Simplex Elements Stochastic Collocation for Uncertainty Propagation in Robust Design Optimization

Jeroen A.S. Witteveen,∗ Gianluca Iaccarino†

Center for Turbulence Research, Stanford University, Building 500, Stanford, CA 94305–3035, USA

Robust design optimization has become essential to make high–performance aerospace designs insensitive to uncertainties in the environment and the design parameters. The uncertainty propagation step is often computationally the most intensive in this process, which limits the number of uncertainties that can be taken into account. In this paper, a Simplex Elements Stochastic Collocation (SESC) uncertainty propagation method is developed that matches the robustness of Monte Carlo (MC) simulation and that achieves a constant order of convergence $O(p)$ with an increasing number of uncertainties, which is larger than the $O(1)$ of MC. This is accomplished by increasing the polynomial interpolation degree $p$ with increasing dimensionality according to $p = O(n_ξ) - 1$, with $n_ξ$ the number of uncertain parameters. Accurate error estimates are derived that are used as adaptive refinement measure and refinement stopping criterion for non–uniform distributions and unbounded parameter domains.

I. Introduction

A utomatic design optimization has become a vital aspect in aerodynamic design owing to the increasing market competition in aerospace industry. Deterministic optimization tools are widely used based on, for example, adjoint gradient methods, genetic algorithms, or game theory. However, in practice, aeronautical designs do not operate exactly at their design point due to physical variability in the environment caused by, for example, atmospheric turbulent fluctuations. Production tolerances also result in aerodynamic shapes that slightly deviate from the design geometry. These small variations can deteriorate the performance of deterministically optimized designs. It is, therefore, necessary to account for these uncertainties in the optimization process using robust design optimization. Uncertainty Quantification (UQ) is then used in the optimization loop instead of a deterministic simulation.

The UQ process consists of characterization, propagation, and evaluation of the uncertainties in the physical system. The characterization step identifies the relevant sources of uncertainty in terms of their probability distributions. These input uncertainties are then propagated through the computational model to determine their effect on design quantities of interest. This is computationally the most intensive step of the UQ pipeline, which often limits the number of uncertainties that can be taken into account. This is especially true for robust design optimization, in which the uncertainty propagation is performed inside an optimization loop. The last step is the evaluation of the probability distributions and statistical moments of the outputs of interest.

A method that is widely used for the uncertainty propagation step is Monte Carlo (MC) simulation, which is based on performing an ensemble of deterministic computations for randomly sampled parameter values. It is generally acknowledged to be the most robust uncertainty propagation method available, which is an essential property for its potential use in robust design optimization. It is also a popular method

∗ Postdoctoral Fellow, Member AIAA, Phone: +1 650 723 9601, Fax: +1 650 723 9617, jasw@stanford.edu.
† Assistant Professor, Member AIAA, jops@stanford.edu.
because its rate of convergence is independent of the number of random input parameters. However, this constant rate of convergence is only of the order $O = \frac{1}{2}$, which is a condensed notation for $O(n_{s}^{1/2})$ with $n_{s}$ the number of samples. This means that for decreasing the error by a factor 10, the number of samples has to be increased by a factor 100. This property makes MC computationally impractical for the aerodynamic design process, since deterministic computational fluid dynamics (CFD) simulations can already be computationally intensive.

The Stochastic Collocation (SC) method has been developed as a more efficient alternative to MC. It is a Stochastic Galerkin method based on sampling in selected Gauss quadrature points in probability space and Lagrangian polynomial interpolation, which can lead to spectral convergence for smooth responses. For the extension to multiple random parameters a tensor product can be used. However, this results in an exponential increase of the number of samples with dimension and a deteriorating convergence rate. Sparse grid formulations have been applied to alleviate this effect. Also separated solution approximations have been developed to achieve a linear increase of computational costs with dimension. Gauss quadrature samples have the disadvantage that they do not coincide with their higher order quadrature points, which would be useful for establishing convergence. Nested quadrature rules, such as Clenshaw–Curtis quadrature, can be used as an alternative. However, they offer little flexibility especially in higher dimensions as the number of abscissas on each stochastic coordinate axis doubles approximately from one refinement level to the next.

It is widely acknowledged that another problem is that discontinuities in the response surface or numerical noise can result in an oscillatory approximation due to the global polynomial interpolation in the SC method. Multi–element methods and shock capturing approaches have been proposed for a more robust interpolation of the samples. The multi–element approaches are usually based on employing a single–element method independently in multiple hypercube elements discretizing the probability space. For higher order interpolations these methods can still result in local oscillations and overshoots can be present even for low approximation orders. Often not all samples in an element can be reused after refinement and tensor product extensions to higher dimensions are employed, which compromises the efficiency of multi–element discretizations.

In this paper a Simplex Elements Stochastic Collocation (SESC) method is introduced for robust and efficient multi–element uncertainty quantification. It is an extension of the Adaptive Stochastic Finite Elements (ASFE) method based on Newton–Cotes quadrature sampling in simplex elements from first and second degree quadrature to higher order polynomial interpolation. The method converges by adaptively refining one element at a time, where all samples are reused after the hierarchical refinements. The refinement is automatically terminated when an error estimate reaches a stopping criterion. The number of samples is limited further by the location of a considerable number of the Newton–Cotes quadrature points on the boundaries of the simplex elements, such that samples are shared by adjacent elements. The proposed method matches the robustness properties of MC in terms of the Extremum Diminishing (ED) and Total Variation Diminishing (TVD) properties extended to probability space. It can also achieve a constant order of convergence with dimension, which is higher than the $O = \frac{1}{2}$ of MC. This is accomplished by increasing the polynomial interpolation degree with an increasing number of uncertainties. This result shows that higher order polynomial interpolation is essential for efficient uncertainty quantification in case of many input uncertainties.

The SESC formulation based on first degree Newton–Cotes quadrature is discussed in section II. Error estimates are derived and tested in application to an analytical test function. The ED and TVD robustness properties for probability space are defined in section III. The extension of SESC to higher order polynomial interpolation is introduced in section IV for the uniform distribution, the non–uniform beta distribution, and the normal distribution on an unbounded support. Conclusions are summarized in section V.

II. Simplex Elements Stochastic Collocation

First SESC with first degree Newton–Cotes quadrature is considered based on a general formulation of a multi–element uncertainty propagation method. It is applied to an analytical test function with up to three random parameters. Results are compared to error estimates, which can be used as solution–based refinement measure and refinement stopping criterion.
A. General formulation

Consider the following computational problem for output of interest \( u(x,t,\xi) \)

\[
\mathcal{L}(x,t,\xi;u(x,t,\xi)) = S(x,t,\xi),
\]

with appropriate initial and boundary conditions. Operator \( \mathcal{L} \) and source term \( S \) are defined on domain \( D \times T \times \Xi \), where \( x \in D \) and \( t \in T \) are the spatial and temporal dimensions with \( D \subset \mathbb{R}^d \), \( d = \{1,2,3\} \), and \( T = \mathbb{R}^+ \). Randomness is introduced in (1) and its initial and boundary conditions in terms of \( n_L \) uncorrelated second–order random parameters \( \xi(\omega) = \{\xi_1(\omega_1), \ldots, \xi_n(\omega_n)\} \in \Xi \) with parameter space \( \Xi \subset \mathbb{R}^{n_L} \). The symbol \( \omega = \{\omega_1, \ldots, \omega_n\} \in \Omega \subset \mathbb{R}^{n_L} \) denotes realizations in the probability space \( (\Omega,F,P) \) with \( F \subset 2^\Omega \) the \( \sigma \)-algebra of events and \( P \) a probability measure. The random variables \( \omega \) are by definition standard uniformly distributed as \( U(0,1) \). Random parameters \( \xi(\omega) \) can have any arbitrary probability density \( f_\xi(\xi) \). The argument \( \omega \) will be dropped from here on for brevity of the notation. The aim of uncertainty propagation is then to find the probability distribution of \( u(x,t,\xi) \) and its statistical moments \( \mu_{u_i}(x,t) \) given by

\[
\mu_{u_i}(x,t) = \int_{\Xi} u(x,t,\xi)^i f_\xi(\xi) d\xi.
\]

A multi–element UQ method computes these integrals as a summation of integrals over \( n_e \) non–overlapping elements \( \Xi_j \)

\[
\mu_{u_i}(x,t) = \sum_{j=1}^{n_e} \int_{\Xi_j} u(x,t,\xi)^i f_\xi(\xi) d\xi,
\]

with

\[
\sum_{j=1}^{n_e} \int_{\Xi_j} f_\xi(\xi) d\xi = 1.
\]

In SESC the integrals in the elements are estimated by approximating response surface \( u(\xi) \) by an interpolation \( w(\xi) \) of determinstic samples \( v = \{v_1, \ldots, v_n\} \), with \( n_s \) the number of samples. Here the arguments \( x \) and \( t \) are omitted for clarity of the notation. Non–intrusive uncertainty quantification method \( q \) then consists of a sampling method \( g \) and an interpolation method \( h \), for which holds \( w(\xi) = q(u(\xi)) = h(g(u(\xi))) \). The sampling method \( g \) selects the sampling points \( \xi_k \) for \( k = 1, \ldots, n_s \) and returns the sampled values \( v \), with \( v_k = h(g(u(\xi))) = u(\xi_k) \). Sample \( v_k \) is computed by solving (1) for realization \( \xi_k \) of the random parameter vector \( \xi \)

\[
\mathcal{L}(x,t,\xi_k;v_k(x,t)) = S(x,t,\xi_k).
\]

The interpolation of the samples \( w(\xi) = h(v) \) consists in SESC of a piecewise polynomial function

\[
w(\xi) = w_j(\xi), \quad \text{for } \xi \in \Xi_j,
\]

with \( w_j(\xi) \) the polynomial interpolation of order \( p \) of the samples \( v_j = \{v_{k_j,0}, \ldots, v_{k_j,n_j}\} \) at the sampling points \( \{\xi_{k_j,0}, \ldots, \xi_{k_j,n_j}\} \) in element \( \Xi_j \), where \( k_{j,l} \in \{0, \ldots, n_s\} \) for \( j = 1, \ldots, n_e \), \( l = 0, \ldots, N \), and

\[
N + 1 = \frac{(n_e + p)!}{n_e!p!}.
\]

The polynomial interpolation \( w_j(\xi) \) in element \( \Xi_j \) can then be formulated in terms of a truncated Polynomial Chaos expansion\(^7\text{–}^21\)

\[
w_j(\xi) = a_0 \Gamma_{j,0} + \sum_{i_1=1}^{n_L} a_{i_1} \Gamma_{j,1}(\xi_{i_1}) + \sum_{i_1=1,i_2=1}^{n_L} a_{i_1,i_2} \Gamma_{j,2}(\xi_{i_1},\xi_{i_2}) + \ldots + \sum_{i_1=1,i_2=1}^{n_L} \cdots \sum_{i_p=1}^{n_L} a_{i_1,\ldots,i_p} \Gamma_{j,p}(\xi_{i_1},\ldots,\xi_{i_p}),
\]

with multi–dimensional basis polynomials \( \Gamma_{j,\tilde{p}} \) of exact degree \( \tilde{p} \). Expression (8) can be written in the following more convenient shorter notation

\[
w_j(\xi) = \sum_{i=0}^{N} c_{j,i} \Psi_{j,i}(\xi),
\]
with a one–to–one correspondence between the basis polynomials \( \Gamma_{j,p} \) and \( \Psi_{j,l} \), and the coefficients \( a_{i_1 \ldots i_p} \) and \( c_{j,l} \). If the response approximation \( w_j(x, t, \xi) \) depends in addition to the random parameters \( \xi \) also on spatial \( x \) and time \( t \) coordinates, then there occurs a separation of variables in terms of the random parameters, \( \Psi_{j,l}(\xi) \), and the spatial and temporal dimensions, \( c_{j,l}(x, t) \). The polynomial coefficients \( c_{j,l} \) can be determined from\(^{11,18} \)

\[
\begin{bmatrix}
\Psi_{j,0}(\xi_{k,j,0}) & \Psi_{j,1}(\xi_{k,j,1}) & \cdots & \Psi_{j,N}(\xi_{k,j,N}) \\
\Psi_{j,0}(\xi_{k,j,1}) & \Psi_{j,1}(\xi_{k,j,1}) & \cdots & \Psi_{j,N}(\xi_{k,j,N}) \\
\vdots & \vdots & \ddots & \vdots \\
\Psi_{j,0}(\xi_{k,j,N}) & \Psi_{j,1}(\xi_{k,j,N}) & \cdots & \Psi_{j,N}(\xi_{k,j,N})
\end{bmatrix}
\begin{bmatrix}
\epsilon_{j,0} \\
\epsilon_{j,1} \\
\vdots \\
\epsilon_{j,N}
\end{bmatrix}
= \begin{bmatrix}
\psi_{k,j,0} \\
\psi_{k,j,1} \\
\vdots \\
\psi_{k,j,N}
\end{bmatrix}.
\] (10)

The piecewise polynomial approximation \( w(\xi) \) of response surface \( u(\xi) \) is eventually found by substituting (9) into (6) with \( c_{j,l} \) from (10) for \( j = 1, \ldots, n_e \) and \( l = 1, \ldots, N \). The probability distribution function and the statistical moments \( \mu_{w_i} \) of \( u(\xi) \) given by (3) are then approximated by the probability distribution and the moments \( \mu_{w_i} \) of \( w(\xi) \)

\[
\mu_{w_i}(x, t) \approx \mu_{w_i}(x, t) = \sum_{j=1}^{n_e} \int_{\Xi_j} w_j(x, t, \xi)^i f(\xi) d\xi,
\] (11)

in which the multi–dimensional integrals are evaluated using Monte Carlo integration of response surface approximation \( w(\xi) \) with \( n_{\text{MC}} \gg n_e \) samples. This is a fast operation, since it only involves sampling of piecewise polynomial function \( w(\xi) \) through (6) and no additional evaluations of the exact response \( u(\xi) \) by solving an equation similar to (5).

B. First degree Newton–Cotes quadrature

SESC uses simplex elements \( \Xi_j \) to discretize parameter space \( \Xi \), which are equivalent to triangles in a two–dimensional parameter space for \( n_\xi = 2 \) random parameters \( \xi = \{\xi_1, \xi_2\} \), see Figure 1 where standard uniformly distributed parameters \( \xi \sim U(0, 1) \) are considered. Simplex elements are used because they are the natural elements for higher–dimensional Newton–Cotes quadrature rules. For example, the first degree Newton–Cotes quadrature sampling points, shown in Figure 1 by the dots, are located in the vertices of the elements. The samples \( v \) are, therefore, shared by adjacent elements, which reduces the required number of samples compared to using fully independent elements based on, for example, Gauss quadrature in hypercube elements. The interpolation of these samples results in a straightforward piecewise linear approximation \( u(\xi) \) of the response surface \( u(\xi) \).

The initial mesh consists of samples in the vertices of the hypercube parameter space and one sample at the mean value \( \mu_\xi \) of the random parameters \( \xi \), see Figure 1a. The discretization is subsequently refined by refining one element at a time that contains the highest probability \( \Omega_j \)

\[
\hat{\Omega}_j = \int_{\Xi_j} f(\xi) d\xi.
\] (12)

The refinement can be parallelized by refining the \( n_{\text{ref}} > 1 \) elements with the highest \( \hat{\Omega}_j \) at a time. Other refinement criteria than probability \( \Omega_j \) are considered in application to the beta distribution in section IV. The element flagged for refinement is split into two elements by adding a sample at the center of the longest edge of the element \( \Xi_j \), see Figure 1b for a discretization refined to \( n_e = 32 \) elements. The choice of Newton–Cotes quadrature points has the advantage that all samples are reused after the refinement of an element owing to the hierarchical structure of Newton–Cotes quadrature.

C. Error estimates

Define the local error \( \varepsilon(\xi) \) in the response surface approximation \( w(\xi) \) with respect to the exact solution \( u(\xi) \) as

\[
\varepsilon(\xi) = w(\xi) - u(\xi).
\] (13)
µerror norm. Since SESC converges the response surface approximation \( w(\xi) \) by subsequently refining elements, error (13) can be computed exactly at the refinement of an element \( \Xi_j \) in the new sampling point \( \xi_{k,j,\text{ref}} \)

\[
\varepsilon(\xi_{k,j,\text{ref}}) = w(\xi_{k,j,\text{ref}}) - u(\xi_{k,j,\text{ref}}) = w(\xi_{k,j,\text{ref}}) - v_{k,j,\text{ref}}.
\]

This expression for the error is used below to derive error estimates for the statistical moments and an \( L_2 \) error norm. The mean values \( \mu_u \) and \( \mu_w \) of an identical Monte Carlo sampling of the exact response surface \( u(\xi) \) and its approximation \( w(\xi) \) are given by

\[
\mu_u = \frac{1}{n_{\text{MC}}} \sum_{k=1}^{n_{\text{MC}}} u(\xi_{MC,k}), \quad \mu_w = \frac{1}{n_{\text{MC}}} \sum_{k=1}^{n_{\text{MC}}} w(\xi_{MC,k}),
\]

with \( n_{\text{MC}} \) the number of Monte Carlo sampling points \( \xi_{MC,k} \). The absolute error \( \varepsilon_\mu \) between the two is

\[
\varepsilon_\mu = |\mu_w - \mu_u| = \frac{1}{n_{\text{MC}}} \sum_{k=1}^{n_{\text{MC}}} (w(\xi_{MC,k}) - u(\xi_{MC,k})),
\]

which can be written as

\[
\varepsilon_\mu = \frac{1}{n_{\text{MC}}} \left| \sum_{j=1}^{n_e} \sum_{k=1}^{n_{\text{MC},j}} (w(\xi_{MC,k,j}) - u(\xi_{MC,k,j})) \right|,
\]

with \( n_{\text{MC},j} \) the number of Monte Carlo samples in element \( \Xi_j \), \( \sum_{j=1}^{n_e} n_{\text{MC},j} = n_{\text{MC}}, \xi_{MC,k,j} \in \Xi_j, \), and \( k,j \in \{1,\ldots,n_{\text{MC}}\} \) for \( k = \{1,\ldots,n_{\text{MC}}\} \), \( j = 1,\ldots,n_e \). The unknown values \( u(\xi_{MC,k,j}) \) can now be approximated using (13) and (14) by

\[
u(\xi_{MC,k,j}) = w(\xi_{MC,k,j}) - \varepsilon(\xi_{MC,k,j}) \approx w(\xi_{MC,k,j}) - \tilde{\varepsilon}_j,
\]

where the notation \( \tilde{\varepsilon}_j = \varepsilon(\xi_{k,j,\text{ref}}) \) is used. Substituting (18) into (17) gives the following approximation of the absolute error \( \tilde{\varepsilon}_\mu \) in \( \mu_w \)

\[
\tilde{\varepsilon}_\mu = \frac{1}{n_{\text{MC}}} \sum_{j=1}^{n_e} n_{\text{MC},j} \tilde{\varepsilon}_j.
\]

Error \( \tilde{\varepsilon}_j = \varepsilon(\xi_{k,j,\text{ref}}) \) is, however, the local error at the sampling point \( \xi_{k,j,\text{ref}} \) in the response surface approximation \( w(\xi) \) before the refinement of element \( \Xi_j \). The error after refinement \( \hat{\varepsilon}_j \) can be estimated as follows

\[
\hat{\varepsilon}_j = \frac{\tilde{\varepsilon}_j}{2^n},
\]
with $p$ the polynomial order of the response surface approximation $w(\xi)$ and $n_\xi$ the number of random parameters. The factor 2 accounts for the splitting the element at the refinement into two elements, and the exponent $(p + 1)/n_\xi$ represents the actual order of convergence of the method. This relation between the polynomial degree $p$, dimensionality $n_\xi$, and order of convergence $O$ is demonstrated in section IV. The error estimate $\hat{\varepsilon}_\mu$ for $\varepsilon_\mu$ is then finally obtained by replacing $\hat{\varepsilon}_j$ in (19) by $\hat{\varepsilon}_j$ from (20)

$$\hat{\varepsilon}_\mu = \frac{1}{n_{\text{bMC}} 2^{p + 1}} \left| \sum_{j=1}^{n_{\text{bMC}}} n_{\text{MC}} \hat{\varepsilon}_j \right| = \frac{\hat{\varepsilon}_\mu}{2^{p + 1}}.$$  

(21)

This expression is expected to be a slightly conservative estimate of $\varepsilon_\mu$, since it is based on the approximation of the error uniformly in the element $\hat{\varepsilon}_j$ by the exact error $\varepsilon(\xi_{j,\text{ref}})$ in the refinement sampling point $\xi_{j,\text{ref}}$ in the center of the longest edge of the element $\Xi_j$. Closer to the vertices of the element the local error is likely to be smaller than $\varepsilon(\xi_{j,\text{ref}})$ in absolute sense. The slightly conservative nature of the error estimate $\hat{\varepsilon}_\mu$ can be considered a desirable property for using it as a refinement stopping criterion. Similarly an approximation $\hat{\varepsilon}_\sigma$ and estimate $\hat{\varepsilon}_\sigma$ of the error $\varepsilon_\sigma$ in the standard deviation

$$\varepsilon_\sigma = |\sigma_w - \sigma_u|,$$

(22)

can be derived to be

$$\hat{\varepsilon}_\sigma = \left| \sigma_w - \sqrt{\frac{1}{n_{\text{bMC}}} \sum_{j=1}^{n_{\text{bMC}}} \sum_{k=1}^{n_{\text{MC}}} (w(\xi_{MC_{j,k}}) - \hat{\varepsilon}_j)^2} - \left( \frac{1}{n_{\text{bMC}}} \sum_{j=1}^{n_{\text{bMC}}} \sum_{k=1}^{n_{\text{MC}}} (w(\xi_{MC_{j,k}}) - \hat{\varepsilon}_j) \right)^2 \right|,$$

(23)

and

$$\hat{\varepsilon}_\sigma = \left| \sigma_w - \sqrt{\frac{1}{n_{\text{bMC}}} \sum_{j=1}^{n_{\text{bMC}}} \sum_{k=1}^{n_{\text{MC}}} \left( w(\xi_{MC_{j,k}}) - \hat{\varepsilon}_j \right)^2} - \left( \frac{1}{n_{\text{bMC}}} \sum_{j=1}^{n_{\text{bMC}}} \sum_{k=1}^{n_{\text{MC}}} \left( w(\xi_{MC_{j,k}}) - \hat{\varepsilon}_j \right) \right)^2 \right|,$$

(24)

where the expression $\sigma_u = \sqrt{\mu_u - \mu_\xi}$ is used with $\mu_\xi$, defined by (2). In the same way also the corresponding relations for higher order statistical moments can be derived. The error between $w(\xi)$ and $u(\xi)$ can also be expressed in terms of the $L_2$ norm of the difference between the Monte Carlo samples of the two functions $w_{\text{MC}} = \{w(\xi_{MC_1}), \ldots, w(\xi_{MC_{n_{\text{MC}}}})\}$ and $u_{\text{MC}} = \{u(\xi_{MC_1}), \ldots, u(\xi_{MC_{n_{\text{MC}}}})\}$ as follows

$$\varepsilon_2 = \sqrt{\frac{1}{n_{\text{MC}}} \sum_{k=1}^{n_{\text{MC}}} (w(\xi_{MC_k}) - u(\xi_{MC_k}))^2}.$$  

(25)

This represents essentially the Euclidean distance between the two vectors $w_{\text{MC}}$ and $u_{\text{MC}}$ weighted by the non–uniform Monte Carlo sampling. The following error approximation $\hat{\varepsilon}_L_2$ and estimate $\hat{\varepsilon}_L_2$ can be derived for this $L_2$ error measure

$$\hat{\varepsilon}_L_2 = \sqrt{\frac{1}{n_{\text{bMC}}} \sum_{j=1}^{n_{\text{bMC}}} n_{\text{MC}_j} \hat{\varepsilon}_j^2},$$  

(26)

and

$$\hat{\varepsilon}_L_2 = \frac{1}{2^{p + 1}} \sqrt{\frac{1}{n_{\text{bMC}}} \sum_{j=1}^{n_{\text{bMC}}} n_{\text{MC}_j} \hat{\varepsilon}_j^2} = \frac{\hat{\varepsilon}_L_2}{2^{p + 1}}.$$  

(27)

The structure of the error estimate $\hat{\varepsilon}_L_2$ as a summation of non–negative terms $n_{\text{MC}_j} \hat{\varepsilon}_j^2$, makes it particularly suitable to be used as solution–based refinement criterion. The element $\Xi_j$ with the highest contribution $n_{\text{MC}_j} \hat{\varepsilon}_j^2$ to $\hat{\varepsilon}_L_2$ can then be refined at each refinement step to minimize $\hat{\varepsilon}_L_2$. This refinement criterion is applied in section IV.
D. Analytical test function

The properties of SESC are illustrated in application to the analytical test function

\[ u(\xi) = \arctan(\xi : \xi^* - \xi_1^2), \]  

(28)

with \( \xi^* = \{\xi^*_1, \ldots, \xi^*_n_\xi\} \in [0, 1]^{n_\xi} \) a vector of \( n_\xi \) arbitrary values. The random parameters \( \xi \) have standard uniform distributions \( \mathcal{U}(0, 1) \). The non-uniform beta distribution and the normal distribution with an unbounded parameter range are used in section IV. First a one-dimensional probability space with \( n_\xi = 1 \) random parameter \( \xi_1 \) is considered.

The error convergence of first degree SESC is shown in Figure 2 for the error in the approximation of the mean \( \varepsilon_\mu \) (16) and standard deviation \( \varepsilon_\sigma \) (22), and for \( L_2 \) error measure \( \varepsilon_{L_2} \) (25) as function of the number of elements \( n_e \) and the number of samples \( n_s \). The order of convergence given in the logarithmic plots is determined by a least squares linear fit through the last 5 data points in the asymptotic regime. The linear fit shown in Figure 2 is indistinguishable from the actual error. The errors \( \varepsilon_\mu, \varepsilon_\sigma, \) and \( \varepsilon_{L_2} \) are similar in absolute sense and converge with second order, \( O = 2 \), as function of the number of elements \( n_e \). An indication of the theoretical error convergence rate of MC simulation of order \( O = \frac{1}{2} \) is also given in Figure 2b for comparison. As function of the number of samples \( n_s \) the order is slightly higher than in Figure 2a, owing to the decreasing average number of samples per element \( n_s/n_e \) with increasing \( n_e \) shown in Figure 3. This effect is caused by the sharing of samples by adjacent elements. For \( n_\xi = 1 \), the ratio \( n_s/n_e \) approaches an asymptotic value of unity. In addition, the error as function of \( n_s \) is of more practical importance, since the computational costs are dominated by solving the \( n_s \) deterministic problems (5).

![Figure 2](image-url1)

(a) Function of \( n_e \)

![Figure 2](image-url2)

(b) Function of \( n_s \)

**Figure 2.** Error for \( n_\xi = 1 \) and \( p = 1 \) with the corresponding order of convergence \( O \).

![Figure 3](image-url3)

**Figure 3.** Average number of samples per element \( n_s/n_e \) for \( n_\xi = 1 \) and \( p = 1 \).
In Figure 4 error convergence as function of \( n_s \) is compared with that of the error approximations and estimates for the mean, \( \tilde{\varepsilon}_\mu (19) \) and \( \hat{\varepsilon}_\mu (21) \), standard deviation, \( \tilde{\varepsilon}_\sigma (23) \) and \( \hat{\varepsilon}_\sigma (24) \), and \( L_2 \) norm, \( \tilde{\varepsilon}_{L_2} (26) \) and \( \hat{\varepsilon}_{L_2} (26) \), respectively. The error approximations and, especially, the error estimates give an excellent prediction of the error for the mean, the standard deviation, as well as the \( L_2 \) norm. The error estimates are slightly conservative as expected, which makes them suitable to be used as a reliable stopping criterion for the refinement.

![Figure 4: Approximation and estimate of the error for \( n_s = 17 \) and \( p = 1 \) with the corresponding order of convergence \( O \).](attachment:image)

Figure 4d shows the distribution of the local error \( \varepsilon(\xi_1) \) as function of the random parameter \( \xi_1 \) for \( n_s = 17 \). It illustrates that the error is zero at the uniformly spaced sampling points, and that the error attains a local maximum in absolute sense in the interior of the elements. The local error estimates \( \varepsilon(\xi_{k,j,\text{ref}})/2^{\sqrt{p+1}/n_s} \), from (14) and (20), shown by the dots at the refinement points \( \xi_{k,j,\text{ref}} \) give a good representation of the local error, where the subscript 1 is dropped for clarity of the notation. This is an indication that the error estimate \( \tilde{\varepsilon}_j = \varepsilon(\xi_{k,j,\text{ref}}) \) is a suitable candidate for a solution–based refinement criterion. The error estimates \( \varepsilon(\xi_{k,j,\text{ref}})/2^{\sqrt{p+1}/n_s} \) have the character of a convex hull around the actual error \( \varepsilon(\xi) \), which leads to the slightly conservative error estimation. At the start of the simulation, no interpolation \( w(\xi) \) is available, such that \( \varepsilon(\xi_{k,j,\text{ref}}) = -v_{k,j,\text{ref}} \) is used for (14). This leads to the relatively high error estimates for the first data point in Figures 4a to 4c. Below the \( L_2 \) error norm is considered only, since it gives a good indication of the error in the statistical moments and it gives a smoother convergence in higher dimensions, since it is not an integral quantity as opposed to the moments.
E. Curse–of–dimensionality

The real problem in uncertainty propagation, and interpolation and integration in general, is the approximation in higher–dimensional spaces. In UQ the dimensionality of probability space is determined by the number of random parameters \( n_\xi \). Figure 5 shows the error convergence of first degree SESC with increasing \( n_\xi \in \{1, 2, 3\} \) for the \( L_2 \) error norm \( \varepsilon_{L_2} \) and its estimate \( \hat{\varepsilon}_{L_2} \) as function of \( n_\xi \). The order of convergence decreases from \( O = 2.023 \) at \( n_\xi = 1 \) to \( O = 0.7485 \) for \( n_\xi = 3 \), which is only slightly better than the theoretical \( O = \frac{1}{2} \) of MC. The error estimate accurately predicts this behavior, while it becomes slightly more conservative with increasing dimension. The order approximately follows the relation \( O = \frac{p + 1}{n_\xi} = 2/n_\xi \), which is demonstrated further in section IV. This implies that for \( n_\xi \geq 4 \), i.e. \( O \leq 0.5 \), first degree SESC essentially becomes redundant in the sense that it does no longer achieve a higher rate of convergence than MC. Therefore, in section IV, higher order polynomial interpolation is considered to remedy this effect.

![Figure 5. Error and estimate for \( n_\xi = \{1, 2, 3\} \) and \( p = 1 \) with the corresponding order of convergence \( O \).](image)

III. Extremum Diminishing and Total Variation Diminishing robustness

The reliability of first degree SESC is guaranteed by the Extremum Diminishing (ED) and Total Variation Diminishing (TVD) robustness concepts extended to probability space. These well–established robustness properties were first developed in the context of spatial finite volume discretizations of CFD problems by Jameson\(^{11} \) and Harten,\(^{9} \) respectively, to avoid unphysical values due to overshoots and undershoots at the approximation of shock discontinuities in physical space. In their extension to probability space, these concepts ensure that no non–zero probabilities are predicted for unphysical realizations due to overshoots and undershoots in the interpolation of the samples. ED and TVD can be defined in probability space as follows.

**Definition 1. (Extremum Diminishing)** A set of samples \( \mathbf{v} \) is Extremum Diminishing (ED) with respect
to response surface $u(\xi)$ if
\[
\min(v) \geq \min(u(\xi)) \land \max(v) \leq \max(u(\xi)).
\] (29)

Sampling method $g$ is ED if the resulting set of samples $v$ is ED for all $u(\xi)$. Approximation $w(\xi)$ of response surface $u(\xi)$ is ED if
\[
\min(w(\xi)) \geq \min(u(\xi)) \land \max(w(\xi)) \leq \max(u(\xi)).
\] (30)

Uncertainty quantification method $q$ is ED if the resulting approximation $w(\xi)$ is ED for all $u(\xi)$.

The TVD concept is extended to probability space by the following definition.

**Definition 2. (Total Variation Diminishing)** A set of samples $v$ is Total Variation Diminishing (TVD) with respect to response surface $u(\xi)$ if
\[
TV(v) \leq TV(u).
\] (31)

Sampling method $g$ is TVD if the resulting set of samples $v$ is TVD for all $u(\xi)$. Approximation $w(\xi)$ of response surface $u(\xi)$ is TVD if
\[
TV(w) \leq TV(u).
\] (32)

Uncertainty quantification method $q$ is TVD if the resulting approximation $w(\xi)$ is TVD for all $u(\xi)$.

Total Variation (TV) in probability space is defined below in analogy to its definition in a one–dimensional physical space.\(^9\)

**Definition 3. (Total Variation)** The total variation (TV) of response surface $u(\xi)$ in the space $\Xi$ of random parameter $\xi$ is
\[
TV(u) = \int_\Xi \left| \frac{\partial u}{\partial \xi} \right| d\xi.
\] (33)

The TV of the continuous and piecewise continuously differentiable approximation $w(\xi)$ is
\[
TV(w) = \sum_{j=1}^{n_x} TV(w_j) = \sum_{j=1}^{n_x} \int_{\Xi_j} \left| \frac{\partial w_j}{\partial \xi} \right| d\xi.
\] (34)

The TV of a discrete set of samples $v$ is
\[
TV(v) = \sum_{k=1}^{n_x-1} |v_{k+1} - v_k|,
\] (35)

for $\xi_1 < \xi_2 < \cdots < \xi_{n_x}$.

It can be shown that MC and any sampling method $g$ satisfy the ED and TVD definitions. First degree SESC is also ED and TVD for arbitrary $n_\xi$,\(^{23}\) which makes the method in principle as robust as MC. This is an advantageous property in robust design optimization, in which the reliability of the employed methods is essential. It also is an important property in higher–dimensional probability spaces, in which it is largely impossible to visually inspect the quality of response surface approximations.

**IV. Higher order polynomial interpolation**

A higher order interpolation is pursued using higher degree Newton–Cotes quadrature points to obtain a constant order of convergence with increasing dimensionality. The higher order method is applied to a non–uniform beta distribution and the normal distribution with unbounded support. Remaining issues are also addressed.
for the initial discretization of a two-dimensional parameter space using the relation
as opposed
Higher degree SESC is applied to test function \((B)\) of higher degree Newton–Cotes quadrature points. In addition, the total number of
sampling locations \(\xi_k\) in an element is increased from \(n_L + 1\) to \(N + 1\) \((7)\) using higher degree Newton–Cotes
quadrature points, which can be derived from the location of the element vertices. The higher order sampling
points for \(p = \{2, 4\}\) are shown in Figure 6 for the initial discretization of a two-dimensional parameter space
\(\Xi\) with \(n_L = 2\) and \(n_e = 4\).

A. Higher degree Newton–Cotes quadrature

Higher degree Newton–Cotes quadrature points are used in addition to the vertices of the simplex elements
to construct the higher order polynomial approximation \(w(\xi)\) of response surface \(u(\xi)\). The data structure
of the vertices and the elements remains the same as for first degree SESC. In addition, the total number of
sampling locations \(\xi_k\) in an element is increased from \(n_L + 1\) to \(N + 1\) \((7)\) using higher degree Newton–Cotes
quadrature points, which can be derived from the location of the element vertices. The higher order sampling
points for \(p = \{2, 4\}\) are shown in Figure 6 for the initial discretization of a two-dimensional parameter space
\(\Xi\) with \(n_L = 2\) and \(n_e = 4\).

Higher order interpolation does not automatically satisfy the ED requirements in Definition 1 as opposed
to first degree SESC. The ED property is no longer guaranteed, if the higher order polynomial \(w_j(\xi)\) with
\(\xi \in \Xi_j\) does not conserve the extrema of the samples \(v_j\) in element \(\Xi_j\), i.e.,
\[
\min_{\Xi_j} w_j(\xi) < \min v_j \lor \max_{\Xi_j} w_j(\xi) > \max v_j. \quad (36)
\]

In practice these inequalities are checked using the Monte Carlo sampling of the response approximation \(w_j(\xi_{MC_{j,k}})\) for computing the moments in \((11)\) with \(\xi_{MC_{j,k}} \in \Xi_j\), \(k_{j,k} \in \{1, \ldots, n_{MC_{j}}\}\), and \(k = \{1, \ldots, n_{MC_{j}}\}\)
\[
\min_{k_{j,k}} w_j(\xi_{MC_{j,k}}) < \min v_j \lor \max_{k_{j,k}} w_j(\xi_{MC_{j,k}}) > \max v_j. \quad (37)
\]

If expression \((37)\) holds, \(w_j(\xi)\) in element \(\Xi_j\) is replaced by a piecewise linear function using a sub–
triangulation of the samples \(v_j\), which is by definition ED.

B. Constant order of convergence

Higher degree SESC is applied to test function \((B)\), for which the results for the order of convergence \(O\) are
summarized in Table 1. The convergence rate decreases with increasing \(n_L\) as already observed in section II,
and increases with increasing \(p\). The convergence as function of \(n_e\) is slightly faster than as function of
\(n_e\), which closely follows the relation \(O = (p + 1)/n_L\) also given in Table 1. It is suggested to counteract
the decreasing convergence order \(O\) for increasing \(n_L\) by simultaneously increasing the polynomial degree
according to \(p = O n_L - 1\). This results, in principle, in a constant order of convergence \(O\) with dimensionality
\(n_L\), which can be chosen higher than the \(O = \frac{1}{2}\) of MC. For example, a constant \(O = 2\) follows from
\(p = \{1, 3, 5, \ldots\}\) for \(n_L = \{1, 2, 3, \ldots\}\) as highlighted in Table 1.

The \(L_2\) error norm \(\varepsilon_{L_2}\) and its estimate \(\hat{\varepsilon}_{L_2}\) are shown in Figure 7 using the relation \(p = O n_L - 1\)
with \(n_L = \{1, 2, 3\}\) and \(O = 2\). The refinement is continued until the number of vertices is \(n_e = 10^3\) for
\(n_L = \{1, 2, 3\}\). Although the absolute error increases approximately one order of magnitude per additional

\[\text{American Institute of Aeronautics and Astronautics}\]
### Table 1. Order of convergence $O$ of the $L_2$ error.

<table>
<thead>
<tr>
<th>$p$</th>
<th>$n_\xi$</th>
<th>$\text{function of } n_s$</th>
<th>$\text{function of } n_e$</th>
<th>$O = \frac{p+1}{n_\xi}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>2.023 1.081 0.748</td>
<td>2.003 1.006 0.641</td>
<td>2.000 1.000 0.667</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>3.007 1.532 1.075</td>
<td>2.992 1.478 0.991</td>
<td>3.000 1.500 1.000</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>4.049 <strong>2.019</strong> 1.369</td>
<td>3.996 1.971 1.296</td>
<td>4.000 2.000 1.333</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>5.168 2.536 1.629</td>
<td>4.971 2.491 1.565</td>
<td>5.000 2.500 1.667</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>6.333 3.074 <strong>2.011</strong></td>
<td>6.011 3.031 1.947</td>
<td>6.000 3.000 2.000</td>
</tr>
</tbody>
</table>

random parameter, the order of convergence remains with $O = 2$ significantly larger than the theoretical MC value of $O = \frac{1}{2}$ as opposed to using constant $p = 1$ as in Figure 5. This result shows that increasing the polynomial interpolation degree $p$ with increasing dimensionality $n_\xi$ is essential for efficient uncertainty quantification in case of many input uncertainties. Estimate $\hat{\varepsilon}_{L_2}$ gives a good representation of the error for all cases, with a slightly more conservative estimate for increasing $n_\xi$ and $p$.

![Figure 7](image-url)

**Figure 7.** Error and estimate for SESC with $n_\xi = \{1, 2, 3\}$, $O = 2$, and $p = O n_\xi - 1$ with the actual order of convergence.

### C. Non–uniform distribution

Uncertainty propagation concerns the approximation $w(\xi)$ of a response $u(\xi)$ in a weighted parameter space $\Xi$, weighted by the probability distribution of the random input parameters $\xi$. It is, therefore, essential to demonstrate the convergence behavior of UQ methods also for non–uniform input distributions. Here the
The unimodal symmetric beta distribution is considered with probability density
\[ f_\xi(\xi) = \frac{\xi^{\beta_1-1}(1-\xi)^{\beta_2-1}}{B(\beta_1, \beta_2)}, \]
with beta function \( B(\beta_1, \beta_2) \) and parameter values \( \beta_1 = \beta_2 = 4 \). For this non-uniform distribution, three refinement criteria are compared:

1. Probability measure \( \Omega_j \)
   \[ \Omega_j = \int_{\Xi_j} f_\xi(\xi) d\xi \approx \frac{n_{\text{MC},j}}{n_{\text{MC}}}; \]  

2. Volume measure \( \Xi_j \)
   \[ \Xi_j = \int_{\Xi_j} d\xi; \]  

3. Solution based \( L_2 \) error estimate measure from (27)
   \[ \hat{\varepsilon}_{L_2,j} = \frac{n_{\text{MC},j} \varepsilon_j^2}{n_{\text{MC},j}}. \]  

The refinement is performed for measure \( \hat{\varepsilon}_{L_2} \), until the discretization contains \( n_\varepsilon = 10^3 \) vertices for \( n_\xi = \{1, 2, 3\} \) and \( p = O(n_\xi - 1) \). The refinement based on measures \( \Omega_j \) and \( \Xi_j \) is then performed until the same number of samples \( n_s \) is reached as for \( \hat{\varepsilon}_{L_2,j} \) and \( n_\varepsilon = 10^3 \). This procedure is chosen to make a fair comparison based on equal computational costs \( n_s \), since the different refinement criteria lead to discretizations with different \( n_s \) for the same \( n_\varepsilon \) or \( n_\varepsilon \).

The results for error estimate \( \hat{\varepsilon}_{L_2,j} \) and actual error \( \varepsilon_{L_2,j} \) are given in Tables 2 and 3. Refinement criterion \( \hat{\varepsilon}_{L_2,j} \) successfully minimizes the error estimate \( \hat{\varepsilon}_{L_2} \) in Table 2 compared to the \( \Omega_j \) and \( \Xi_j \) measures for all \( n_\xi \). The actual error \( \varepsilon_{L_2} \) for measure \( \hat{\varepsilon}_{L_2} \) in Table 3 is only slightly lower than for volume refinement and marginally higher for \( n_\xi = 3 \). Surprisingly for this probability weighted approximation problem, probability refinement performs consistently worse with an order of magnitude higher errors. This is caused by the dominating interpolation errors at the boundaries of parameter space \( \Xi \). The decaying probability in the tails of the beta distribution more effectively reduces their integral effect on \( \varepsilon_{L_2} \) for uniform volume refinement than for the relatively large elements and local errors at the boundary for probability refinement. It is also remarkable that the errors for constant \( n_\varepsilon = 10^3 \) decrease for increasing dimension \( n_\varepsilon \). This can be explained by an increasing number of elements \( n_\xi \) and, consequently, an increasing \( n_s \) with increasing \( n_\varepsilon \) for constant \( n_\varepsilon \). This effect can also be seen in Figure 7. Solution based refinement measure \( \varepsilon_{L_2,j} \) (41) can be further improved by taking into account the number of samples \( n_{\text{MC},j} \) required to refine element \( \Xi_j \). This number is not constant, since the elements after refinement can share new samples with previously refined neighbors. The modified refinement measure is then
\[ \frac{\hat{\varepsilon}_{L_2,j}}{n_{\text{MC},j}}. \]

Based on the assumption that the error reduction at refinement is proportional to the error \( \varepsilon_{L_2,j} \) in element \( \Xi_j \), this formulation estimates the largest error reduction per additional sample. This modified refinement measure adds, however, significantly to the computational time.

<table>
<thead>
<tr>
<th>( n_\xi )</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Omega_j )</td>
<td>1.141 \cdot 10^{-7}</td>
<td>8.497 \cdot 10^{-8}</td>
<td>1.094 \cdot 10^{-8}</td>
</tr>
<tr>
<td>( \Xi_j )</td>
<td>9.359 \cdot 10^{-9}</td>
<td>4.728 \cdot 10^{-9}</td>
<td>1.642 \cdot 10^{-9}</td>
</tr>
<tr>
<td>( \hat{\varepsilon}_{L_2,j} )</td>
<td>8.091 \cdot 10^{-9}</td>
<td>6.572 \cdot 10^{-10}</td>
<td>6.113 \cdot 10^{-11}</td>
</tr>
</tbody>
</table>

It is not to be expected that unweighted refinement based on \( \Xi_j \) results in the minimum error in this weighted integration problem. Therefore, also a mixed refinement measure \( M_j \) is considered which blends volume and probability refinement as follows
\[ M_j = (1 - \alpha) \Xi_j + \alpha \Omega_j, \]
the resulting SESC triangulation of a two–dimensional parameter space Ξ with $\xi = 0$. For the limit case $\alpha = 0$ holds $M_j = \bar{\xi}_j$ and $\alpha = 1$ gives $M_j = \bar{\Omega}_j$. The results in Table 4 show that a slight contribution of probability refinement $\bar{\Omega}_j$ at $\alpha = 20\%$ reduces the error with respect to pure volume refinement, $\alpha = 0$. Both the optimal value of $\alpha$ and the relative improvement increase moderately with dimensionality $n_\xi$. Mixed refinement measure $M_j$ has also the advantage that it can easily be applied to problem with a large number of output quantities of interest as opposed to solution based measure $\tilde{\epsilon}_{L_2}$.

Table 3. Error $\epsilon_{L_2}$ for $n_v = 1000$ and the beta distribution.

<table>
<thead>
<tr>
<th>$n_\xi$</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{\Omega}_j$</td>
<td>$9.481 \cdot 10^{-8}$</td>
<td>$3.541 \cdot 10^{-8}$</td>
<td>$2.137 \cdot 10^{-9}$</td>
</tr>
<tr>
<td>$\bar{\xi}_j$</td>
<td>$6.959 \cdot 10^{-9}$</td>
<td>$2.088 \cdot 10^{-9}$</td>
<td>$2.550 \cdot 10^{-10}$</td>
</tr>
<tr>
<td>$\tilde{\epsilon}_{L_2}$</td>
<td>$5.914 \cdot 10^{-9}$</td>
<td>$1.103 \cdot 10^{-9}$</td>
<td>$4.818 \cdot 10^{-10}$</td>
</tr>
</tbody>
</table>

with $\bar{\xi} = \sum_{j=1}^{n_\xi} \bar{\xi}_j$ and parameter $\alpha \in [0, 1]$. For the limit case $\alpha = 0$ holds $M_j = \bar{\xi}_j$ and $\alpha = 1$ gives $M_j = \bar{\Omega}_j$. The results in Table 4 show that a slight contribution of probability refinement $\bar{\Omega}_j$ at $\alpha = 20\%$ reduces the error with respect to pure volume refinement, $\alpha = 0$. Both the optimal value of $\alpha$ and the relative improvement increase moderately with dimensionality $n_\xi$. Mixed refinement measure $M_j$ has also the advantage that it can easily be applied to problem with a large number of output quantities of interest as opposed to solution based measure $\tilde{\epsilon}_{L_2}$.

Table 4. Error $\epsilon_{L_2}$ for mixed refinement criterion $M_j$, $n_v = 1000$, and the beta distribution.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$n_\xi$</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>6.959 $\cdot 10^{-9}$</td>
<td>2.088 $\cdot 10^{-9}$</td>
<td>2.550 $\cdot 10^{-10}$</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>6.560 $\cdot 10^{-9}$</td>
<td>9.282 $\cdot 10^{-10}$</td>
<td>1.256 $\cdot 10^{-10}$</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>6.560 $\cdot 10^{-9}$</td>
<td>1.038 $\cdot 10^{-9}$</td>
<td>1.023 $\cdot 10^{-10}$</td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>6.560 $\cdot 10^{-9}$</td>
<td>1.442 $\cdot 10^{-9}$</td>
<td>7.968 $\cdot 10^{-11}$</td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>6.780 $\cdot 10^{-9}$</td>
<td>1.614 $\cdot 10^{-9}$</td>
<td>8.131 $\cdot 10^{-11}$</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>7.684 $\cdot 10^{-9}$</td>
<td>1.629 $\cdot 10^{-9}$</td>
<td>1.303 $\cdot 10^{-10}$</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>9.481 $\cdot 10^{-8}$</td>
<td>3.541 $\cdot 10^{-8}$</td>
<td>2.137 $\cdot 10^{-9}$</td>
<td></td>
</tr>
</tbody>
</table>

In Figure 8 the resulting SESC triangulation of a two–dimensional parameter space $\Xi$ with $n_\xi = 2$ is shown for measure $M_j$ with $\alpha = 0.2$ and $n_v = 250$. As a reference also a Monte Carlo sample with $n_{\text{MC}} = 10^4$ is shown. For completeness also the results of the refinement measures for the uniform distribution of Section B are given in Table 5. For the uniform distribution, the results of refinement measures $\bar{\Omega}_j$, $\bar{\xi}_j$, and $M_j$ coincide. The errors $\epsilon_{L_2}$ are slightly higher for the uniform distribution than for the beta distribution with $M_j$ and $\alpha = 0.2$. This demonstrates that SESC is equally or more suitable for probabilistically weighted interpolation problems than for unweighted parameter studies, which is an essential property for a UQ method.

![Figure 8](image_url)

(a) SESC with $M_j$, $\alpha = 0.2$, and $n_v = 250$

(b) MC with $n_{\text{MC}} = 10^4$

Figure 8. MC and SESC discretizations of parameter space $\Xi$ with $n_\xi = 2$ for the beta distribution.
Table 5. Error $\varepsilon_{L_2}$ for $n_v = 1000$ and the uniform distribution.

<table>
<thead>
<tr>
<th>$n_\xi$</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Omega_j$</td>
<td>$8.536 \cdot 10^{-9}$</td>
<td>$1.983 \cdot 10^{-9}$</td>
<td>$3.024 \cdot 10^{-10}$</td>
</tr>
<tr>
<td>$\Xi_j$</td>
<td>$8.536 \cdot 10^{-9}$</td>
<td>$1.983 \cdot 10^{-9}$</td>
<td>$3.024 \cdot 10^{-10}$</td>
</tr>
<tr>
<td>$\hat{\varepsilon}_{L_2}$</td>
<td>$7.042 \cdot 10^{-9}$</td>
<td>$1.860 \cdot 10^{-9}$</td>
<td>$1.876 \cdot 10^{-9}$</td>
</tr>
</tbody>
</table>

D. Unbounded parameter domain

Another issue for multi-element UQ methods are probability distributions with unbounded support. Unbounded random parameter domains are frequently encountered in terms of, for example, the normal distribution, which naturally emerges from the central limit theorem and the Karhunen–Loève expansion. Multi-element discretizations are usually based on elements $\Xi_j$ with finite volume $0 < \Xi_j < \infty$, such that infinite parameter spaces $\Xi$ need to be truncated to a bounded support in order to discretize them using a finite number of elements. This truncation introduces an undesirable stochastic bias error into the UQ results.

SESC is based on Monte Carlo integration of the piecewise polynomial response surface approximation $w(\xi)$ for computing the statistical moments. The Monte Carlo sampling points $\xi_{MC_k}$ always have bounded abscissas $\xi_{MC_k} \in (-\infty, \infty)^{n_\xi}$ for all $k = 1, \ldots, n_{MC}$, even for unbounded distributions such as the normal distribution. This property is used in SESC to truncate the infinite parameter space $\Xi$ to a bounded space $\Xi'$, without affecting the approximation of the statistical moments and the probability distribution. In Figure 9 the SESC discretization of two-dimensional parameter space $\Xi$ with $n_\xi = 2$ weighted by a normal input distribution is shown as an example for measure $\bar{M}_j$ with $\alpha = 0.2$, $n_v = 250$ and $n_{sMC} = 10^4$. The normal distribution used, $\mathcal{N}(\mu, \sigma)$, has the same mean and standard deviation as the uniform distribution $\mathcal{U}(0, 1)$ of section B. The maximum distance between $\xi_{MC_k}$ and $\mu_{\xi_i}$ with $i = 1, \ldots, n_\xi$ and $k = 1, \ldots, n_{sMC}$ for the random Monte Carlo sample of Figure 9b is 1.299. This leads for symmetry to the truncated parameter space $\Xi' = [-0.799, 1.799]^{n_\xi}$ of Figure 9a.

![SESC with \(\bar{M}_j\), \(\alpha = 0.2\), and \(n_v = 250\)](image)

![MC with \(n_{sMC} = 10^4\)](image)

Figure 9. MC and SESC discretizations of parameter space $\Xi$ with $n_\xi = 2$ for the normal distribution.

Results for the refinement criteria $\Omega_j$, $\Xi_j$, $\hat{\varepsilon}_{L_2}$, and $\bar{M}_j$ with $\alpha = 0.2$ are given in Table 6 for the normal distribution. The errors $\varepsilon_{L_2}$ for $\hat{\varepsilon}_{L_2}$ and $\bar{M}_j$ are comparable in magnitude and smaller than those for $\Omega_j$ and $\Xi_j$. All errors are slightly larger than for the uniform and beta distribution due to the larger parameter space $\Xi'$ and the larger input standard deviation $\sigma_{\xi}$ compared to the beta distribution.

E. Remaining issues

We have demonstrated that higher degree SESC maintains a constant order of convergence $O$ with increasing dimensionality $n_\xi$, which is higher than the $O = \frac{1}{2}$ of MC. However, MC has another special property that
the number of samples $n_{\text{Minit}}$ in its initial approximation is with $n_{\text{Minit}} = 1$ independent of $n_L$. The number of samples $n_{\text{init}}$ in the initial SESC discretization increases, however, rapidly with increasing $n_L$, especially when the relation $p = \mathcal{O}n_L^{-1}$ is used as indicated in Table 7. This second aspect of the curse–of–dimensionality can also be identified in Figure 7. This effect is partly caused by the increasing asymptotic average number of Newton–Cotes quadrature samples per element $n_s/n_e$ with increasing $p$ also given in Table 7 and the tensorial grid of samples in higher dimensions visible in, for example, Figure 1b.

<table>
<thead>
<tr>
<th>$p$</th>
<th>$n_{\text{init}}$</th>
<th>$n_{\text{n}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3 5 9</td>
<td>1.00 0.53 0.24</td>
</tr>
<tr>
<td>2</td>
<td>5 13 35</td>
<td>2.00 2.06 1.60</td>
</tr>
<tr>
<td>3</td>
<td>7 25 91</td>
<td>3.00 4.59 5.09</td>
</tr>
<tr>
<td>4</td>
<td>9 41 189</td>
<td>4.00 8.13 11.70</td>
</tr>
<tr>
<td>5</td>
<td>11 61 341</td>
<td>5.00 12.66 22.64</td>
</tr>
</tbody>
</table>

V. Conclusions

A higher degree Simplex Elements Stochastic Collocation (SESC) method is introduced for efficient uncertainty propagation of multiple uncertain parameters. It approximates the response surface $u(\xi)$ using a piecewise polynomial interpolation of higher degree Newton–Cotes quadrature points in a simplex elements discretization of parameter space $\Xi$. The method converges by adaptively refining one element at a time, where all samples are reused after the hierarchical refinements, until an error estimate reaches a stopping criterion. SESC is proven to achieve the same robustness as Monte Carlo (MC) simulation in terms of the Extremum Diminishing (ED) and Total Variation Diminishing (TVD) concepts in probability space.

In order to counteract the curse–of–dimensionality in terms of a decreasing order of convergence $\mathcal{O}$ for an increasing number of random input parameters $n_\xi$, it is suggested to simultaneously increase the polynomial degree according to $p = \mathcal{O}n_L^{-1} - 1$. This is shown to result in a constant order of convergence $\mathcal{O}$ with increasing dimensionality $n_L$ higher than the $\mathcal{O} = \frac{1}{2}$ of MC. This result shows that higher order polynomial interpolation is essential for efficient uncertainty quantification in case of many input uncertainties. Also four refinement criteria are considered for non–uniform probability distributions and unbounded parameter domains in terms of the beta and normal distribution, respectively. A mix of a probability and volume refinement criterion with blending parameter value $\alpha = 0.2$ performs slightly better than pure volume refinement and an $L_2$ error estimate measure. Pure probability refinement gives the largest error of these four measures. The derived error estimates give an excellent, slightly conservative indication of the actual error over the whole range of $n_\xi$ and $p$ considered.

The increase of the polynomial order $p$ with increasing dimensionality $n_L$ results, however, in a fast increase of the number of samples in the initial SESC discretization. Randomized refinement sampling will be considered in future work to reduce this aspect of the curse–of–dimensionality, since random sampling is known to be a very effective sampling strategy in higher dimensions.
References


