}{\nu} \sum_{i=1}^{\nu} J(\xi^i), \tag{2}$$

where ξ^i is the ith realization of Ξ. The convergence rate of Monte Carlo methods is $O(\nu^{\frac{1}{2}})$, which is independent of the number of random variables. However, this slow convergence rate requires a large sample size, which hinders the application of Monte Carlo simulation when $J(\xi)$ is relatively expensive to compute. While the convergence rate of Monte Carlo methods can be improved with different sampling techniques [20–23], these sampling strategies are still not sufficient on their own to make Monte Carlo tractable for computationally demanding analyses. In addition, propagation using Monte Carlo methods also results in noisy objective and constraint functions, which presents challenges for gradient-based optimization [24].

B. Stochastic Collocation

Another uncertainty propagation method of interest is the stochastic collocation method [25], which expands J in a series of random variable realizations as follows:

$$J(\xi) \approx \sum_{i=1}^{N_p} L_i(\xi) J(\xi^i). \tag{3}$$

Here, N_p is the number of collocation points, and $L_i(\xi)$ is the multi-dimensional tensor-product application of 1D Lagrange polynomials. For example, for $\xi \in \mathbb{R}^2$, we have

$$L_{(i-1)\nu+j}(\xi) = L_{(i-1)\nu+j}(\xi_1, \xi_2) = \hat{L}_i(\xi_1)\hat{L}_j(\xi_2),$$

where the ith 1D Lagrange polynomial is given by

$$\hat{L}_i(u) = \prod_{\substack{j=1 \\ j \neq i}}^{\nu} \frac{u - u_j}{u_i - u_j}. \tag{4}$$

Because stochastic collocation uses multiple realizations of random variables to approximate QoI moments, it is considered to be a stochastic expansion method and is similar to polynomial chaos expansion [26–28].

The number and location of the collocation points, u_i, in Eq. (4) is critical to the quality of the expansion. Let J denote an arbitrary QoI that depends on a single random variable realization ξ, with mean μ_ξ, and standard deviation σ. Then the expected value of J can be written using a 1D quadrature of the form

$$\mu_J \approx \sum_{i=1}^{\nu} w_i J(\xi^i), \tag{5}$$

where w_i are the quadrature weights. For example, evaluating μ_J when Ξ has a Gaussian distribution involves the approximation

$$\mu_J = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} J(\xi) e^{-\frac{(\xi-\mu_\xi)^2}{2\sigma^2}} d\xi \approx \frac{1}{\sqrt{\pi}} \sum_{i=1}^{\nu} \tilde{w}_i J(\mu_\xi + \sqrt{2}\sigma\tilde{\xi}^i), \tag{6}$$

where $\tilde{\xi}^i$ and $\tilde{w}_i$ are the Gauss-Hermite quadrature locations and weights, respectively.

The 1D stochastic collocation shown in Eq. (5) can be extended to multiple dimensions using tensor products as follows:

$$\mu_J \approx \sum_{i_1=1}^{\nu_1} \dots \sum_{i_n=1}^{\nu_n} J(\xi_1^{i_1}, \dots, \xi_n^{i_n})(w_1^{i_1} \cdot w_2^{i_2} \cdot \dots \cdot w_n^{i_n}). \tag{7}$$

One can immediately see in Eq. (7) that tensor-product stochastic collocation methods suffer from the curse of dimensionality: the cost of the method grows as $N_p = \prod_{j=1}^{n} \nu_j$. For example, for a function of 10 independent random variables, a stochastic collocation approximation using just 4 quadrature points in each direction would require more than a million evaluations of J. In addition, the interpolation error for stochastic collocation has a convergence rate of $O(N_p^{\frac{-\gamma}{n}})$, assuming that $J(\xi)$ is γ times continuously differentiable. Hence, for a fixed number of collocation points, N_p, the accuracy decreases as the dimension, n, increases [29]. Work has been done to partially ameliorate the effects of dimensionality by using sparse grid techniques [30–36].

III. Hessian-based Dimension Reduction

Traditional uncertainty propagation methods, including those described in Section II, require a certain number of samples to produce accurate statistics. However, in the context of ordinary/partial differential equation-constrained optimization, one can only afford a limited number of QoI evaluations. Thus, it becomes important to choose a sampling strategy that can capture the behavior of the QoI efficiently. One way to do this is to identify the most nonlinear directions by some means and sample only along these directions. This is the strategy adopted here.

The proposed propagation method performs stochastic collocation in a subspace created by estimating the dominant directions of the Hessian of the QoI in isoprobabilistic space. For the purposes of explaining the method, consider a 2D quadratic QoI which is a function of a random variable Ξ, with realization $\xi = [\xi_1, \xi_2]$. Assume that Ξ has a standard bivariate normal distribution, i.e., $\Xi \sim \mathcal{N}([0,0]^T, [1,1]^T)$. We write the particular quadratic QoI as

$$\begin{aligned} J(\xi) &= 50\hat{\xi}_1^2(\xi) + \hat{\xi}_2^2(\xi), \\ \hat{\xi} &= R(\theta) * \xi, \end{aligned} \tag{8}$$

where $R(\theta)$ is a 2D rotation matrix and $\hat{\xi}$ are the rotated variables. The rotation is introduced so that we can explore the impact of Hessian eigenvectors that are not aligned with the random variable axes. For concreteness, suppose that the rotation matrix rotates the random variables by $\theta = 60°$. Then the eigenvalues, Λ, and eigenvectors, V, of the QoI Hessian, $\nabla^2_{\hat{\xi}} J$, are given by

Algorithm 1: Reduced Space Collocation

Data: random variable mean μ_ξ,
random variable covariance Σ,
1D parametric space quadrature location $\boldsymbol{q}$, and
1D parametric space quadrature weights $\boldsymbol{w}$
Result: Expected surrogate value $\bar{\mu}_J$

```
1  Λ, V ← eigfact(∇²_ξ J) in isoprobabilistic space
2  Get indices, ind, of dominant eigenvalues in Λ
3  V_dominant = V[:, ind]
4  μ̄_J = 0.0
5  n̄ ← length(ind)        (dimension of the approximation)
6  for i_1 = 1: length(q) do
7      for i_2 = 1: length(q) do
8          ⋱
9          for i_n̄ = 1: length(q) do
10             ξ̄ = [q_{i_1}, ..., q_{i_n̄}]
11             w̄ = w_{i_1} · w_{i_2} · ... · w_{i_n̄}
12             μ̄_J += w̄ J(μ_ξ + √(2Σ) V_dominant ξ̄)
13         end
14         ⋰
15     end
16 end
17 μ̄_J = μ̄_J / π^{n̄/2}
```

$$\Lambda = \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 100 \end{bmatrix} \qquad \text{and} \qquad \mathsf{V} = \begin{bmatrix} \mathsf{V}_1 & \mathsf{V}_2 \end{bmatrix} = \begin{bmatrix} \frac{-\sqrt{3}}{2} & \frac{-1}{2} \\ \frac{-1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}. \tag{9}$$

In the given example, $\lambda_2 > \lambda_1$, so V_2 is the dominant direction as shown in Fig. 1a. Finally, stochastic collocation is applied along the dominant direction V_2, treating V_1 as the invariant direction. This approximation is shown in Fig. 1b, where there is no variation in function values along V_1. More generally, the variation along V_1 could be linear and this approximation would remain equally accurate. Note that in higher dimensions, V_1 and V_2 are a collection of eigenvectors containing invariant and dominant directions respectively.

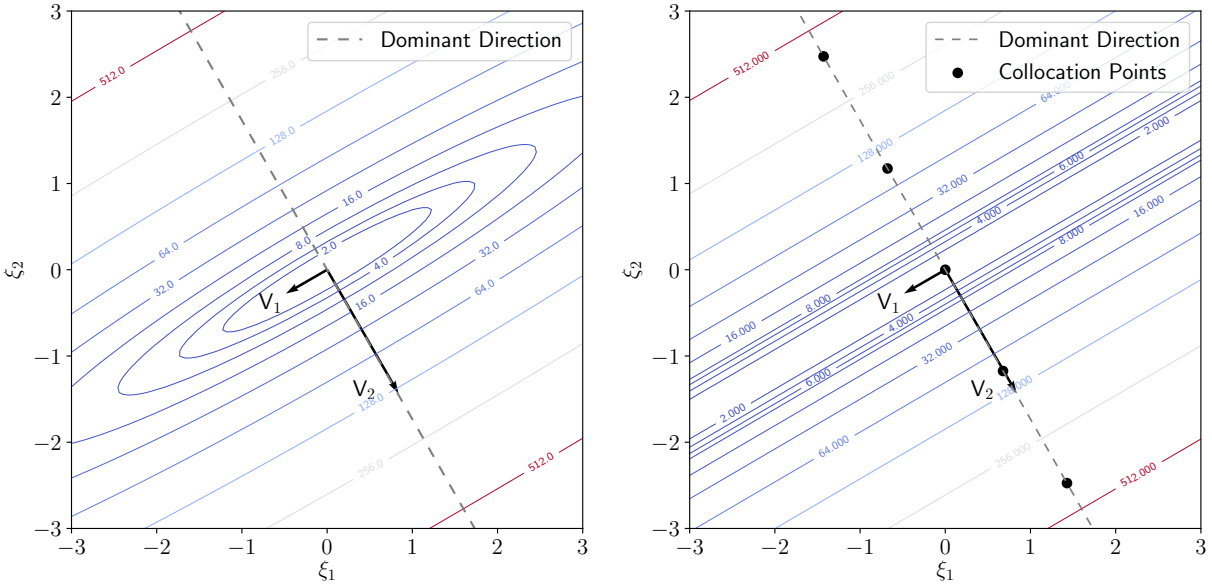
We reiterate the importance of using the isoprobabilistic space for computing the dominant directions. This is illustrated in Fig. 2, where we plot the QoI contours with respect to the original random variables, ξ, and isoprobabilistic random variables, $\bar{\xi}$. For the case when $\sigma = [2, 0.1]$, the dominant direction, V_2, is significantly different when compared to V_2 for the same QoI in Fig. 1a, where the random variables have the same standard deviation.

The proposed method is summarized in Algorithm 1 for multi-dimensional tensor-product stochastic collocation with equal numbers of quadrature points in every dimension. We would like to draw the reader's attention to line 12 in Algorithm 1, where the QoI is evaluated at a realization

$$\mu_\xi + \sqrt{2\Sigma}\mathsf{V}_{dominant}\bar{\xi}.$$

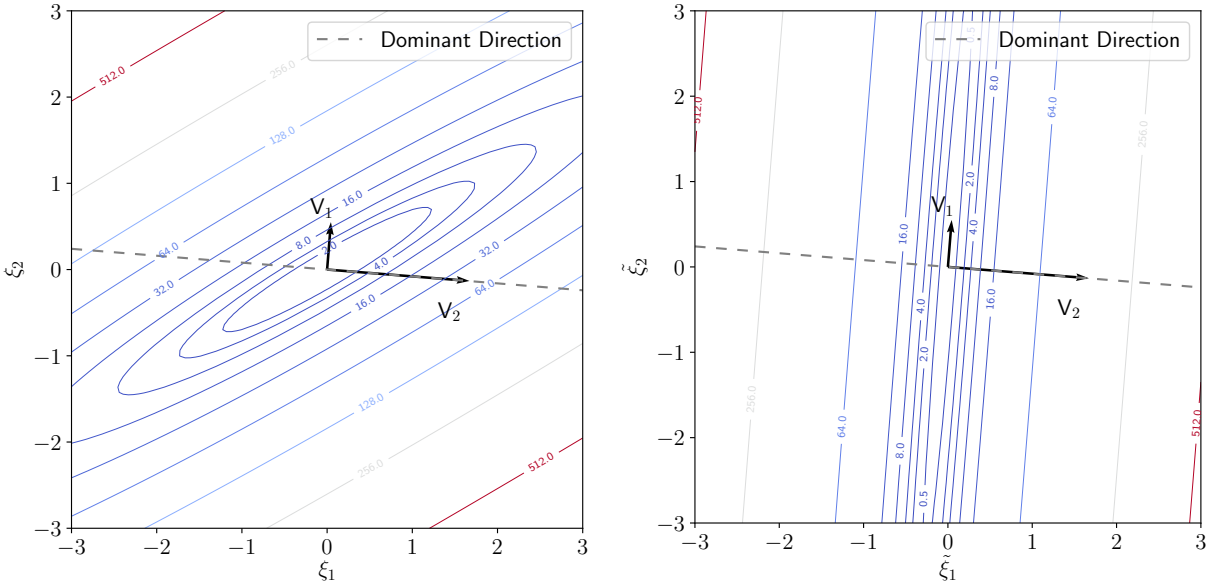
Similar to Eq. (6), this requires scaling $\mathsf{V}_{dominant}\bar{\xi}$, in the uncorrelated random variable space by a factor of $\sqrt{2\Sigma}$. For this example, we assume that the random variables are uncorrelated, so $\sqrt{2\Sigma} = \mathsf{diag}(\sqrt{2\sigma_i})$. If the random variables are correlated, then they must be transformed into an uncorrelated space for collocation. This is acheived by spectral decomposition of the covariance matrix [37, 38].

Finally, the proposed dimension reduction is independent of the (non-intrusive) uncertainty propagation scheme. For example, one can use a different number of collocation points in each dominant direction. In addition, it is possible to use a completely different propagation method, such as Monte Carlo simulation, in the subspace defined by $\mathsf{V}_{dominant}$.



(a) Original random variable space **(b) Resulting approximation of J**

Figure 1 Eigenvectors, $[\mathsf{V}_1, \mathsf{V}_2]$, of $\nabla^2_\xi J$ and the consequent dominant direction used for stochastic collocation when ξ has a standard normal distribution.



(a) Original random variable space **(b) Isoprobabilistic space**

Figure 2 Contour plot of $J(\xi)$, and the corresponding dominant direction when ξ has $\sigma = [2, 0.1]$. $[\mathsf{V}_1, \mathsf{V}_2]$ are the eigenvectors of $\nabla^2_\xi J$ in the isoprobabilistic space.

A. Arnoldi Iteration

Algorithm 1 requires the eigenvalues and eigenvectors of the QoI Hessian with respect to the uncertain variables. For the differential-equation based simulations we are interested in, computing the Hessian of J requires n additional linearized solutions [39, 40], which is impractical for large-scale problems. Therefore, we need a computationally inexpensive method of approximating the spectrum (i.e. eigenvalues) and eigenvectors of the Hessian.

To this end, we rely on Arnoldi's method [41], which is a well known Krylov subspace method for spectral analysis [42, 43]. Estimating the eigenmodes of the Hessian using Arnoldi's method requires Hessian-vector products, $(\nabla^2_\xi J)z_j$, where z_j is an element of an orthonormal basis defined by Arnoldi's method. This matrix-vector product can be approximated by taking the directional derivative of the gradient using a forward finite-difference approximation:

$$(\nabla^2_\xi J)z_j \approx \frac{\nabla_\xi J(\xi + \alpha z_j) - \nabla_\xi J(\xi)}{\alpha}, \tag{10}$$

where $\alpha > 0$ is a finite step length. A modified Arnoldi's method that uses this finite-difference approximation is provided in Algorithm 2.

At the end of its mth iteration, Arnoldi's method produces an $m \times m$ upper Hessenberg matrix, H_m. The eigenvalues of the symmetric part of H_m provide good estimates for the dominant eigenvalues of $\nabla^2_\xi J$ [44]. The corresponding Ritz-approximate eigenvectors of $\nabla^2_\xi J$ can be obtained by multiplying the eigenvectors of the symmetric part of H_m with the orthonormal bases $\mathsf{Z}_m = [z_1, z_2, \ldots, z_m]$.

Since the Hessian is a symmetric matrix, the astute reader may wonder why we are not using the Lanczos algorithm, which is the preferred method for spectral analysis of symmetric matrices. We use Arnoldi's method because the finite-difference approximations introduce errors into the Hessian-vector products, and these errors effectively perturb the Hessian and lead to a loss of symmetry. The symmetric part of H_m is used to help recover symmetry and obtain strictly real eigenvalue approximations.

The reader may also notice that the orthonormal basis in Arnoldi's method is seeded with the negative of the normalized gradient. We use the negative gradient due to historical reasons related to Newton's method for optimization, where a linear system of the form

$$\left(\nabla^2_\xi J\right) p = -\nabla_\xi J \tag{11}$$

is solved. p in Eq. (11) is a trial step in the optimization routine. We direct the reader to [44] for more information on the use of Arnoldi's method for optimization.

IV. Results

A. 2D Quadratic Problem

In order to investigate the proposed method, we compare the mean computed by Algorithm 1 against the exact value for the 2D quadratic problem described in Eq. (8) at $\mu_\xi = [0, 0]^T$. The goal of this study is to investigate the impact of numerically integrating only along the reduced stochastic space. Given that the problem is only 2 dimensional, the exact Hessian of the QoI is used, i.e. we do not use Arnoldi's method in this case. The relative approximation error is computed as

$$\epsilon = \left| \frac{\bar{\mu}_J - \mu_J}{\mu_J} \right|, \tag{12}$$

where $\bar{\mu}_J$ is the approximated mean. We evaluate the accuracy of our method for $\theta \in [0, 90°]$, and standard deviations ranging between 0.1 and 2 for the random variables, by identifying the maximum relative error for values of μ_ξ between $[-2, -2]$ and $[2, 2]$ at every θ and standard deviation ratio.

Consider Fig. 3, where the approximation error is plotted as a function of the rotation angle, θ, for a given standard deviation. The rotation angle enables us to study the effect of the relative orientation of the random variable axes compared to the axes of the contours of $\nabla^2_\xi J$; in particular, $\theta = 0°$ means that the dominant eigenvector of the Hessian is aligned with the random variable with the highest standard deviation, while $\theta = 90°$ means that the dominant eigenvector is aligned with the variable with the lowest standard deviation. As can be seen in Fig. 3a, the approximation error remains constant with variation in θ when the two standard deviations are equal. However, Fig. 3b shows that the

Algorithm 2: Arnoldi's Method with Finite-Difference-Based Products

Data: $m > 0$, $\alpha > 0$, ξ^0, and
$g_0 = \nabla_\xi J(\xi^0)$
Result: approximate eigenvalues, $\Lambda_m = [\lambda_1, \lambda_2, \ldots, \lambda_m]$, and
approximate eigenvectors, $\mathsf{V}_m = [v_1, v_2, \ldots, v_m]$

set $z_1 = -g_0/\|g_0\|$
for $j = 1, 2, \ldots, m$ **do**
  evaluated perturbed gradient, $g_j = \nabla_\xi J(\xi^0 + \alpha z_j)$
  compute $z_{j+1} = (g_j - g_0)/\alpha$
  for $i = 1, \ldots, j$ **do** (Modified Gram-Schmidt)
    $h_{i,j} = z_{j+1}^T z_i$
    $z_{j+1} \leftarrow z_{j+1} - h_{i,j} z_i$
  end
  compute $h_{j+1,j} = \|z_{j+1}\|$
  if $\|h_{j+1,j}\| = 0$ **then** (check for breakdown)
    set $m = j$ and break
  end
  $z_{j+1} \leftarrow z_{j+1}/h_{j+1,j}$
end
compute eigen-decomposition of the symmetric part of H_m, i.e. $\frac{1}{2}\left[\mathsf{H}_m + \mathsf{H}_m^T\right]\tilde{\mathsf{V}}_m = \tilde{\mathsf{V}}_m \Lambda_m$.
compute the approximate eigenvectors $\mathsf{V}_m = \mathsf{Z}_m \tilde{\mathsf{V}}_m$

approximation error increases when the dominant eigenvector of the Hessian is not aligned with the random variable with the highest standard deviation.

A more comprehensive study of the effects of the standard deviations is given in Fig. 4a where, for $\theta = 90°$, the approximation error first increases with increase in the ratio of $\frac{\sigma_1}{\sigma_2}$ before decreasing. This is because the influence of the standard deviations on the dominant direction exceeds that of the QoI when $\frac{\sigma_1}{\sigma_2} \approx 7.33$, resulting in a dominant direction that tries to align itself with the direction with the higher standard deviation, ξ_1. As a consequence, the dominant direction starts to become more aligned with the variable with the highest standard deviation. Failing to account for the standard deviation by computing $\nabla^2_\xi J$ in the original random variable space, instead of the isoprobabilistic space, results in a poorer approximation of μ_J as $\frac{\sigma_1}{\sigma_2}$ increases. This behavior is shown in Fig 4b.

B. Multidimensional Synthetic Quadratic

Next we apply Algorithm 1 in higher dimensions. In order to do so in a systematic manner, we need to have control over the test problem on which we verify our method. To this end, we define a multidimensional quadratic problem, henceforth referred to as the Hadamard Quadratic, given by

$$J(\xi) = \xi^T \mathsf{V} \Lambda \mathsf{V}^T \xi, \tag{13}$$

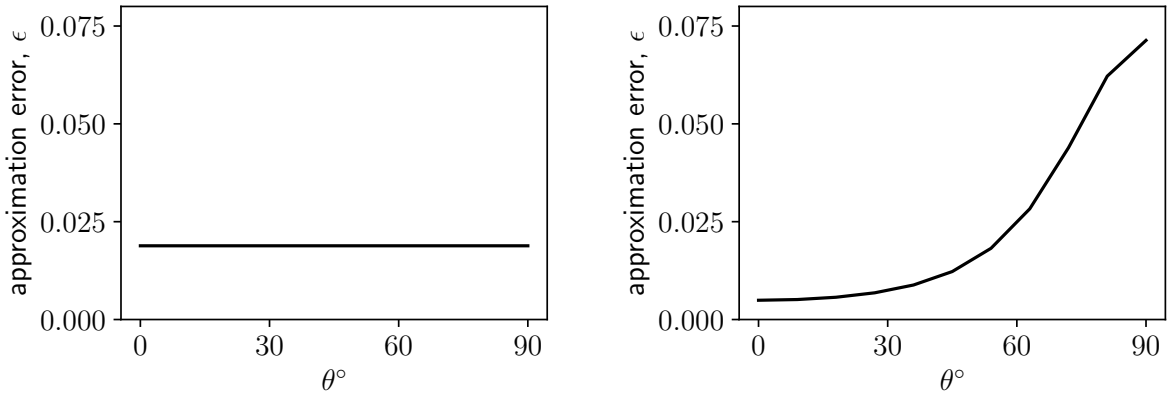
where V is an orthonormalized Hadamard matrix whose columns represent synthetic eigenvectors. Λ is a diagonal matrix that contains synthetic eigenvalues with a predetermined decay rate given by

$$\Lambda_{i,i} = \frac{1}{i^r},$$

where r is a user defined exponent; we consider values of $r = \frac{1}{2}$, 1, or 2. Since $J(\xi)$ is a quadratic, μ_J can be computed analytically as

$$\mu_J = \mathrm{trace}(\mathsf{V}\Lambda\mathsf{V}^T\Sigma) + \mu_\xi^T \mathsf{V}\Lambda\mathsf{V}^T \mu_\xi. \tag{14}$$

Recognizing that Hessian computation and the subsequent eigen decomposition is prohibitively difficult for large-scale problems, we now include the modified Arnoldi iteration, shown in Algorithm 2, to estimate the most dominant



(a) $\sigma = [0.1, 0.1]$ **(b)** $\sigma = [0.2, 0.1]$

Figure 3 Maximum relative approximation error for different orientation of V and different values of σ using Algorithm 1.

eigenmodes. This raises two important questions; how many Arnoldi iterations should be carried out, and how many dominant directions should be considered for uncertainty propagation?

Since a tensor product multidimensional stochastic collocation scheme becomes computationally infeasible when applied to more than 10 dimensions in serial, we use Algorithm 2 to accurately estimate the 10 most dominant directions. To assess the accuracy of the 10 eigenvalues obtained from Arnoldi's method, we compare them with the exact eigenvalues using the following metric:

$$\tau = \frac{||\lambda_{1:10} - \lambda_{1:10}^{exact}||_2}{||\lambda_{1:10}^{exact}||_2},$$

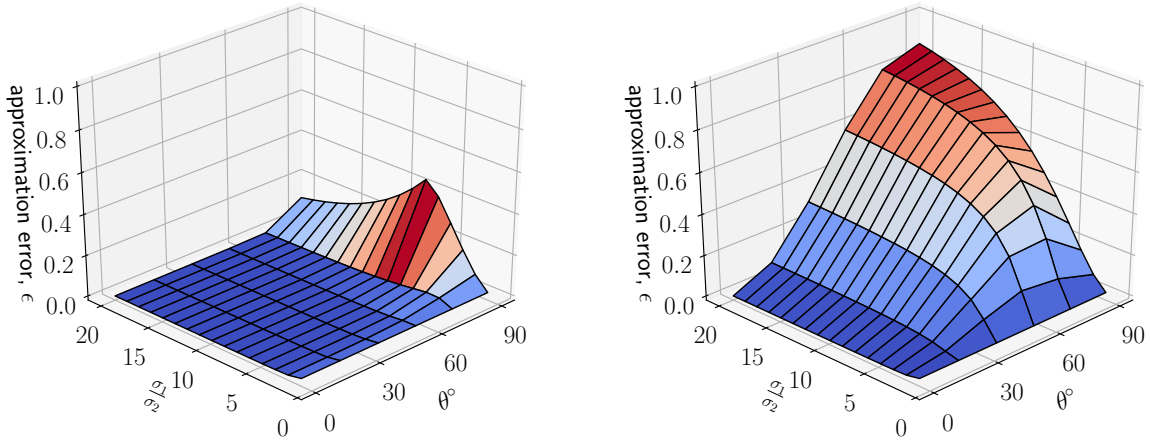
where $\lambda_{1:10}$ denotes a vector of the 10 estimated eigenvalues with largest magnitude, and similarly for the exact eigenvalutes in $\lambda_{1:10}^{\text{exact}}$.

In order to obtain the statistical performance of the algorithm over a range of random variable variances, we run ten instances of the Hadamard quadratic of a given dimension and eigenvalue decay rate, using a uniform random number generator between 0 and 1 to generate values of μ_ξ and σ. Fig. 5 shows the accuracy of the estimated eigenvalues. The blue dot shows the average error over the ten instances and the error bars indicate the maximum and minimum of those errors.

Arnoldi's method does an excellent job of estimating the dominant modes when the eigenvalues decay rapidly, as seen for the case of $\lambda_i = \frac{1}{i^2}$. On the other hand, it requires roughly 40 iterations for $\lambda_i = \frac{1}{\sqrt{i}}$ to achieve $\tau < O(10^{-4})$. Notice that the convergence of the eigenvalues with Arnoldi's method is not affected by the number of random variables in the problem, as it takes the same number of iterations to achieve reasonable accuracy in eigenvalues across all three dimensions considered (64, 128, 256).

Next, we look at the quality of our approximate mean, $\bar{\mu}_J$, as a function of number of directions used for uncertainty propagation. To ensure that the first ten eigenvalues are accurately estimated using Arnoldi's method, a maximum of $m = 71$ Arnoldi iterations was allowed for this test; for problems with fewer than 71 dimensions, the Arnoldi iteration terminated when $m = n$, otherwise it terminated after 71 iterations. Again, in addition to generating μ_ξ and σ randomly, we run ten instances of every case to avoid statistical aberrations. The corresponding results for our approximation error from Eq. (12) are shown in Fig 6.

As expected, our proposed method performs well when the eigenmodes of the QoI decay rapidly. For all cases, the approximation improves as the number of directions used for collocation increases, although the rate of improvement in the approximation decreases when the QoI Hessian has slowly decaying eigenvalues. We also find that as the number of random variables increases, discarding invariant dimensions results in diminishing effects on the error in our approximation. This is because the majority of the nonlinearity is captured by the dominant directions and the remaining invariant directions have a very small contribution to the QoI mean value.



(a) $V_{dominant}$ computed in isoprobabilistic space **(b) $V_{dominant}$ computed in original random variable space**

Figure 4 Variation of relative approximation error with $\frac{\sigma_1}{\sigma_2}$ and θ.

V. Summary and Conclusions

The ability to accurately compute statistical moments with a computationally tractable framework is critical for robust optimization of large problems. In this paper, we have proposed a method that aims to achieve this capability by exploiting the characteristics of the problem at hand. In addition, propagating uncertainties using non-intrusive stochastic collocation allows us to extend our method for use with existing software libraries for uncertainty analysis.

Using Hessian information allows us to identify the most nonlinear directions of a QoI, so we can focus effort on these directions instead of allocating valuable computational resources in directions with minimal impact on the QoI statistics. We also demonstrated the importance of considering the probability distribution of the input random variables by identifying the most nonlinear direction in the isoprobabilistic space, which helps us to identify the best compromise in creating an approximation.

Accurate Hessian information for the QoI is not always available or is expensive to compute. However, accurate gradients are becoming increasingly available in analysis codes. This is where Arnoldi's method proves useful: we employed a modified Arnoldi's method that uses the QoI gradient information to generate a reasonably accurate approximation to the eigenmodes of the Hessian.

Finally, we demonstrated that our method can be applied to problems with up to 256 random variables and rapidly decaying eigenmodes. This is significant as it brings us closer to performing robust design optimization on practical problems which, in general, contain many uncertain parameters.

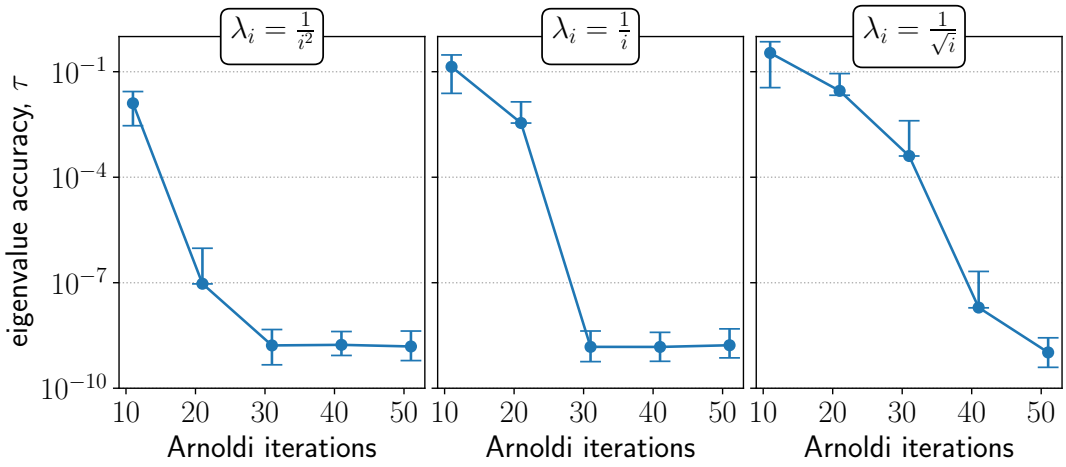
Acknowledgements

This work was supported in part by the U.S. Air Force Research Laboratory (AFRL) under the Michigan-AFRL Collaborative Center in Aerospace Vehicle Design (CCAVD), with Dr. Philip Beran as the task Technical Monitor. The computing infrastructure for this work was provided by RPI's Scientific Computation Research Center. The authors gratefully acknowledge this support.

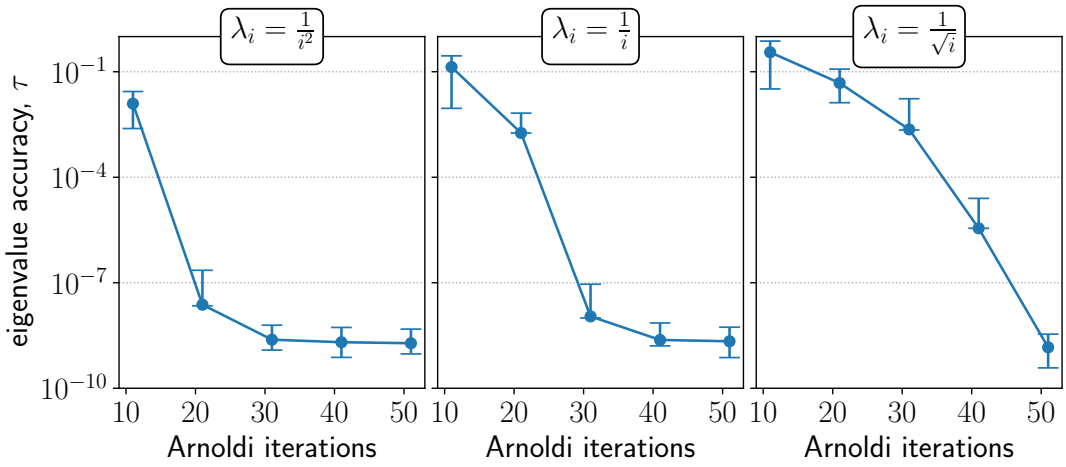
References

[1] Keane, A., and Nair, P., *Computational Approaches for Aerospace Design: The Pursuit of Excellence*, 1st ed., Wiley, West Sussex, England, 2005, Chap. 8, pp. 327–357.

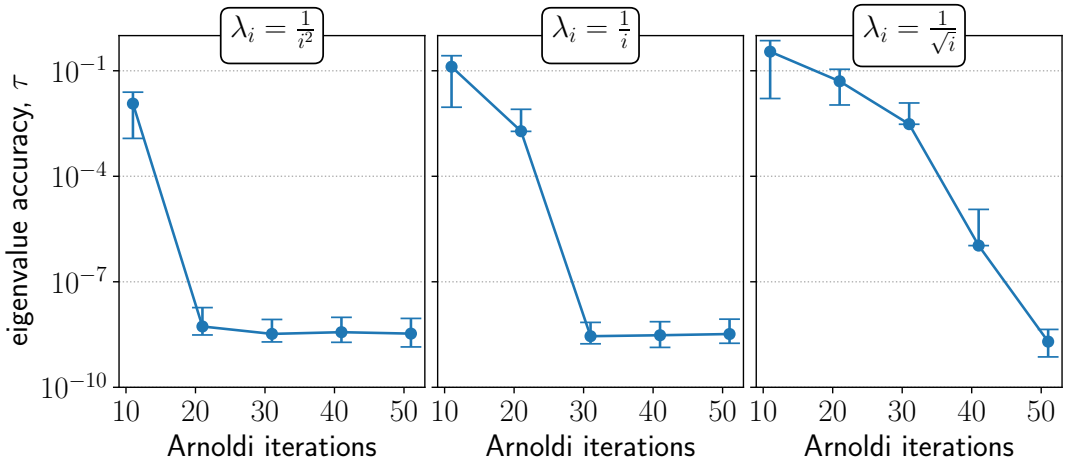
[2] Chen, W., Allen, J. K., Tsui, K.-L., and Mistree, F., "A Procedure for Robust Design: Minimizing Variations Caused by Noise



(a) Problem dimension, n = 64

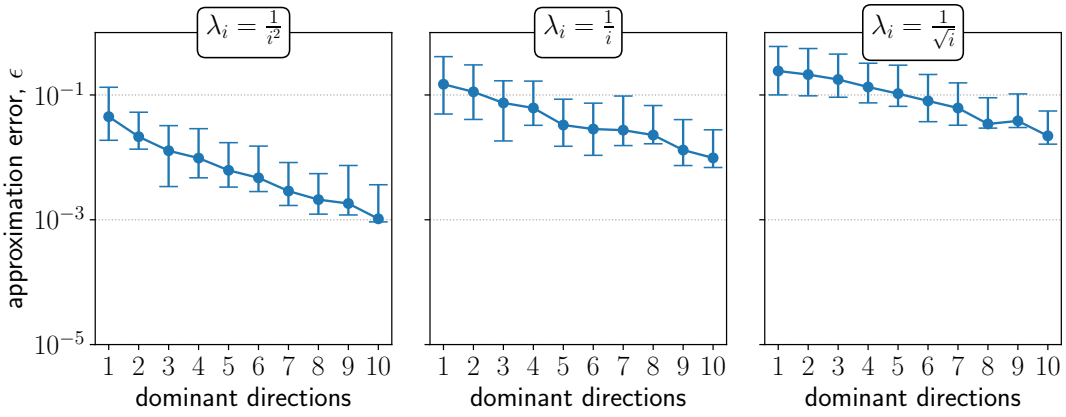


(b) Problem dimension, n = 128

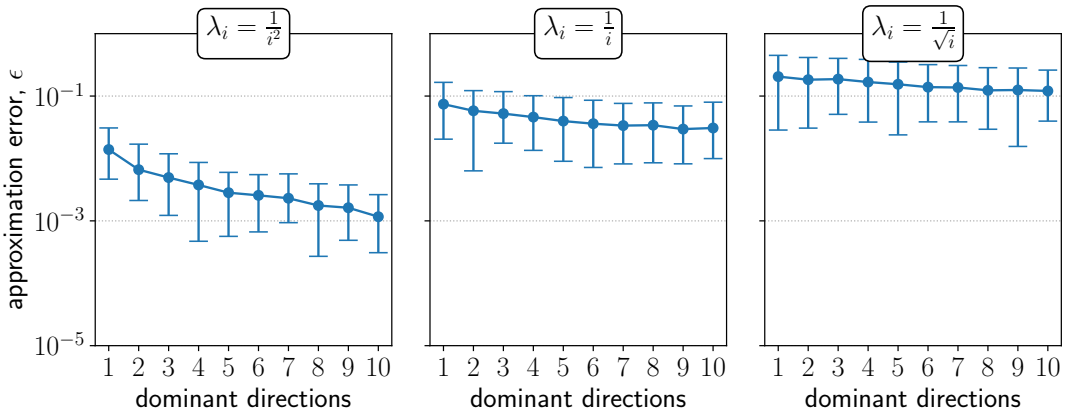


(c) Problem dimension, n = 256

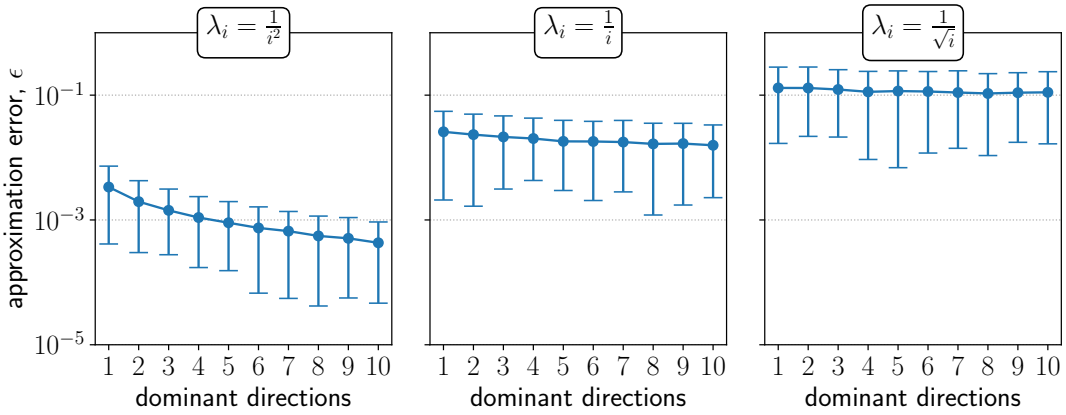
Figure 5 Eigenvalue accuracy, τ, with respect to number of Arnoldi iterations



(a) Problem dimension, n = 16



(b) Problem dimension, n = 64



(c) Problem dimension, n = 256

Figure 6 Approximation error, ϵ, for Hadamard quadratic as a function of number of dominant directions used for uncertainty propagation.

Factors and Control Factors," *Journal of Mechanical Design*, Vol. 118, No. 4, 1996, pp. 478–485. doi:10.1115/1.2826915, URL http://dx.doi.org/10.1115/1.2826915.

[3] Taguchi, G., *System of Experimental Design*, UNIPUB/Kraus International Publications, New York, 1987.

[4] Phadke, M., *Quality Engineering Using Robust Design*, Prentice Hall, Englewood Cliffs, NJ, 1989.

[5] Smith, R., *Uncertainty Quantification: Theory, Implementation, and Applications*, SIAM, 2013, Chap. 1, pp. 1–10.

[6] Eldred, M., Giunta, A., Wojtkiewicz, S., and Trucano, T., "Formulations for Surrogate-Based Optimization Under Uncertainty," American Institute of Aeronautics and Astronautics, 2002. doi:10.2514/6.2002-5585, URL https://doi.org/10.2514/6.2002-5585.

[7] Ghate, D., and Giles, M., *Inexpensive Monte Carlo uncertainty analysis*, Tata McGraw-Hill, New Delhi, 2006, pp. 203–210.

[8] Wright, M. J., Bose, D., and Chen, Y.-K., "Probabilistic Modeling of Aerothermal and Thermal Protection Material Response Uncertainties," *AIAA Journal*, Vol. 45, No. 2, 2007, pp. 399–410. doi:10.2514/1.26018, URL https://doi.org/10.2514/1.26018.

[9] Perez, R., "Uncertainty Analysis of CFD Via Polynomial Chaos," Ph.D. thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA, 2008.

[10] Chalot, F., Dinh, Q., Herbin, E., Martin, L., Ravachol, M., and Roge, G., "Estimation of the Impact of Geometrical Uncertainties on Aerodynamic Coefficients Using CFD," American Institute of Aeronautics and Astronautics, 2008. doi:10.2514/6.2008-2068, URL https://doi.org/10.2514/6.2008-2068.

[11] Chantrasmi, T., Doostan, A., and Iaccarino, G., "Padé–Legendre approximants for uncertainty analysis with discontinuous response surfaces," *Journal of Computational Physics*, Vol. 228, No. 19, 2009, pp. 7159 – 7180. doi:https://doi.org/10.1016/j.jcp.2009.06.024, URL http://www.sciencedirect.com/science/article/pii/S002199910900326X.

[12] Chung, H.-S., and Alonso, J., "Using gradients to construct cokriging approximation models for high-dimensional design optimization problems," American Institute of Aeronautics and Astronautics, 2002. doi:10.2514/6.2002-317, URL https://doi.org/10.2514/6.2002-317.

[13] Laurenceau, J., and Sagaut, P., "Building Efficient Response Surfaces of Aerodynamic Functions with Kriging and Cokriging," *AIAA Journal*, Vol. 46, No. 2, 2008, pp. 498–507. doi:10.2514/1.32308, URL https://doi.org/10.2514/1.32308.

[14] Yamazaki, W., Rumpfkeil, M., and Mavriplis, D., "Design Optimization Utilizing Gradient/Hessian Enhanced Surrogate Model," American Institute of Aeronautics and Astronautics, 2010. doi:10.2514/6.2010-4363, URL https://doi.org/10.2514/6.2010-4363.

[15] Rumpfkeil, M., Yamazaki, W., and Dimitri, M., "A Dynamic Sampling Method for Kriging and Cokriging Surrogate Models," American Institute of Aeronautics and Astronautics, 2011. doi:10.2514/6.2011-883, URL https://doi.org/10.2514/6.2011-883.

[16] Rumpfkeil, M. P., "Optimizations Under Uncertainty Using Gradients, Hessians, and Surrogate Models," *AIAA Journal*, Vol. 51, No. 2, 2012, pp. 444–451. doi:10.2514/1.J051847, URL https://doi.org/10.2514/1.J051847.

[17] Rumpfkeil, M. P., "Robust design under mixed aleatory/epistemic uncertainties using gradients and surrogates," *Journal of Uncertainty Analysis and Applications*, Vol. 1, No. 1, 2013, p. 7. doi:10.1186/2195-5468-1-7, URL https://doi.org/10.1186/2195-5468-1-7.

[18] Constantine, P. G., Dow, E., and Wang, Q., "Active Subspace Methods in Theory and Practice: Applications to Kriging Surfaces," *SIAM Journal on Scientific Computing*, Vol. 36, No. 4, 2014, pp. A1500–A1524. doi:10.1137/130916138, URL https://doi.org/10.1137/130916138.

[19] Dwight, R., and Han, Z.-H., "Efficient Uncertainty Quantification Using Gradient-Enhanced Kriging," American Institute of Aeronautics and Astronautics, 2009. doi:10.2514/6.2009-2276, URL https://doi.org/10.2514/6.2009-2276.

[20] Hammersley, J., and Handscomb, D., *Monte Carlo Methods*, Chapman and Hall, New York, 1964.

[21] Helton, J. C., and Davis, F., "Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems," *Reliability Engineering & System Safety*, Vol. 81, No. 1, 2003, pp. 23 – 69. doi:https://doi.org/10.1016/S0951-8320(03)00058-9, URL http://www.sciencedirect.com/science/article/pii/S0951832003000589.

[22] Tari, M., and Dahmani, A., "Refined descriptive sampling: A better approach to Monte Carlo simulation," *Simulation Modelling Practice and Theory*, Vol. 14, No. 2, 2006, pp. 143 – 160. doi:https://doi.org/10.1016/j.simpat.2005.04.001, URL http://www.sciencedirect.com/science/article/pii/S1569190X0500047X.

[23] Caflisch, R. E., "Monte Carlo and quasi-Monte Carlo methods," *Acta Numerica*, Vol. 7, 1998, p. 1–49. doi:10.1017/S0962492900002804.

[24] Beyer, H.-G., and Sendhoff, B., "Robust optimization – A comprehensive survey," *Computer Methods in Applied Mechanics and Engineering*, Vol. 196, No. 33, 2007, pp. 3190 – 3218. doi:https://doi.org/10.1016/j.cma.2007.03.003, URL http://www.sciencedirect.com/science/article/pii/S0045782507001259.

[25] Mathelin, L., and Hussaini, M. Y., "A stochastic collocation algorithm for uncertainty analysis," Tech. Rep. CR-2003-212153, NASA, 2003.

[26] Mathelin, L., Hussaini, M. Y., and Zang, T. A., "Stochastic approaches to uncertainty quantification in CFD simulations," *Numerical Algorithms*, Vol. 38, No. 1, 2005, pp. 209–236. doi:10.1007/BF02810624, URL https://doi.org/10.1007/BF02810624.

[27] Eldred, M., and Burkardt, J., "Comparison of Non-Intrusive Polynomial Chaos and Stochastic Collocation Methods for Uncertainty Quantification," American Institute of Aeronautics and Astronautics, 2009. doi:10.2514/6.2009-976, URL https://doi.org/10.2514/6.2009-976.

[28] Padulo, M., Campobasso, M. S., and Guenov, M. D., "Novel Uncertainty Propagation Method for Robust Aerodynamic Design," *AIAA Journal*, Vol. 49, No. 3, 2011, pp. 530–543. doi:10.2514/1.J050448, URL https://doi.org/10.2514/1.J050448.

[29] Smith, R., *Uncertainty Quantification: Theory, Implementation, and Applications*, SIAM, 2013, Chap. 10, pp. 207–237.

[30] Smolyak, S., "Quadrature and interpolation formulas for tensor products of certain classes of functions," *Soviet Mathematics, Doklady*, Vol. 4, 1963, pp. 240–243.

[31] Nobile, F., Tempone, R., and Webster, C. G., "A Sparse Grid Stochastic Collocation Method for Partial Differential Equations with Random Input Data," *SIAM Journal on Numerical Analysis*, Vol. 46, No. 5, 2008, pp. 2309–2345. doi:10.1137/060663660.

[32] Nobile, F., Tempone, R., and Webster, C. G., "An Anisotropic Sparse Grid Stochastic Collocation Method for Partial Differential Equations with Random Input Data," *SIAM Journal on Numerical Analysis*, Vol. 46, No. 5, 2008, pp. 2411–2442. doi:10.1137/070680540.

[33] Gerstner, T., and Griebel, M., "Numerical integration using sparse grids," *Numerical Algorithms*, Vol. 18, No. 3, 1998, p. 209. doi:10.1023/A:1019129717644, URL https://doi.org/10.1023/A:1019129717644.

[34] Barthelmann, V., Novak, E., and Ritter, K., "High dimensional polynomial interpolation on sparse grids," *Advances in Computational Mathematics*, Vol. 12, No. 4, 2000, pp. 273–288. doi:10.1023/A:1018977404843, URL https://doi.org/10.1023/A:1018977404843.

[35] Frauenfelder, P., Schwab, C., and Todor, R. A., "Finite elements for elliptic problems with stochastic coefficients," *Computer Methods in Applied Mechanics and Engineering*, Vol. 194, No. 2, 2005, pp. 205 – 228. doi:https://doi.org/10.1016/j.cma.2004.04.008, URL http://www.sciencedirect.com/science/article/pii/S0045782504003299.

[36] Xiu, D., and Hesthaven, J. S., "High-Order Collocation Methods for Differential Equations with Random Inputs," *SIAM Journal on Scientific Computing*, Vol. 27, No. 3, 2005, pp. 1118–1139. doi:10.1137/040615201.

[37] Jäckel, P., *Monte Carlo methods in finance*, Wiley and Sons, Chichester, West Sussex, England, 2002.

[38] Jäckel, P., "A note on multivariate Gauss-Hermite quadrature," 2005. http://jaeckel.16mb.com/ANoteOnMultivariateGaussHermiteQuadrature.pdf[retrieved 07 Nov 2017].

[39] Rumpfkeil, M. P., and Mavriplis, D. J., "Efficient Hessian Calculations Using Automatic Differentiation and the Adjoint Method with Applications," *AIAA Journal*, Vol. 48, No. 10, 2010, pp. 2406–2417. doi:10.2514/1.j050451, URL https://doi.org/10.2514/1.j050451.

[40] Griewank, A., and Walther, A., *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation, Second Edition*, Society for Industrial and Applied Mathematics, 2008.

[41] Arnoldi, W. E., "The Principle of Minimized Iterations in the Solution of the Matrix Eigenvalue Problem," *Quart. Appl. Math.*, Vol. 9, 1951, pp. 17–29.

[42] Saad, Y., "A Flexible Inner-Outer Preconditioned GMRES Algorithm," *SIAM Journal on Scientific Computing*, Vol. 14, No. 2, 1993, pp. 461–469. doi:10.1137/0914028, URL https://doi.org/10.1137/0914028.

[43] Saad, Y., *Iterative Methods for Sparse Linear Systems*, 2nd ed., Society for Industrial and Applied Mathematics, 2003. doi:10.1137/1.9780898718003, URL http://epubs.siam.org/doi/abs/10.1137/1.9780898718003.

[44] Hicken, J., and Ashley, A., "Arnoldi-based Sampling for High-dimensional Optimization using Imperfect Data," American Institute of Aeronautics and Astronautics, 2015. doi:10.2514/6.2015-2943, URL https://doi.org/10.2514/6.2015-2943.