

Lectures

Advanced Methods of Stochastic and Optimization in Industrial Applications

Dirk Roos

ADVANCED METHODS OF STOCHASTIC AND OPTIMIZATION IN INDUSTRIAL APPLICATIONS

Dirk Roos*

DYNARDO – Dynamic Software and Engineering GmbH, Weimar, Germany

ABSTRACT: In real case applications within the virtual prototyping process, it is not always possible to reduce the complexity of the physical models and to obtain numerical models which can be solved quickly. Usually, every single numerical simulation takes hours or even days. Although the progresses in numerical methods and high performance computing, in such cases, it is not possible to explore various model configurations, hence efficient surrogate models are required.

The paper gives an overview about advanced methods of meta-modeling. In addition, some new aspects are introduced to improve the accuracy and predictability of surrogate models, commonly used in numerical models for automotive applications. Whereby, the main topic is reducing the necessary number of design evaluations, e.g. finite element analysis within global variance-based sensitivity and robustness studies. In addition, the similar approach can be used to perform optimization and stochastic analysis and to create synthetic meta-models for experimental data.

Stochastic analysis and optimization of technical systems have gained increasing attention in engineering practice in recent years. While in reliability analysis, occurrence of rare events, that violate the safety or performance criteria of a system are computed, robustness assessment deals with the sensitivity towards natural scatter of the system parameters in a high probability range. In order to obtain meaningful results from any of these assessments, an accurate model of the underlying random parameters has to be used. There exist physical or geometrical quantities which vary randomly along the geometry of a structure, such as distributed loads, Young's modulus, material thickness etc. The spatial correlation of such parameters can be taken into account by modelling them as random field.

Design for Six Sigma is a concept to optimize the design such that the parts conform with Six Sigma quality, i.e. quality and reliability are explicit optimization goals even with variations e.g. in manufacturing, design configurations and environment. This paper presents an adaptive response surface method in the design and random space to solve reliability-based optimization problems under uncertainties. The probabilistic and optimization tasks are performed with the **optiSLang**, **SoS** and **SLang** software packages.

KEYWORDS: surrogate models, robustness evaluation, random fields, stochastic finite elements, reliability analysis, robust design optimization

1 INTRODUCTION

Within many engineering fields, structural design must meet challenging demands of cost-effectiveness. This leads to light-weight, highly flexible, and consequently vulnerable structures. In order to assess potential risks associated with such a design, the structural analysis must take into account available information on the randomness of loads and structural properties. It means that the analysis should include reliability estimates in an appropriate manner. Methods of multidisciplinary optimization have obtained an increasing importance in the design of engineering problems for improving the design performance and reducing costs. The virtual

prototyping is an interdisciplinary process. Such a multidisciplinary approach requires to run different solvers in parallel and to handle different types of constraints and objectives. Arbitrary engineering software and complex non-linear analyses have to be connected. Resulting optimization problems may become very noisy, very sensitive to design changes or ill-conditioned for mathematical function analysis (e.g. non-differentiable, non-convex, non-smooth).

Considering the properties of the computational analysis realistically it is necessary to take into account some uncertainty. This uncertainty can be conveniently described in terms of probability measures, such as distribution functions. It is a major goal of stochastic computational analysis to relate the uncertainties of the input variables to the un-

*Corresponding author: Dr.-Ing. Dirk Roos, DYNARDO – Dynamic Software and Engineering GmbH, Luthergasse 1d, D-99423 Weimar, Germany, E-Mail: dirk.roos@dynardo.de

certainty of the model response performance. The stochastic models for these uncertain parameters can be one of the following:

- Time-independent random problems (Reliability analysis)
 - Random variables (constant in time and space)
 - Random fields (correlated fluctuations in space)
- Time-dependent random problems (First passage reliability analysis)
 - Random processes (correlated fluctuations in time)

Within a computer implementation it becomes necessary to discretize random fields/processes in time and space. This immediately leads to a considerable number of random variables.

The aims of stochastic analysis include safety measures such as time-invariant failure probabilities or safety indices, response variability in terms of standard deviations or time-variant first passage probabilities. The available solution methods for these tasks are exact methods (numerical integration, Monte Carlo simulation – MCS), approximate (first and second order reliability method – FORM/SORM, response surface method - RSM) or other techniques (e.g. based on Markov vector theory in stochastic dynamics).

The stochastic analysis software **SLang** (– the Structural Language) includes several methods solving all off these stochastic models. Currently, **optiSLang** (– the optimizing Structural Language) supports methods to solve nonlinear optimization and to analyse random variables only. In addition, the **SoS** (– Statistics on Structures) add-on tool to **optiSLang** provides methods to solve random fields.

2 META-MODELING

2.1 INTRODUCTION

Meta-modeling is one of the most popular strategy for design exploration within nonlinear optimization [1–3] and stochastic analysis [4–7]. Moreover, the engineer has to calculate the general trend of physical phenomena or would like to re-use design experience on different projects. Due to the inherent complexity of many engineering problems it is quite alluring to approximate the problem and to solve other design configurations in a smooth sub-domain by applying a surrogate model ([8, 9]). Starting from a reduced number of simulations, a surrogate model of the original physical problem can be used to perform various possible design configurations without computing any further analyses. So, the engineer may apply a classical design of experiment or stochastic

sampling methods for more than $n = 10 \dots 15$ input parameters to calculate the simulations.

2.2 POLYNOMIAL LEAST SQUARE APPROXIMATION

These well-distributed results can be used to create the meta-models. To simulate complex real world systems the response surface methodology is becoming very popular and is widely applied in many scientific areas. In general, response surface methodology is a statistical method for constructing smooth approximations to functions in a multi-dimensional space. In design studies, e.g. design optimization or reliability analysis, a response surface is generated with appropriate approximation functions on a suitable set of discrete support points distributed throughout the design space of interest.

A commonly used approximation method of model responses, objectives, constraints and state functions

$$y(\mathbf{x}) \mapsto \hat{y}(\mathbf{x})$$

is the regression analysis. Usually, the approximation function is a first or second order polynomial ([10–12]) as shown in Figure 2. As an example in the ($n = 2$)-dimensional case, k -responses ($k = 0, \dots, N$) will be approximated using a least square quadratic polynomial in the following form:

$$y_k = \beta_0 + \beta_1 x_{1k} + \beta_2 x_{2k} + \beta_{11} x_{1k}^2 + \beta_{22} x_{2k}^2 + 2\beta_{12} x_{1k} x_{2k} + \varepsilon_k \quad (1)$$

Herein the term ε_k represents the approximation errors. The approximate coefficients β can be calculated using the least square postulate

$$S = \sum_{k=1}^m \varepsilon_k^2 = \varepsilon^T \varepsilon \rightarrow \min$$

Of course the accuracy of the approximation compared to the real problem has to be checked and verified. For reasonably smooth problems, the accuracy of response surface approximations improves as the number of points increases. However, this effect decreases with the degree of oversampling. An attractive advantage of the response surface methodology is the smoothing by approximating the subproblem. Especially for noisy problems like crash analysis, for which the catch of global trends is more important and the local noise may not be meaningful, a smoothing of the problem may be advantageous. However, the recommended area of application is restricted to reasonably smooth problems with a small number of input variables, because linear and quadratic functions are possibly weak approximations near and far from certain support points. And using polynomials higher than second order may only result in higher local accuracy with many suboptima. Because of that in the last years, different

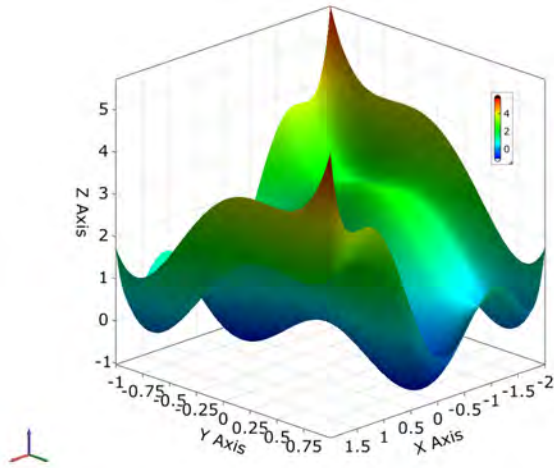


Figure 1: Original model response function $z(x, y)$.

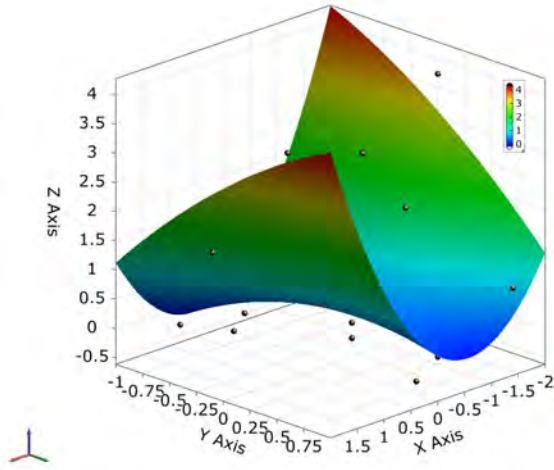


Figure 2: Polynomial least square approximation of a given set of support points. Quadratic approximation function $\hat{z}(x, y)$ using quadratic regression.

advanced surrogate models have been developed to improve the accuracy and predictability of surrogate models.

2.3 STEPWISE RESPONSE SURFACES

Within the stepwise response surface approach the response surface is determined by backward stepwise regression ([13]), so that the square and cross terms can be absorbed into the model automatically according to their actual contribution, which is calculated by repeated variance analysis. The quadratic backward stepwise regression begins with a model that includes all constant, linear and quadratic terms of the least square polynomial approximation (1). By deleting trivial regressors one at a time this approach develops a stepwise final regression model which only contains regression coefficients which have large effects on the responses. This method is

often applied for multiobjective optimization of vehicles. e.g. in [14–16]. Therewith the number n of linear and quadratic terms of the regression model can be dramatically reduced.

2.4 SHEPARD INTERPOLATION

A well known method that among all scattered data for an arbitrary number of variables is the [17] method. Shepard's method and its generalization (e.g. [18]) is a statistical interpolation averaging the known values of the original function which exactly interpolates the values of the data. The most relevant drawback of this method is that the interpolated values are always constrained between the maximum and minimum values of the data set.

2.5 KRIGING MODELS

Kriging was originally developed to model spatial variations in measured geological models ([19]). These models are inherently random, and in most cases the variation about a constant mean value is gaussian. Furthermore, the use of Kriging models is also becoming popularity for approximating optimization and stochastic problems ([20]). Kriging is an accurate surrogate model for this type of application due to its flexibility to approximate many different and complex response functions. In addition, it is also suitable for deterministic models since it interpolates the support data points and provide a confidence interval about a prediction result of the approximation model. The flexibility of this method is a result of using a set of parameters to define the model but the process of selecting the best set of parameters for a Kriging model has a few drawbacks. These parameters must be found via a constrained iterative search, a computationally expensive process with respect of the assumption gaussian random process ([21]).

2.6 ADVANCED MOVING LEAST SQUARE APPROXIMATION

Moving least square (MLS) functions can approximate locally clustered support point samples with higher local approximation quality. In addition MLS improve the response surface model using additional support points. MLS is formulated as

$$\hat{y}(\mathbf{x}) = \sum_{i=1}^{n_b} h_i(\mathbf{x}) a_i(\mathbf{x}) = \mathbf{h}^T(\mathbf{x}) \mathbf{a}(\mathbf{x}) \quad (2)$$

with a predefined number of basis terms n_b , a vector of basis functions $\mathbf{h}$ and the associated vector of the coefficients $\mathbf{a}$. [22] formulates a local MLS approximation as

$$\hat{y}(\mathbf{x}, \mathbf{x}_j) = \sum_{i=1}^{n_b} h_i(\mathbf{x}_j) a_i(\mathbf{x}) = \mathbf{h}^T(\mathbf{x}_j) \mathbf{a}(\mathbf{x})$$

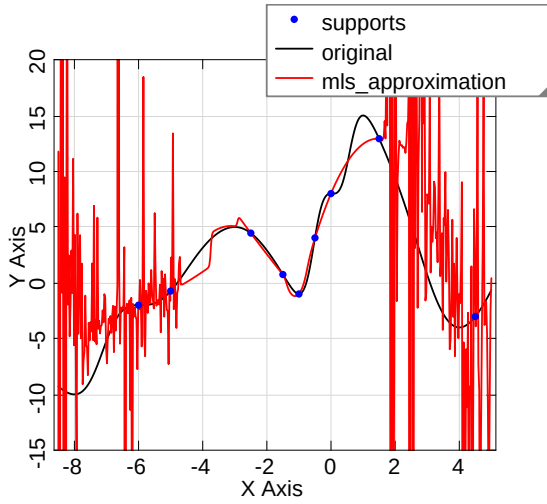


Figure 3: MLS approximation with weighting function (5) and $D = 1$. For some support points ($-2.5 \leq x \leq 0$) the approximation error is very small but for coordinates where the support points have larger distances the shape function is not continuous.

with $j = 1, \dots, n_s$ support points. The approximate coefficient vector $\mathbf{a}$ can be calculated using the weighted least square postulate

$$\begin{aligned} S(\mathbf{x}) &= \sum_{j=1}^{n_s} w(\mathbf{x} - \mathbf{x}_j) (\hat{y}(\mathbf{x}, \mathbf{x}_j) - y(\mathbf{x}_j))^2 \\ &= \sum_{j=1}^{n_s} w(\mathbf{x} - \mathbf{x}_j) \left(\sum_{i=1}^{n_b} h_i(\mathbf{x}_j) a_i(\mathbf{x}) - y(\mathbf{x}_j) \right)^2 \\ &= (\mathbf{H}\mathbf{a} - \mathbf{g})^T \mathbf{W}(\mathbf{x}) (\mathbf{H}\mathbf{a} - \mathbf{g}) \rightarrow \min \end{aligned} \quad (3)$$

with the weighting function $w(\mathbf{x} - \mathbf{x}_j)$ and

$$\begin{aligned} \mathbf{g} &= [y(\mathbf{x}_1) \ y(\mathbf{x}_2) \ \dots \ y(\mathbf{x}_{n_s})]^T \\ \mathbf{H} &= [\mathbf{h}(\mathbf{x}_1) \ \mathbf{h}(\mathbf{x}_2) \ \dots \ \mathbf{h}(\mathbf{x}_{n_s})]^T \\ \mathbf{h}(\mathbf{x}_j) &= [h_1(\mathbf{x}_j) \ h_2(\mathbf{x}_j) \ \dots \ h_{n_b}(\mathbf{x}_j)]^T \\ \mathbf{W}(\mathbf{x}) &= \text{diag}[w(\mathbf{x} - \mathbf{x}_1) \ w(\mathbf{x} - \mathbf{x}_2) \ \dots \ w(\mathbf{x} - \mathbf{x}_{n_s})] \end{aligned}$$

The least square error $S(\mathbf{x})$ may be a minimum in case that the partial gradients are zero.

$$\frac{\partial S(\mathbf{x})}{\partial \mathbf{a}} = 0$$

So using the Equation (3) a linear equation system gives an estimation of the coefficient vector $\mathbf{a}$

$$\mathbf{a}(\mathbf{x}) = \mathbf{M}^{-1}(\mathbf{x}) \mathbf{B}(\mathbf{x}) \mathbf{g} \quad (4)$$

with

$$\begin{aligned} \mathbf{M}(\mathbf{x}) &= \mathbf{H}^T \mathbf{W}(\mathbf{x}) \mathbf{H} \\ \mathbf{B}(\mathbf{x}) &= \mathbf{H}^T \mathbf{W}(\mathbf{x}) \end{aligned}$$

Cause the matrix of the basis function $\mathbf{M}(\mathbf{x})$ should be non-singular always a sufficient number of n_s

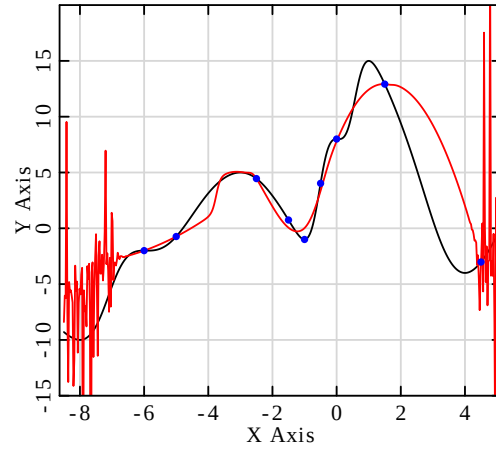


Figure 4: MLS approximation with weighting function (5) and $D = 2$. Now the interval where the shape function is continuous is larger but the error for points with $-2.5 \leq x \leq 0$ increases and for marginally coordinates where there are no support points the shape function is still not continuous.

immediate neighbor support points have to be available. The number must be at least as large as number of the basis terms. The Equation (4) inserted in (2) gives the approximation function

$$\hat{y}(\mathbf{x}) = \mathbf{h}^T(\mathbf{x}) \mathbf{M}^{-1}(\mathbf{x}) \mathbf{B}(\mathbf{x}) \mathbf{g}$$

An accurate as possible approximation quality requires a weighting function which is larger than zero $w(\mathbf{x} - \mathbf{x}_j) > 0$ and monotonically decreasing $w(\|\mathbf{x} - \mathbf{x}_j\|)$ inside of a small sub space $\Omega_s \subset \Omega$. So the influence of supports far from the actual coordinates is unimportant. An uniform weighting is given by a symmetry condition $w(\mathbf{x} - \mathbf{x}_j) = w(\mathbf{x}_j - \mathbf{x}) = w(\|\mathbf{x} - \mathbf{x}_j\|)$. Usually, an exponential function is used in this way:

$$w(\|\mathbf{x} - \mathbf{x}_j\|) = \begin{cases} e^{-\left(\frac{\|\mathbf{x} - \mathbf{x}_j\|}{D\alpha}\right)^2} & \|\mathbf{x} - \mathbf{x}_j\| \leq D \\ 0 & \|\mathbf{x} - \mathbf{x}_j\| > D \end{cases} \quad (5)$$

with a constant

$$\alpha = \frac{1}{\sqrt{-\log 0.001}}$$

and a influence radius D to choose. It is obvious that the smaller D the better the response values of the support points fit the given values. But as mentioned above at least n_b support points have to be available in every point to be approximated. Therefore it is possible that a D has to be chosen which leads to a large shape function error at the support points - see Figures 3, 4 and 5. To avoid these problems a new regularized weighting function was introduced

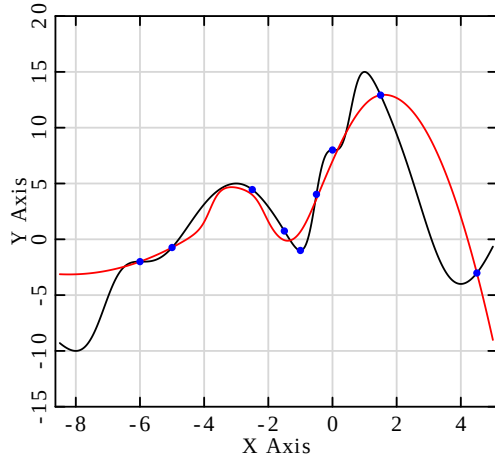


Figure 5: MLS approximation with weighting function (5) and $D = 3$. The shape function is completely continuous in the contemplated interval but the shape function error at the support points is very large.

by [23]:

$$w_R(\|\mathbf{x} - \mathbf{x}_j\|) = \begin{cases} \frac{\hat{w}_R(\|\mathbf{x} - \mathbf{x}_j\|)}{\sum_{i=1}^{n_s} \hat{w}_R(\|\mathbf{x} - \mathbf{x}_i\|)} & \|\mathbf{x} - \mathbf{x}_j\| \leq D \\ 0 & \|\mathbf{x} - \mathbf{x}_j\| > D \end{cases} \quad (6)$$

with $0 < j \leq n_s$ and

$$\hat{w}_R(d) = \frac{\left(\left(\frac{d}{D}\right)^2 + \varepsilon\right)^{-2} - (1 + \varepsilon)^{-2}}{(\varepsilon)^{-2} - (1 + \varepsilon)^{-2}}; \varepsilon \ll 1 \quad (7)$$

It is recommended by the authors to use the value

$$\varepsilon = 10^{-5}$$

As a matter of fact a large D is needed to approximate for coordinates where no support points are around and a small D is needed for coordinates where are a lot of support points in order to reach a minimal approximation error. To comply with this conditions it is necessary to use a function $d(\mathbf{x})$ for the influence radius instead of a constant D . One possibility to get such a function is to properly scale and flip horizontal a function f_{osp} , which represents the occurrence of the support points

Another problem within optimization and stochastic analysis could be the fact that the approximation for marginal coordinates follows the global trend of the given support points. This may lead to marginal approximation values which differ from the approximation values of the support points for orders of magnitudes. Then it could be useful to add some support points (border points) to force the approximation for marginal coordinates to averaged values.

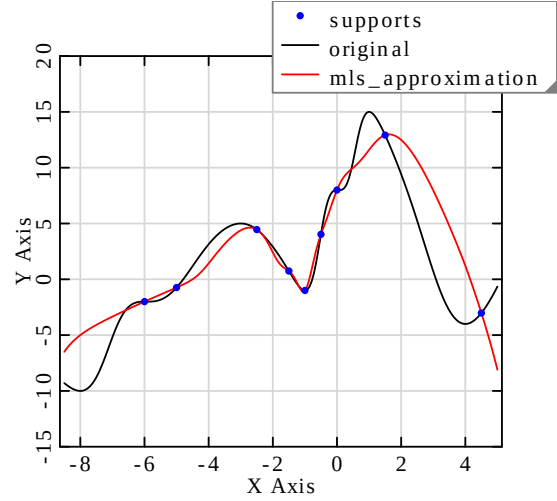


Figure 6: MLS approximation with regularized weighting function (7) and $D = 10$.

The values for the additional support points are calculated with the Shepard interpolation. Whereby the weighted interpolation function of the response surface is

The smoothing of the interpolation is numerically controlled by the smoothing exponent m .

2.7 SUPPORT VECTOR MACHINES

2.7.1 Classification

A very efficient tool for classification purposes are Support Vector Machines (SVM), which is a method from the statistical learning theory. This method was firstly proposed by [24] and became popular in the last decade. Fundamental publications from this period are [25], [26] and [27]. The algorithmic principle is to create a hyperplane, which separates the data into two classes by using the maximum margin principle.

The linear separator is a hyperplane which can be written as

$$f(\mathbf{x}) = \langle \mathbf{w}, \mathbf{x} \rangle + \alpha_0 \quad (8)$$

where $\mathbf{w}$ is the parameter vector that defines the normal to the hyperplane and α_0 is the threshold. In Figure 7 a linear separation is shown for a set of points. The two classes are associated with -1 and $+1$. The SVM principle is to maximize the distance between the hyperplane and the two classes. This can be seen in Figure 7. This principle can be written as an optimization problem

$$\max_{w,b} \min_i \{ \|\mathbf{x} - \mathbf{x}_i\| : \langle \mathbf{w}, \mathbf{x} \rangle + \alpha_0 = 0 \} \quad (9)$$

where

$$\min_i \{ \|\mathbf{x} - \mathbf{x}_i\| : \langle \mathbf{w}, \mathbf{x} \rangle + \alpha_0 = 0 \} \quad (10)$$

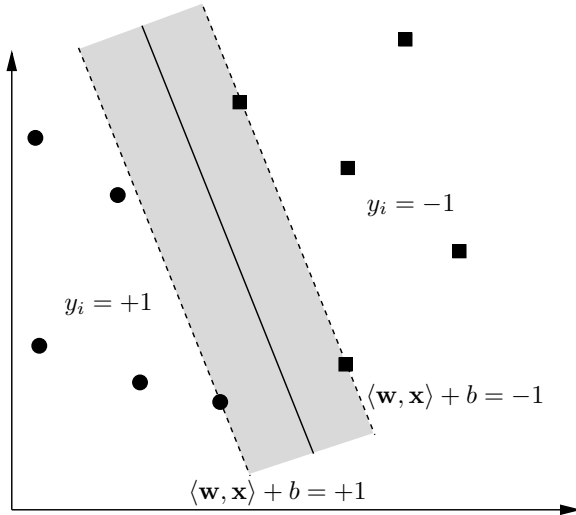


Figure 7: SVM: linear separation by a hyperplane

is the minimum distance from the training points to the hyperplane. By assuming that the minimal distance is equal to one

$$\min_{i=1..n} |\langle \mathbf{w}, \mathbf{x}_i \rangle + \alpha_0| = 1 \quad (11)$$

we obtain for the margin width

$$\Delta = \frac{2|\langle \mathbf{w}, \mathbf{x}_i \rangle + \alpha_0|}{\|\mathbf{w}\|} = \frac{2}{\|\mathbf{w}\|} \quad (12)$$

By introducing Lagrange multipliers $\alpha_i \geq 0$ we obtain the following unconstrained form of the problem

$$\mathbf{w} = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i \quad \text{with} \quad \sum_{i=1}^n \alpha_i y_i = 0 \quad (13)$$

whereby only the training points for which the Lagrange multipliers are strictly positive $\alpha_i > 0$, the so-called support vectors, are needed for the function evaluation

$$\mathbf{w} = \sum_{j=1}^s \alpha_j y_j \mathbf{x}_j \quad (14)$$

with $s < n$.

For nonlinear separable classes the training data are mapped nonlinearly into a higher-dimensional feature space and a linear separation is constructed there. This is illustrated in Figure 8. The transformation $\psi(\mathbf{x})$ which is realized as an inner product

$$f(\mathbf{x}) = \sum_{i=1}^s \alpha_i y_i \langle \psi(\mathbf{x}_i), \psi(\mathbf{x}) \rangle + \alpha_0 \quad (15)$$

can be substituted by a kernel function

$$K(\mathbf{x}, \mathbf{y}) = \langle \psi(\mathbf{x}), \psi(\mathbf{y}) \rangle \quad (16)$$

which leads finally to the expression

$$f(\mathbf{x}) = \sum_{i=1}^s \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + \alpha_0 \quad (17)$$

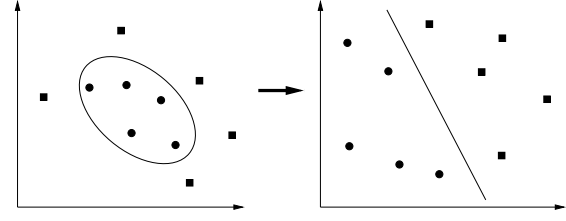


Figure 8: SVM: nonlinear projection into feature space

where explicit knowledge of the nonlinear mapping is not needed. Often used kernel types are the Gaussian kernel

$$K(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{2D^2}\right) \quad (18)$$

and the polynomial kernel

$$K(\mathbf{x}, \mathbf{y}) = (\langle \mathbf{x}, \mathbf{y} \rangle + \theta)^p \quad (19)$$

During the training of the support vector machines the Lagrange multiplier of the training points have to be determined by minimizing the primal objective function obtained from Eq.(9) which reads

$$L_p(\boldsymbol{\alpha}) = \frac{1}{2} \sum_{i=1}^s \sum_{j=1}^s y_i y_j \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j) - \sum_{i=1}^s \alpha_i \quad (20)$$

Many algorithms can be found in literature, see [27]. We use one of the fastest methods, the sequential minimal optimization algorithm proposed by [28] for this training. In this algorithm the Lagrange multipliers will be updated pair-wisely by solving the linear constraint conditions.

2.7.2 Regression

Regression based on Support Vector Machines was introduced by [29]. In [30] and [31] a detailed introduction is published. In the SVR approach an error tolerance function

$$L_\epsilon(y) = \begin{cases} 0 & |f(\mathbf{x}) - y| < \epsilon \\ |f(\mathbf{x}) - y| - \epsilon & |f(\mathbf{x}) - y| \geq \epsilon \end{cases} \quad (21)$$

which is called ϵ -insensitive loss function is defined. The optimization task is defined as

$$\text{minimize} \frac{1}{n} \sum_{i=1}^n |f(\mathbf{x}_i, \mathbf{w}) - y_i|_\epsilon + \|\mathbf{w}\|^2. \quad (22)$$

The output of the Support Vector Regression reads

$$f(\mathbf{x}) = \sum_{i=1}^n (\alpha_i^* - \alpha_i) K(\mathbf{x}_i, \mathbf{x}) + \alpha_0 \quad (23)$$

where α_i^* and α_i are the Lagrange multipliers. The primal form of the objective function reads

$$L_p(\alpha^*, \alpha) = \epsilon \sum_{i=1}^n (\alpha_i^* + \alpha_i) - \sum_{i=1}^n y_i (\alpha_i^* - \alpha_i) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\alpha_i^* - \alpha_i) (\alpha_j^* - \alpha_j) K(\mathbf{x}_i, \mathbf{x}_j) \quad (24)$$

subjected to the constraints

$$\sum_{i=1}^n (\alpha_i^* - \alpha_i) = 0, \quad 0 \leq \alpha_i^*, \alpha_i. \quad (25)$$

Again it is recommended to use the sequential minimal optimization algorithm for the determination of the Lagrange multipliers.

2.8 ARTIFICIAL NEURAL NETWORKS

2.8.1 Introduction

Artificial neural networks are inspired by biological counterparts. The brain consist of a large number (10^{11}) of highly connected elements (approximately 10^4 connections per element) - neurons. In our simplified approach, a neuron can be considered as consisting of three parts - the dendrites which serve as sensors for electrical signals and carry it into the cell body, the cell body sums up all the inputs and processes the information and finally the axon which transmits the output via the synapsis to other neurons. Within this complex system, a large number of complex information can be stored. Learning during lifetime consists of a modification of the connections and a varying weight of each connection before being processed in the cell body. Artificial neural networks are used in many areas of engineering. The main applications can be divided into pattern recognition and regression.

2.8.2 Multilayer Perceptron

The multilayer perceptron is one of the widely used artificial neural networks, especially suitable for regression analysis. It consist of an input layer a certain number of hidden layers and an output layer. Each neuron in the input layer represents a single input parameter. The neurons of the input layer are connected to neurons of the first hidden layer. The number of hidden layers and the number of neurons is variable and should be chosen with respect to the complexity of the system response and the number of available training patterns. Only forward connections are allowed in a multilayer perceptron. The neurons in the final output layer represent the output parameter of the regression model. A schematic drawing of the layout is given in Fig. 50. The output a_i^l of neuron i in layer l is calculated as

$$a_i^l = h \left(\sum_{j=1}^{N_i^l} w_{ji}^{l-1} v_j^{l-1} + b_i^l \right), \quad (26)$$

where h is the activation function, N_i^l is the number of connections to the previous layer, w_{ji}^{l-1} corresponds to the weights of each connection and b_i^l is the bias, which represents the constant part in the activation function. In Fig.51 commonly used activation functions are illustrated.

2.8.3 Training

The training of the neural network is an optimization procedure, in which the weights and biases of the neural network are determined. For this purpose, a certain a number of training samples M with corresponding pairs of inputs $\mathbf{p}_i$ and outputs $\mathbf{o}_i$ is required. The mean square error F , which is the average value of the difference between the approximated response and the outputs, is used as objective in the optimization procedure.

$$F(\mathbf{w}, \mathbf{b}) = \frac{1}{M} \sum_{i=1}^M e_i \quad (27)$$

$$e_i = |\mathbf{a}(\mathbf{p}_i) - \mathbf{o}_i|^2 \quad (28)$$

In general, different training algorithms can be applied to solve this optimization procedure. The most important are the standard back-propagation algorithm [32], RPROP [33], the conjugate gradient method [34] and the scaled conjugate gradient algorithm [35]. In this paper, the Levenberg Marquardt algorithm [36] has been used, since for small systems with up to 1000 free parameters it was found to be faster compared to other methods. For all these methods, the gradient $\mathbf{g}$ of the objective function F with respect to the free parameters $\mathbf{x}$ (weights $\mathbf{w}$ and biases $\mathbf{b}$) is to be calculated. This can be performed with a variation of the back-propagation algorithm [37]:

$$\begin{aligned} \mathbf{G} = \frac{\partial F}{\partial \mathbf{x}_j} &= \frac{2}{M} \sum_{q=1}^M \sum_{i=1}^N e_i^q(\mathbf{x}) \frac{\partial e_i^q(\mathbf{x})}{\partial \mathbf{x}_j} \\ &= \frac{2}{M} \sum_{q=1}^M [\mathbf{J}^q(\mathbf{x})]^T \mathbf{e}^q(\mathbf{x}) \end{aligned} \quad (29)$$

$$\mathbf{J}^q(\mathbf{x}) = \begin{bmatrix} \frac{\partial e_1^q(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial e_1^q(\mathbf{x})}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial e_N^q(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial e_N^q(\mathbf{x})}{\partial x_n} \end{bmatrix}, \quad (30)$$

where $\mathbf{J}$ describes the sensitivity of the outputs with respect to the free parameters and N is the dimension of the output vector. The Hessian can be ex-

pressed in a similar way as

$$\begin{aligned}
 \mathbf{H} &= \frac{\partial^2 F}{\partial x_j \partial x_k} \\
 &= 2 \sum_{q=1}^M \sum_{i=1}^N \left[\frac{\partial e_i^q(\mathbf{x})}{\partial x_k} \frac{\partial e_i^q(\mathbf{x})}{\partial x_j} + e_i^q(\mathbf{x}) \frac{\partial^2 e_i^q(\mathbf{x})}{\partial x_j \partial x_k} \right] \\
 &= \frac{2}{M} \sum_{q=1}^M \left\{ [\mathbf{J}^q(\mathbf{x})]^T \mathbf{J}^q(\mathbf{x}) + 2\mathbf{S}^q(\mathbf{x}) \right\} \\
 \tilde{\mathbf{H}} &= \frac{2}{M} \sum_{q=1}^M \left\{ [\mathbf{J}^q(\mathbf{x})]^T \mathbf{J}^q(\mathbf{x}) \right\}.
 \end{aligned}$$

If we assume $\mathbf{S}^q$ to be small, the exact Hessian $\mathbf{H}$ can be replaced by $\tilde{\mathbf{H}}$. As a result, the update of the parameters in the Newton-iteration can be written as

$$\Delta \mathbf{x}^{(k)} = -\tilde{\mathbf{H}}^{-1} \mathbf{G}. \quad (31)$$

This approach requires the approximated Hessian $\tilde{\mathbf{H}}$ to be invertible. However, this cannot be assured, especially if a high number of neurons in the hidden layer is used and the size of the training set is small. In order to overcome this problem, an additional scalar parameter μ is added to all the diagonal elements of $\tilde{\mathbf{H}}$. In the limit case of $\mu = 0$ the algorithm converges to the Newton method (with the approximated Hessian), whereas for a large parameter μ the update can be approximated by a steepest descent algorithm with learning rate $\frac{1}{2\mu}$:

$$\Delta \mathbf{x}^{(k)} \approx -\frac{1}{2\mu} \mathbf{G}^{(k)}, \text{ for large } \mu. \quad (32)$$

The parameter μ is initially set to a small value (e.g. 0.01), the update $\Delta \mathbf{x}^{(k)}$ and the mean square error $F(\mathbf{x}^{(k)} + \Delta \mathbf{x}^{(k)})$ are calculated. If a reduction of the mean square error is obtained, the next iteration step $k+1$ is performed with $\mu^{(k+1)} = \mu^{(k)}/2$, otherwise the iteration step is repeated with a $\mu^{(k)}$ increased by a factor of 2.

The initial weights have been calculated from a uniform distribution in the interval $[-s, s]$ according to [38], where s is given by

$$s = \sqrt{\frac{3}{L_i}} \quad (33)$$

and L_i is the number of input connections of the neuron. The initial biases are set to zero. Due to the random initialization the training process was repeated 10 times to decrease the influence of the starting values of the weights.

In order to reduce the influence of over-fitting, an early stopping criterion [39] is applied. The data set is divided into a training and validation set. The update of the weights and biases is stopped, if the required accuracy of the mean square error for the training samples is reached, the mean square error

for the validation set starts to increase or the norm of the gradient of the objective function is smaller than a prescribed value.

In general, two strategies for learning can be applied - sequential learning and batch mode learning. In the first approach, the training samples are presented to the learning algorithm separately in a stochastic order (e.g. randomly) and the free parameters are updated for each of these training samples to reduce the difference between the network output and the training sample. In the second approach, the average error for the whole training set is calculated and the update of the free parameters is performed for the full set at once. In this investigation, the batch mode was applied, since the convergence speed of the method increased dramatically compared to the sequential approach.

In the same way, the number of neurons in the single hidden layer is calculated. Starting with only three neurons, the number of neurons is iteratively increased until no significant decrease of the objective function is obtained in the validation data set.

3 STOCHASTIC SAMPLING

3.1 INTRODUCTION

For a global variance-based sensitivity analysis it is recommended to scan the design space with latin hypercube sampling and to estimate the sensitivity with the multivariate statistic based on surrogate models. Results of a global sensitivity study are the global sensitivities of the optimization or random variables due to important responses. So, it is possible to identify the sub domains for optimization and reliability analysis.

3.2 LATIN HYPERCUBE SAMPLING

Latin hypercube sampling (LHS) is an advanced Monte Carlo simulation and a further development of the stratified sampling methodology. In order to reduce the necessary number of samples, each class of any random variable is considered in the same manner ([40]). First, the marginal distribution or cumulative distribution function X_i is subdivided into N classes $\mathbb{D}_j$ with the same probability

$$P[x_i \in \mathbb{D}_j] = \frac{1}{N}, \quad i = 1, \dots, n, \quad j = 1, \dots, N$$

So N^n hypercubes are created with the probability N^{-n} . Using random permutations of the indices i and j , we obtain N samples. For each class, one representative value can be selected random based, or is defined using the mean value for this class ([41]).

Latin hypercube sampling is qualified for simulation of samples for the robustness evaluation. In comparison with the plain Monte Carlo simulation, it requires fewer sampling points to represent the design space. Furthermore, LHS can be used as a selec-

tion procedure to find support points in the experimental design. Whenever, for systematic sampling methods, the number of required support points is unacceptable, LHS is especially suitable to reduce the number of support points.

Spurious linear correlations between the random variables are reduced by using a Cholesky decomposition of the target correlation matrix ([42]). This algorithm requires that the number of samples is higher than the number of stochastic variables $N > n$. If this is not the case this algorithm is not applied, which implies that spurious correlations occur with the same order of magnitude as by using plain Monte Carlo Simulation. This effect can be seen in Table 2. In contrast to the reduced spurious linear correlations the magnitudes of higher order correlation errors, e.g. quadratic dependencies, are not reduced by using this type of Latin Hypercube Sampling.

3.3 LATIN HYPERCUBE SAMPLING WITH MINIMIZED LINEAR CORRELATION ERRORS

Alternative to the standard Latin Hypercube sampling with Cholesky decomposition the spurious linear correlations can be minimized by an internal optimization procedure using a column-pairwise-switch algorithm. This approach works independently of the sample size, thus it is also applied for $N \leq n$. In Table 2 the differences in the correlation errors between both LHS types are given. Due to the internal optimization procedure this advanced Latin Hypercube sampling requires a significant higher numerical effort than the standard method. Based on Table 3 we recommend to use this sampling method only up to 500 samples. For a higher number of samples the computation is quite expensive and the benefit is negligible compared to the standard LHS scheme.

Additional to the generation of an initial design this advanced LHS method can be used to generate additional samples to an existing sample set. Any sampling type of the existing set can be considered, the final set is obtained by minimized the correlation errors of the merged set of initial and additional samples. If the existing set is also generated by Latin Hypercube sampling, the optimal number of additional samples is the double amount of the initial number.

4 MULTIVARIATE STATISTIC

4.1 ANALYSIS OF CORRELATION

Various statistical analysis procedures are available for the subsequent evaluation of correlation of input parameters and the responses. For example, the coefficients of correlation are calculated from all pairwise combinations of both input variables and re-

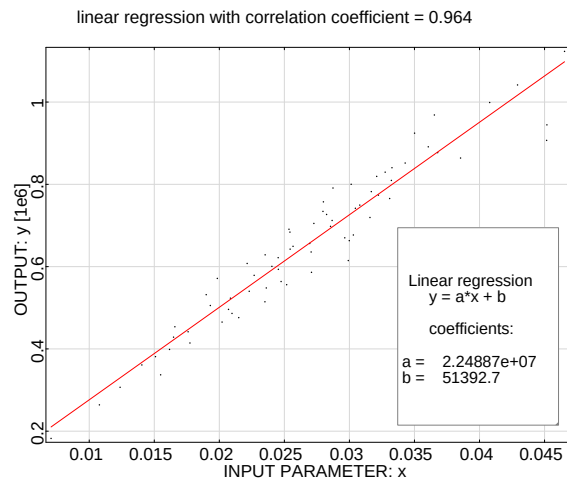


Figure 9: Given set of samples of input parameter $x_i = x$ and response $x_j = y$, linear regression function with correlation coefficient of $\rho_{ij} = 0.964$.

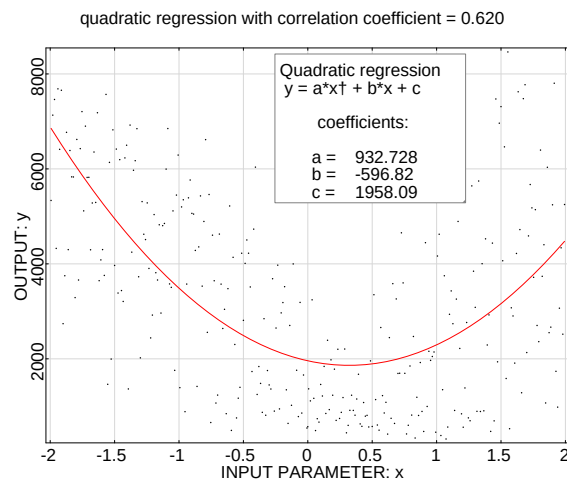


Figure 10: Given set of samples of input parameter $x_i = x$ and response $x_j = y$, quadratic regression function with correlation coefficient of $\rho_{ij} = 0.62$.

sponse according to:

$$\rho_{ij} = \frac{1}{N-1} \frac{\sum_{k=1}^N (x_i^{(k)} - \mu_{x_i})(x_j^{(k)} - \mu_{x_j})}{\sigma_{x_i} \sigma_{x_j}} \quad (34)$$

The quantity ρ_{ij} , called the linear correlation coefficient, measures the strength and the direction of a linear relationship between two variables, as shown in Figure 9. The linear correlation coefficient is sometimes referred to as the Pearson product moment correlation coefficient. The quadratic coefficients of correlation

$$\rho_{ij} = \frac{1}{N-1} \frac{\sum_{k=1}^N (\hat{y}^{(k)}(x_i) - \mu_{\hat{y}(x_i)})(x_j^{(k)} - \mu_{x_j})}{\sigma_{\hat{y}(x_i)} \sigma_{x_j}}$$

is defined as the linear coefficient of correlation (see Equation (34)) between the least-squares fit of a quadratic regression $\hat{y}(x_i)$ of the variable x_j and x_j themselves on the samples $x_i^{(k)}, x_j^{(k)}$, as shown in Figure 10. A correlation greater than 0.7 is generally described as strong, whereas a correlation less than 0.3 is generally described as weak. These values can vary based upon the type of data being examined. All pairwise combinations (i, j) , values can be assembled into a correlation matrix C_{XX} , as shown in Figure 52.

4.2 ANALYSIS OF PREDICTION

The coefficient of determination

$$R_j^2 = \frac{\sum_{k=1}^N (\hat{y}^{(k)}(x_i) - \mu_{X_j})^2}{\sum_{k=1}^N (x_j^{(k)} - \mu_{X_j})^2}$$

with $i = 1, \dots, n$ is a value which indicates the shaking (variance or fluctuation) of responses j of the approximation model, depending on the regression model terms. It is a measure that allows to predict the dependence of a response value from a set of input parameter $i = 1, \dots, n$ in a special regression model context. This R^2 value varies between 0 and 1. The coefficient of determination represents the percent of the data that is the closest to the regression model best fit. For example, $R^2 = 0.868$ based on a linear regression, which means that 86.8% of the total variation in the response value y can be explained by the linear relationship between the regression model containing input parameter. However, a high value of R^2 not always implies a good regression model, because adding of additional approximation model terms increase the coefficient.

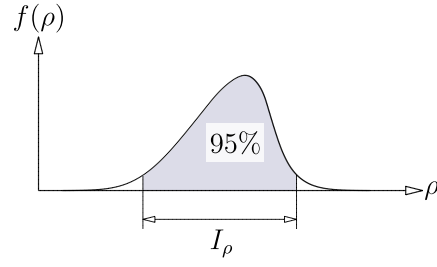


Figure 11: Confidence Interval for coefficient of correlation ρ .

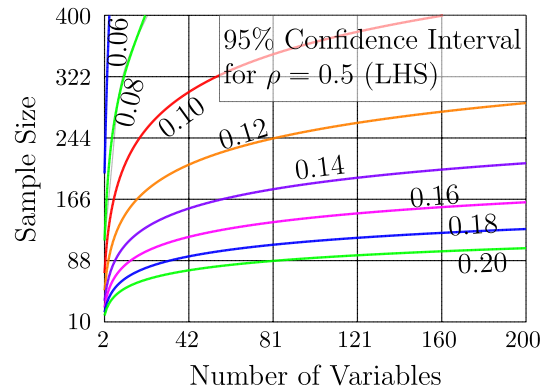


Figure 12: 95% Confidence intervals to estimate a coefficient of correlation of $\rho = 0.5$ using LHS.

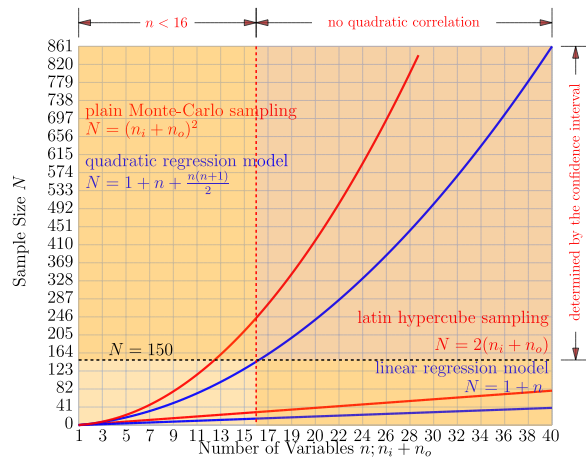


Figure 13: Costs of the robustness evaluation depending on the sampling method, the number of input and output parameters n_i, n_o and the regression model to estimate the prediction values. A commonly used maximal sample number is $N = 150$. So, in case of a full polynomial regression model the number of input parameters is restricted to $n < 16$.

Adjusted coefficient of determination This inadequacy leads to the *adjusted* R^2 coefficient

$$R_{adj}^2 = 1 - \frac{N-1}{N-p} (1 - R^2)$$

who not increase with the number of model terms for a small sample size N . Whereat N is the number of sample points and p the number of regression coefficients. In fact the coefficient decreases with unnecessary model terms. With the comparison of R^2 and R_{adj}^2 it is possible to predict the regression model. A high difference between the coefficients indicates that unnecessary terms are included in the model.

Coefficients of importance An important prediction value to explain the influence of a single input parameter l on a chosen output parameter j , depending on the regression model is the coefficients of importance

$$COI_{jl} = R_j^2 - \frac{\sum_{k=1}^N \left(\hat{y}^{(k)}(x_i) - \mu_{X_j} \right)^2}{\sum_{k=1}^N \left(x_j^{(k)} - \mu_{X_j} \right)^2}$$

which an regression model $\hat{y}(x_i | i \in \{1, \dots, n\} \wedge i \notin \{l\})$. In addition the adjusted coefficients of importance is given by

$$COI_{jl}^{adj} = 1 - \frac{N-1}{N-p} (1 - COI_{jl})$$

4.3 COSTS OF THE ROBUSTNESS EVALUATION

4.3.1 Minimal sample size

In order to obtain stable statistics for large linear correlation coefficients only, there is a relation between the number of design variables $n = n_i$, the number of responses n_o and the necessary number of samples. The simulation procedure will produce N samples for all parameter x_i denoted by x_i^k ; $i = 1 \dots n$; $k = 1 \dots N$. In general it is recommended that the number of samples N be at least equal, but better higher than the number n of design variables.

The recommended minimum number of samples depends on the number of input and output parameters and is given by $N = (n_i + n_o)^2$ for plain Monte Carlo sampling and $N = 2(n_i + n_o)$ for latin hypercube sampling, as shown in Figure 13. This rough estimation may be sufficiently for a relative small number of input parameters up to $n < 40$. A more precise estimation is given as a results of a convergence study of the confidence intervals.

4.3.2 Confidence intervals

Of course, the correlation coefficients are random values themselves. Whereby the variance depends on the number of the calculated samples N . According to a specified confidence interval I_p of e.g. 95% the possible lower and upper bounds of the estimated coefficients of correlation ρ_{ij} can be evaluated, as shown in Figure 11.

The confidence intervals for the estimated coefficients of correlation ρ_{ij} in Figure 12 are computed based on the Fisher's z -transformation. The interval for a significance level of α (i.e. a confidence level of $1 - \alpha$) is given by

$$\left[\tanh \left(z_{ij} - \frac{z_c}{\sqrt{N-3}} \right), \tanh \left(z_{ij} + \frac{z_c}{\sqrt{N-3}} \right) \right]$$

In this Equation, N is the number of samples used for the estimation of ρ_{ij} . The critical value z_c is computed by using the Bonferroni-corrected value for the significance level $\alpha' = \alpha/k$ with k being the number of confidence tests. The transformed variable z is computed from

$$z_{ij} = \frac{1}{2} \log \frac{1 + \rho_{ij}}{1 - \rho_{ij}} \quad (35)$$

and the critical value z_c is given by

$$z_c = \Phi^{-1}(1 - \alpha'/2) \quad (36)$$

where $\Phi^{-1}(\cdot)$ is the inverse cumulative Gaussian distribution function.

In order to study the effect of latin hypercube sampling on the reduction of statistical uncertainty, a comparison of the estimation errors (standard deviations) of the correlation coefficients is carried out. The most important domain in the range of $2 < n < 200$, coefficient of correlation $\rho = 0.5$ is detailed shown in Figure 12. For example, for $n = 160$ input parameters and a tolerated maximum error of 14 percent according to a 95%-confidence interval the necessary number of samples is $N = 200$ using latin hypercube sampling.

4.3.3 Cost of the regression model

The regression models for the coefficient of importance require a minimum number of samples

$$N = 1 + n + \frac{n(n+1)}{2} \quad (37)$$

depending on the constant, linear and quadratic terms of the regression model. So, in case of a full polynomial regression model and an acceptable maximal sample number of $N = 150$ the number of input parameters is restricted to $n < 16$.

To eliminate this limitation, other surrogate models e.g. reduced polynomial regression model, moving least square approximation, artificial neural networks and support vector regression are introduced in the previous sections of this paper.

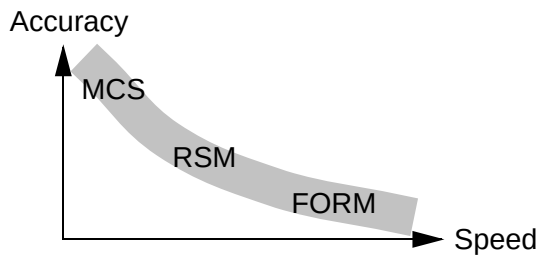


Figure 14: Performance of methods for stochastic analysis. a.) Monte Carlo simulation (MCS), b.) Response Surface Method (RSM), c.) First Order Reliability Method (FORM).

4.4 SIGNIFICANCE FILTER AND REDUCED POLYNOMIAL REGRESSION MODEL

The meta-modeling approaches reported in the literature are usually based on the assumption that a simulation model generates a single, i.e. scalar value response. But most complex engineering simulations yield to multiple responses. To reduce the required number of support points a possible approach is to create a separate meta-model for each response individually.

A significance filter to reduce the number n of relevant input parameters can be based on the differences between the user-defined and sampled quadratic and linear input correlation matrix C_{XX} . Figure 52 shows the linear correlation matrix as a result of the latin hypercube sampling approach. In this example the user-defined input correlation matrix is simple the unite matrix. So, the histogram of the differences can be calculated, as shown in Figure 54, and can be fitted by a probability density function with the 95% – 99%-fractiles of the coefficients of correlation. These quantiles are used as a significance filters for the relevant input parameters with a coefficients of correlation larger than these bounds. Result is a reduced polynomial regression model of the responses, as shown in Figure 53.

5 RELIABILITY ANALYSIS

5.1 INTRODUCTION

Numerical methods, e.g. for structural analysis have been developed quite substantially over the last decades. In particular, finite element methods and closely related approximations became the state of the art. The modeling capabilities and the solution possibilities lead to an increasing refinement allowing for more and more details to be captured in the analysis. On the other hand, however, the need for more precise input data became urgent in order to avoid or reduce possible modeling errors. Such errors could eventually render the entire analysis procedure useless. Typically, not all uncertainties encountered in structural analysis can be reduced by careful modeling since their source lies in the intrinsic

randomness of natural phenomena. It is therefore appropriate to utilize methods based on probability theory to assess such uncertainties and to quantify their effect on the outcome of structural analysis.

Different methods exist, but they all have a limited area of application. Although Monte Carlo methods are most versatile, intuitively clear, and well understood, the computational cost (the number of computational runs required) in many cases is prohibitive. Thus approximations become important, which can be based e.g. on the response surface methods (RSM) or first/second order reliability methods (FORM/SORM). For the feasibility of response surface approaches it is quite essential to reduce the number of variables to a tractable amount. This may require extensive sensitivity analysis in order to identify the relevant random variables. This is particularly important for random variables arising from discretization of random fields or processes. In this context, close coupling between the tools for stochastic and computational analyses is essential. Simplifications can be based on the following items

- Global variance-based sensitivity or robustness analysis of the structural response with respect to the random variables. Again, this aims at reducing the number of random variables needed.
- Concentrate random sampling in the region which contributes most to the total failure probability. This is generally called “importance sampling”. It is important to note that most importance sampling strategies work best with a low number of random variables.
- Approximation of the numerical response by a class of simple mathematical functions. This is the so-called “response surface method”. Again, it is vital that the number of random variables be kept small.

As a very simplistic rule-of-the-thumb, Fig. 14 gives the accuracy/speed ratio for some solution methods as mentioned above. However, there are situations in which MCS can be comparably fast or FORM can be comparably accurate.

5.2 FORMULA OF THE FAILURE PROBABILITY

Safety and reliability analysis warrants the exclusion of damage and design collapse during the life time. Probability of surviving is the numerical quantity of safety and reliability and the probability of failure is the complement.

We can define any undesired or unsafe state of a response as an event $\mathcal{F}$ out of the set of all random variables $\mathbf{X}$ such a way that the assign state function $g(\mathbf{x})$ is less or equal to zero. Generally, failure (i.e. an undesired or unsafe state of the response) is defined in terms of a limit state function $g(\cdot)$, i.e.

by the set $\mathcal{F} = \{\mathbf{X} : g(\mathbf{X}) \leq 0\}$. Frequently, $Z = g(\mathbf{X})$ is called *safety margin*. As indicated

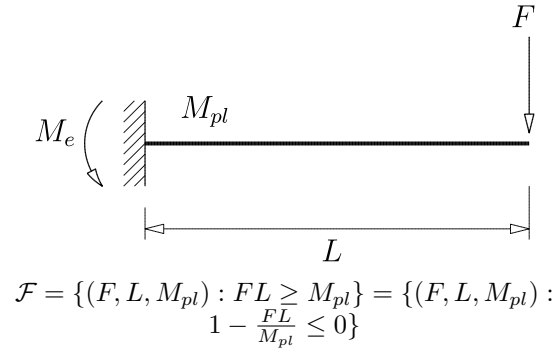


Figure 15: Structural system and several unique failure conditions.

in Fig. 15, the definition of the limit state function is not unique. The failure probability is defined as the probability of the occurrence of the unsafe event $\mathcal{F}$:

$$P(\mathcal{F}) = P[\{\mathbf{X} : g(\mathbf{X}) \leq 0\}] \quad (38)$$

This quantity is *unique*, i.e. not depending on the particular choice of the limit state function.

The response behavior near the failure state is most important in the reliability analysis. The random design parameters, such as loadings, material parameters and geometry, are the set of basic random variables $\mathbf{X}$ which determine the probabilistic response of numerical model. The failure condition is defined by a deterministic limit state function

$$g(\mathbf{x}) = g(x_1, x_2, \dots, x_n) \leq 0$$

as shown in Fig. 10. The failure probability of a design is given by

$$P(\mathcal{F}) = P[\mathbf{X} : g(\mathbf{X}) \leq 0] = \int_{g(\mathbf{x}) \leq 0} \dots \int_{-n}^n f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \quad (39)$$

where $f_{\mathbf{X}}(\mathbf{x})$ is the joint probability density function of the basic random variables.

5.3 RESPONSE SURFACE METHODS FOR RELIABILITY ANALYSIS

5.3.1 Global Polynomial Approximation

Normally, the state function $g(\mathbf{X})$ of a system response is described implicitly, e.g. through an algorithmic procedure within finite element analysis. Alternatively, the original state function can be approximated by a response surface function $\tilde{g}(\mathbf{x})$ which has polynomial form ([43–48]). Additionally, the limit state function $g(\mathbf{x}) = 0$ themselves can be interpolated by second order polynomials ([49, 50]). One of the major advantages of the response surface method lies in its potential to selectively determine the number of structural analyses of the support points. This is especially helpful if some overall knowledge on the system behavior - particularly

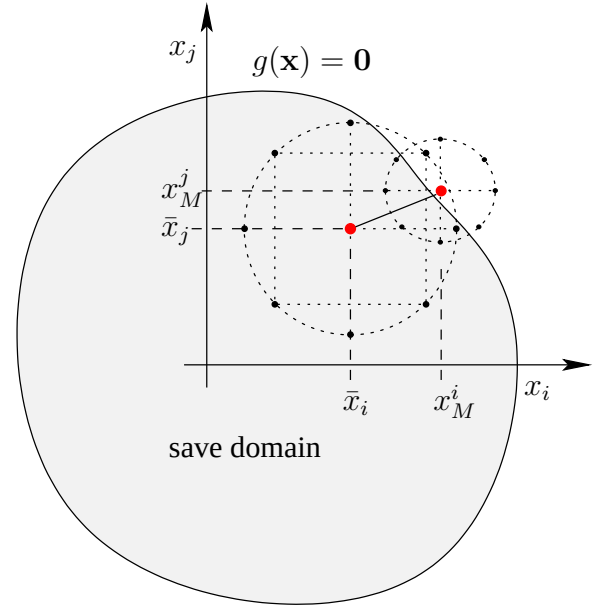


Figure 16: Adaptive design of experiment in the random space.

near to the failure region - is a priori available. By such means the computational effort can be substantially reduced.

On the other hand, the global approximation schemes widely used in the application of the response surface method can be quite misleading due to the lack of information in certain regions of the random variable space. Standard second order polynomial approximations are not sufficiently flexible. So, the estimation of the failure probability using this global approximation leads to large errors, in particular for small failure probabilities $P(\mathcal{F}) < 10^{-2}$ and a number of random parameters of $n > 5$. It is therefore required to avoid such undesirable approximation errors at reasonable computational effort.

5.3.2 Adaptive Response Surface Method

A commonly used approximation method with minimized the regression error within the support point values is the moving least square method. The main advantage of this method is the flexibility for the approximation of highly nonlinear state and limit state functions. The proposed method is suitable for computing the reliability of complex models and is intended to provide reasonably accurate estimates of failure probabilities while maintaining computational efficiency.

Adaptive design of experiment In particular, these response surfaces can be adaptively refined to consistently increase the accuracy of the estimated failure probability. This is especially suitable for the reliability analysis of complex nonlinear structures. An arbitrary number of check points even in high local concentration can be used without approxima-

tion problems. Using deterministic design of experiment, the necessary number of support points become very high with an increasing number of random variables.

To decrease the number of support points in an optimized way, the so called D-optimality criterion is used. A discussion of this criterion is presented by [51]. The effectiveness of a design in satisfying the minimum variance (D-Optimal) criterion is expressed by the D-Efficiency of the design. In [12], more exact specifications of the D-Optimal criteria and further criteria called alphabetic optimality criteria are described. However, the first design of experiment in the first iteration should explore the random space including safe and unsafe domain as accurate as possible. A possible approach is given in [52] with

$$x_i = \bar{x}_i \pm f\sigma_{x_i}$$

whereby

$$f = \Phi^{-1}(P(\mathcal{F})) = 3, \dots, 5$$

is a factor depending on the assumed failure probability. [45] give an efficient possibility to adaption a design of experiment in the next iterations with

$$\mathbf{x}_M = \bar{\mathbf{x}} + (\mathbf{x}_D - \bar{\mathbf{x}}) \frac{g(\bar{\mathbf{x}})}{g(\bar{\mathbf{x}}) - g(\mathbf{x}_D)}$$

with

$$\mathbf{x}_D = E[\mathbf{X}|g(\mathbf{x}) \leq 0]$$

as shown in Figure 5.3.1. This is achieved by a combination of random search strategies (based on the adaptive sampling approach) as well as deterministic search refinement. In such a way for the most practical examples 3 till 6 iteration steps are necessary for a sufficient convergence. So this adaptive design of experiment using a D-optimal linear or quadratic design in combination with the improved moving least square approximation is suitable up to $n \leq 20$ random parameters.

6 RANDOM FIELDS

6.1 INTRODUCTION

Any mechanical structure possesses some natural randomness in its properties which fluctuates over space: deviations from the design geometry, surface roughness, scatter of material properties and distributed loads are few examples. There exist attempts to model uncertainties by few random variables which are generated by CAD programs. This approach is meaningful for few special problems only. For a realistic interpretation of such random fluctuations within an optimization, robustness or reliability analysis, they have to be modelled as random fields. To distinguish the random field approach from the previous one, it is called it non-parametric here.

In a computer implementation, random processes or fields are discretized into a finite set of random variables. The number of variables, however, can be considerably high. Most of the accurate methods for computing structural reliability (such as advanced Monte Carlo Methods) as well as dependable approximate methods (e.g., the Response Surface Method) are sensitive towards the number of variables involved. The present paper proposes a method to model a random field with a reduced set of variables. Robustness evaluation is employed for this purpose, which relates the stochastic input and output quantities and thus helps to identify the most relevant variables. Unlike previous approaches (sketched in sec. 6.3), it does not rely on purely stochastic considerations, but takes into account the structural behaviour as well.

6.2 PROPERTIES

A random field is, in brief, a random function $H(\mathbf{r})$ defined on a spatial structure. The vector $\mathbf{r} \in \mathbb{R}_{Str.}$ points to a location on the structure. Random fields are used, e.g., to study random fluctuations in geometrical or material properties of a mechanical or structural system. In other words, the considered property is a random variable at each point on the structure. Moreover, the random properties at two different locations can be mutually correlated among each others.

Any random variable is characterized by a probability distribution function, which can be parameterized by distribution type and stochastic moments. For random fields, the moments become functions over space as well. *From now on, a Gaussian (or Normal) distribution type is assumed.* In this case, the characterization by first and second moments provides the full information. In particular

$$\mu_H(\mathbf{r}) = E[H(\mathbf{r})] = \int_{-\infty}^{+\infty} h f_H(\mathbf{r}, h) dh \quad (40)$$

denotes the *mean function*, and

$$\begin{aligned} R_{HH}(\mathbf{r}_1, \mathbf{r}_2) &= E[H(\mathbf{r}_1) \cdot H(\mathbf{r}_2)] \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h_1 h_2 f_H(\mathbf{r}_1, \mathbf{r}_2, h_1, h_2) dh_1 dh_2 \end{aligned} \quad (41)$$

the *correlation function*, with $E[\cdot]$ being the expected value operation. R_{HH} is a function of the distance between two points and indicates the amount of linear dependency between the random properties at these locations.

The so-called *correlation length* L_{HH} , which is actually the centre of gravity of the correlation function, is a typical characteristic of R_{HH} . It has to be estimated from manufacturing processes, natural

scatter of material properties, etc. An infinite correlation length yields a structure with random properties, yet without fluctuations within the structure. A zero correlation length yields uncorrelated (in case of the Gaussian distribution independent) variables. Three special cases are important for the further studies. *Homogeneity*: A random field is said to be homogeneous in the wide sense, if the first and second moments are the same at any possible location, i.e.

$$\mu_H(\mathbf{r}) = \mu_H(\mathbf{r} + \boldsymbol{\xi}) \quad \forall \boldsymbol{\xi} \quad (42)$$

$$R_{HH}(\mathbf{r}_1, \mathbf{r}_2) = R_{HH}(\mathbf{r}_1 + \boldsymbol{\xi}, \mathbf{r}_2 + \boldsymbol{\xi}) \quad \forall \boldsymbol{\xi} \quad (43)$$

Isotropy (in the wide sense) claims that the correlation function depends on the distance between the two observed locations $\mathbf{r}_1$, $\mathbf{r}_2$ only, not on the direction:

$$R_{HH}(\mathbf{r}, \mathbf{r} + \boldsymbol{\xi}) = R_{HH}(\|\boldsymbol{\xi}\|) \quad (44)$$

In case of *orthotropy* the correlation function is a product of two or more independent functions defined on orthogonal axes:

$$R_{HH}(\mathbf{x}, \mathbf{y}) = R(\mathbf{x}) \cdot R(\mathbf{y}) \quad (45)$$

6.3 MODELLING

For computational analyses, a random field has to be discretized in order to yield a finite set of random variables $\mathbf{X}$, which are assigned to discrete locations on the observed structure. Since the Finite Element Method is the standard for structural analyses, it is convenient to discretize the random field in the same way as the finite element model. One speaks of Stochastic Finite Elements in this case. The discretization can be oriented at the element mid points, integration points, or nodes. The properties of the random variables are derived from the random field properties explained previously. The spatial discretization should be able to model the variability of the random field. For this purpose, it has been recommended by Der Kiureghian and Ke [53], Hisada and Nakagiri [54] that the distance between two discretization points should be not more than $1/4$ of L_{HH} .

The set of random variables is then characterized by a mean vector and a correlation matrix. It is convenient for the developments that follow to use the covariance matrix instead, which is defined as

$$\mathbf{C}_{XX} : c_{ij} = R_{HH}(\mathbf{r}_i, \mathbf{r}_j) - \mu_H(\mathbf{r}_i) \cdot \mu_H(\mathbf{r}_j) \quad (46)$$

The joint density of all random variables can be modelled with help of the Nataf model [55, 56], given the type and properties of the marginal distributions for each variable.

From now on, *random fields with zero mean vector are considered*. Then the covariance matrix suffices for the characterization of the random variables set.

Random number generators can produce independent random variables only. For the assumed case of Gaussian distributed variables, independence is equivalent to zero correlation. It can be shown that the random variables will be uncorrelated after the following transformation. The covariance matrix is decomposed with help of an eigenvalue analysis:

$$\boldsymbol{\Psi}^T \mathbf{C}_{XX} \boldsymbol{\Psi} = \text{diag}\{\lambda_i\} \quad (47)$$

Therein, $\boldsymbol{\Psi}$ is the matrix of eigenvectors of $\mathbf{C}_{XX}$ stored columnwise, and the eigenvalues are identical to the variances of the uncorrelated random variables Y_i : $\lambda_i = \sigma_{Y_i}^2$. The transformation rule reads

$$\mathbf{Y} = \boldsymbol{\Psi}^T \mathbf{X} \quad (48)$$

and the backward transformation

$$\mathbf{X} = \boldsymbol{\Psi} \mathbf{Y} \quad (49)$$

because the eigenvectors $\boldsymbol{\Psi}$ form an orthonormal basis. Hence it is possible to model the random field by a Karhunen-Loève expansion of the random field [57] which consists of a sum of deterministic shape functions $\boldsymbol{\Psi}$ multiplied by the respective uncorrelated random amplitudes $\mathbf{Y}$.

The eigenvalues are usually stored sorted by magnitude in descending order, which is a measure for their contribution to representing $\mathbf{C}_{XX}$. This opens a way of reducing the usually huge number of variables. Only the random variables with the highest variances are employed for a later Monte Carlo simulation. The quality of approximation of the random field is expressed by the variability fraction [58]

$$Q = \frac{\sum_{i=1}^n \sigma_{Y_i}^2}{\text{trace}(\mathbf{C}_{XX})} ; \quad 0 \leq Q \leq 1 \quad (50)$$

The number of the random variables considered has to be adjusted before the simulation in order to reach a sufficient quality, e.g. $Q > 0.9$. However, although the random field approximation seems good, it may be not suitable for a reliability analysis. Schorling [59] reported this experience for the application to stability analysis of a randomly imperfect shell structure. The truncated series explained above is not suitable, because the criterion is purely stochastic and does not account for the structural behaviour. On the other hand, purely mechanical considerations may not work as well, if they fail to represent the random field properly.

This is the motivation for the development of a procedure which selects those mode shapes and random variables, which contribute “most” – by means of stochastic influence – to a certain structural performance, here the buckling load. The Robustness Analysis, as described briefly before, offers tools for this purpose.

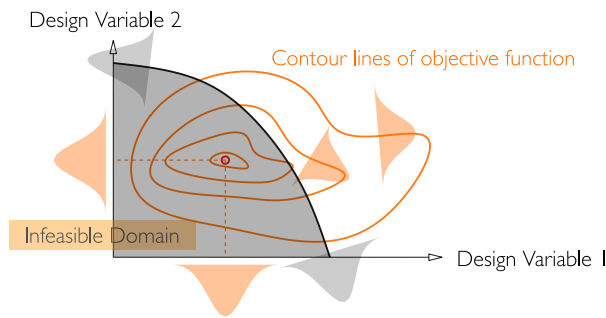


Figure 17: Sources of uncertainty in design optimization.

7 ROBUST DESIGN OPTIMIZATION

7.1 CHALLENGES ON VIRTUAL PROTOTYPING AND MULTIDISCIPLINARY OPTIMIZATION

During the last years, many challenges on virtual prototyping have been occurred. Product life cycles are expected to last for as little as a few months, and more and more customized products are developed, e.g. 1700 car models compared to only 900 ten years ago. The engineer's focus is more and more on "built-in-quality" and "built-in-reliability". The products are developed in the shortest amount of time, and, inspired by that, they have to be safe, reliable and robust. Some markets require optimized product designs to be robust, e.g. defense, aerospace, jet engine, nuclear power, biomedical, oil industry and other mission critical tasks.

At the same time, the numerical models become increasingly detailed and numerical procedures become more and more complex. Substantially more precise data are required for the numerical analysis. Commonly, these data are random parameters. From this it follows that the optimization process includes uncertainties or stochastic scatter of design variables, objective function and restrictions as shown in Figure 17. Furthermore, the optimized designs lead to high imperfection sensitivities and tend to loose robustness. Using a multidisciplinary optimization method, the deterministic optimum design is frequently pushed to the design space boundary. The design properties have no room for tolerances or uncertainties. So the assessment of structural robustness, reliability and safety will be more and more important. Because of that, an integration of optimization and stochastic structural analysis methods is necessary.

7.2 DESIGN FOR SIX SIGMA

Six Sigma is a quality improvement process to optimize the manufacturing process in a way that it automatically produces parts conforming to the Six Sigma quality level, as shown in Figure 7.1. Motorola documented more than \$16 Billion in sav-

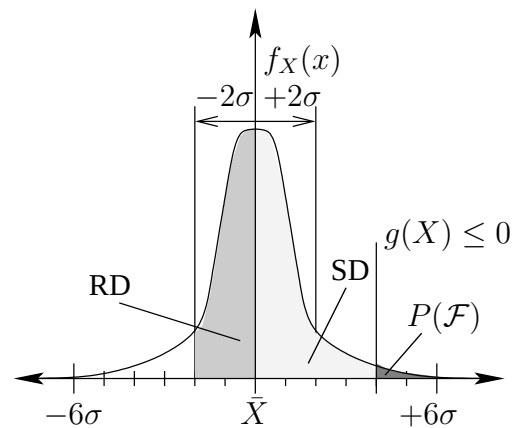


Figure 18: Normal distribution $f_X(x)$ with lower and upper specification limit on 2σ and 6σ level. Robust design (RD) and safety design (SD) ($\geq \pm 2\sigma$) depending on chosen limit state function $g(X) \leq 0$, e.g. stress limit state.

ings as a result of their Six Sigma efforts¹. Since then, hundreds of companies around the world have adopted Six Sigma as a way of doing business.

In contrast, Design for Six Sigma optimizes the design itself such that the part conforms to Six Sigma quality even with variations in manufacturing, i.e. quality and reliability are explicit optimization goals, as shown in Figure 57. Robust design is often synonymous to "Design for Six Sigma" or "reliability-based optimization". The possible sigma levels start at $1,2\sigma$ (robust design optimization) and go up to 6σ (reliability-based design optimization) ([60]), as shown in Table 5.

Within the robust design optimization, the statistical variability of the design parameter is considered. The most general method for solving robust design optimization problems is the well established Monte Carlo simulation method. However, the major shortcoming of this approach is its vast need of computational resources (the number of solver runs required), and these cannot be presumed in general situations.

Optimized designs within the sigma level $\leq \pm 2\sigma$ are characterized as robust design (RD). The objective of the robust design optimization (e.g. [61–63]) is to find a design with a minimal variance of the scattering model responses around the mean values of the design parameters (see [64, 65]).

Other approaches for an evaluation of the design robustness, e.g. the linear approximation of "scattering" solver responses (see e.g. [66]) or the variance estimation in genetic programming (see e.g. [67, 68]), independently of given parameter distributions will not be subject of the following remarks as they are not to be counted to robust design optimization methods in a stricter sense.

In the reliability-based optimization, the optimiza-

¹source: www.isixsigma.com/library/content/c020729a.asp

tion problem can be enhanced by additional stochastic restrictions ensuring that prescribed probabilities of failure can not be exceeded. Furthermore, the probability of failure itself can be integrated into the objective function. Frequently, the search for the optimum by means of deterministic optimization is combined with the calculation of the failure probability, e.g. using the first- order second-moment analysis (FOSM) (e.g. [69]). A more promising combination may under certain circumstances involve the first and second order reliability methods (FORM/SORM) (e.g. [70–72]).

Within the deterministic optimization, a calculation of the failure probability of individual designs has to be performed in order to be able to properly evaluate these designs. Therefore, special attention has to be paid to the cost efficiency of this calculation. As an example, for smooth and well-scaled objective functions with few continuous design parameters, the deterministic optimization as well as the determination of the failure probability that is included within the optimization iteration loop may be performed by means of gradient based programming (e.g. Sequential Quadratic Programming, see [73]).

In [74] a decrease of the numerical expense of these two nested iterations is attempted by substituting the deterministic objective function as well as the limit state function on which the point of largest probability density is searched within FORM by a single objective function in a hybrid design space. However, this leads to an enlargement of the design space for the gradient based programming.

In the reliability-based optimization, frequently approximation function are applied that at the same time approximate the design space and the space of random parameters by means of a meta-model, e.g. in [4–7]. Successful industrial applications of these methods can amongst others be found in [75].

In [76], a linear approximation of the limit state function serves as a constraint of the optimization problem. An improvement of the optimization result is tempted in [77] by taking into account the gradients of the limit state function.

However, in the robust optimization (see [78, 79]) as well, different approximation models in combination with an appropriate variance determination are used, e.g. global polynomial approximations and Kriging models. Their use is restricted to problems with few random variables and few optimization variables ($n \leq 5$).

Reliability-based robust design optimization In reliability-based design optimization, the determin-

istic optimization problem

$$\begin{aligned} f(d_1, d_2, \dots, d_{n_d}) &\rightarrow \min \\ g_k(d_1, d_2, \dots, d_{n_d}) &= 0; \quad k = 1, m_e \\ h_l(d_1, d_2, \dots, d_{n_d}) &\geq 0; \quad l = 1, m_u \\ d_i &\in [d_l, d_u] \subset \mathbb{R}^{n_d} \\ d_l &\leq d_i \leq d_u \\ d_i &= E[X_i] \end{aligned} \quad (51)$$

with n_r random parameters $\mathbf{X}$ and n_d means of the design parameters $\mathbf{d} = E[\mathbf{X}]$ is enhanced by additional m_g random restrictions

$$\int_{g_j(\mathbf{x}) \leq 0}^{n_r} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} - P(\mathbf{X} : g_j(\mathbf{X}) \leq 0) \leq 0; \quad j = 1, m_g \quad (52)$$

with the joint probability density function of the basic random variables $f_{\mathbf{X}}(\mathbf{x})$ and m_g limit state functions $g_j(\mathbf{x}) \leq 0$ (see Figure 10). The probability of failure in (52) is calculated applying the reliability analysis.

Furthermore the objective function can be enhanced by additional criteria such as minimization of the probability of failure $P(\mathcal{F})$

$$f(d_1, d_2, \dots, d_{n_d}, P(\mathcal{F})) \rightarrow \min \quad (53)$$

with

$$P(\mathcal{F}) = \int_{g_j(\mathbf{x}) \leq 0}^{n_r} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \quad (54)$$

Variance-based robust design optimization

Within the robust design optimization, the objective (51) is enhanced by the requirement to minimize the variances $\sigma_{X_i}^2$

$$f(d_1, d_2, \dots, d_{n_d}, \sigma_{X_1}^2, \sigma_{X_2}^2, \dots, \sigma_{X_{n_r}}^2) \rightarrow \min \quad (55)$$

with

$$\sigma_{X_i}^2 = \frac{1}{M-1} \sum_{k=1}^M (x_i^k - \mu_{X_i})^2$$

7.3 ADAPTIVE RESPONSE SURFACES FOR NONLINEAR OPTIMIZATION

7.3.1 Introduction

The response surface methodology (RSM) is one of the most popular strategies for nonlinear optimization. Due to the inherent complexity of many engineering optimization problems it is quite alluring to approximate the problem and to solve the optimization in a smooth sub-domain by applying response surface methodology.

Usually, for a large number of real-life design optimization problems the objectives and constraints are determined as a result of expensive numerical computations. Furthermore, the function values and their derivatives may contain numerical noise and

the calculability of some of the response functions is domain-dependent, e.g. situations when these functions cannot be evaluated at some points of the design space. Especially, to solve this kind of optimization the *adaptive response surface methodology* (ARSM) (see e.g. [66, 80–83]) is developed as a consequent combination of optimization strategy and response surface methodology.

Of course the accuracy of the approximation compared to the real problem has to be checked and verified. Mainly three factors influence the accuracy of a response surface:

1. The number and distribution of support points. The systematic sampling schemes try to place the support points in an optimized way according to the boundary of the design space and the distance of the support points. For reasonably smooth problems, the accuracy of response surface approximations improves as the number of points increases. However, this effect decreases with the degree of oversampling.
2. The choice of the approximation function. In general, higher order functions are more accurate. Linear functions require fewest support points, but are weak approximations. Quadratic functions are most popular. The second order polynomial results in a smooth approximation function and is well scaled for gradient based optimizers. Using polynomials higher than second order may only result in higher local accuracy with many sub-optima.
3. The design space. The overall possible design space is given by the lower and upper boundaries of the optimization parameters. Of course the smaller the approximated subregions, the greater the accuracy. In practical problems we will start with the overall design space and further investigate smaller subregions.

In contrast to the RSM the ARSM uses a subregion of the global parameter range to approximate the responses. Starting with a presumably large subregion the iteration moves and shrinks the subspace till a solution converges to an optimum. This strategy will be denoted as *move limit strategy*. Usually, this is done using low level trial functions (e.g. linear and quadratic polynomial functions). The convergence of the adaption can be improved using advanced moving least square approximation which is very efficient by means of a reduction of the response surface variance using additional support points in the near of the optimal design.

8 NUMERICAL EXAMPLES

8.1 META-MODELING

8.1.1 1D test function

In following examples the approximation quality of the different surrogate models is investigated. The first example is given to analyse the convergence of the approximation quality of the different surrogate models depending on the number N of support points on a non-convex function of the Shepard type:

$$\hat{y}(\mathbf{x}) = \frac{\sum_{i=1}^n y(\mathbf{x}_i) \left(\frac{1}{\|\mathbf{x} - \mathbf{x}_i\| + \epsilon} \right)^2}{\sum_{i=1}^n \left(\frac{1}{\|\mathbf{x} - \mathbf{x}_i\| + \epsilon} \right)^2}$$

with $x_i = 0.75 + z_i$, $z_i = \{-15; -8; -6; -3; -1; 0; 1; 4; 7; 15\}$ and $y_i = \{1; -10; -2; 5; -1; 8; 15; -4; 7; 1\}$, as shown in Figure 56. As meta-models Moving Least Squares with regularized (MLS-R) and exponential (MLS-E) weighting function, Support Vector Regression (SVR) and Multilayer Perceptrons (MLP) are investigated. The optimal influence radius of the exponential weighting function is determined based on an adaptive D -approach. The results indicate, that for the deterministic case the SVR approximation converges much faster than these of MLP and MLS. For a relative large number of support points MLP, SVR and MLS the approximation is almost identical.

8.1.2 2D test function

In this example the well-known two-dimensional test function [84]

$$f(\mathbf{x}) = 0.01 \sum_{i=1}^2 (x_i + 0.5)^4 - 30x_i^2 + 20x_i \quad (56)$$

with $-6 \leq x_i \leq 6$ for the deterministic case and for the case of noisy data values is investigated. For the latter purpose the original function is modified by adding Gaussian noise with zero mean and standard error equal to one to each support point value. In the Figures 19 and 20 the two functions are shown. The test data set, which is of the same size as the support data set, is modified in the same manner. Both data sets are obtained by Latin Hypercube Sampling with uniform distribution. The approximation quality is evaluated by the normalized mean error of 10000 regular distributed evaluation points. In the Figures 21 and 22 the obtained relative mean errors are shown depending on the number of support points. The presented values are the average of 100 random sets for each configuration. As meta-models the polynomial regression, Moving Least Squares with regularized (MLS-Reg) and exponential (MLS-Exp) weighting function, Support Vector Regression (SVR) and Multilayer Perceptrons

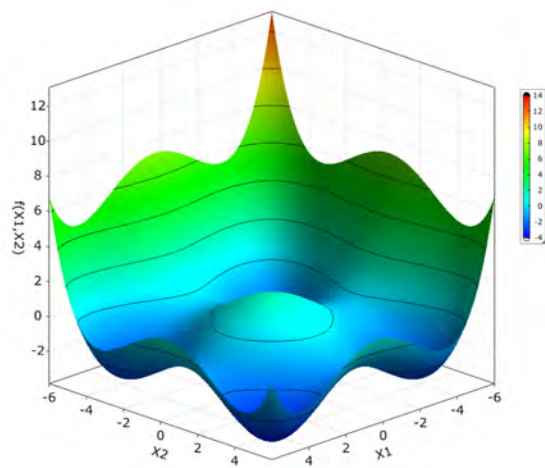


Figure 19: Deterministic 2D test function

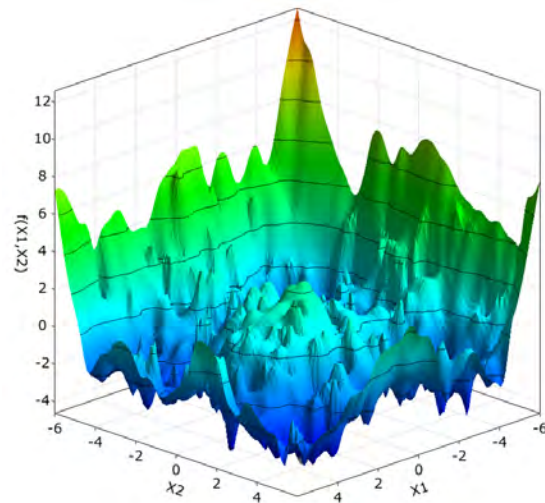


Figure 20: 2D test function with additional Gaussian noise

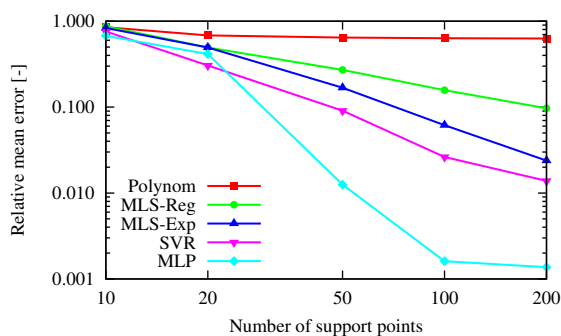


Figure 21: Relative mean errors for the deterministic 2D function

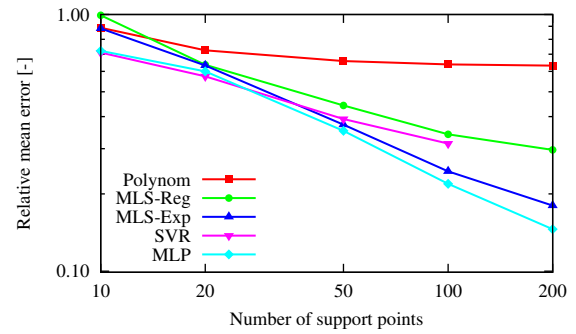
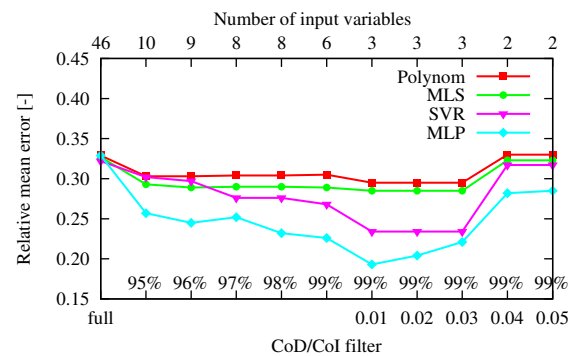


Figure 22: Relative mean errors for the 2D function with additional noise

Response 1



(MLP) are investigated. The optimal influence radius of the exponential weighting function is determined based on the test data set. The results indicate, that for the deterministic case the MLP approximation converges much faster than these of SVR and MLS. For the noisy function the convergence of MLP, SVR and MLS is almost identical. As expected the polynomial regression can not converge to the exact result due to the high complexity of the function.

8.1.3 Industrial application

In this example the applicability of the presented filters and approximation models is investigated for

Response 2

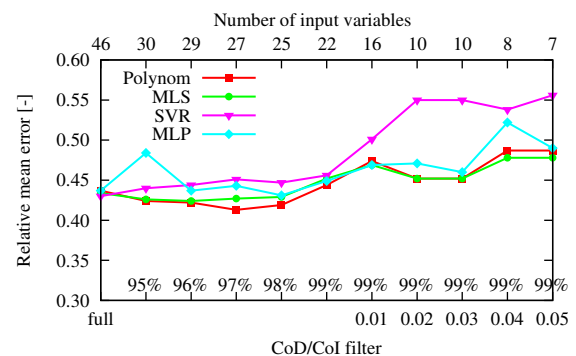


Figure 23: Industrial application: approximation errors for the surrogate models depending on the applied filter parameters

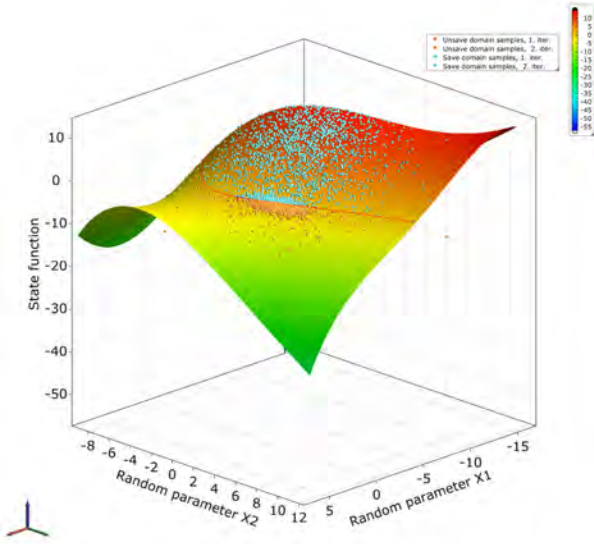


Figure 24: The Quattro Function. Analytical state function $g(\mathbf{x})$ and limit state function $g(\mathbf{x}) = 0$ in combination with direct importance adaptive sampling.

a real problem, where we have 46 input variables and several output quantities which are anonymous. Again, first the significance and importance filter is applied and the remaining set of input variables is used to build up the approximation model. In total, 100 support points for the surrogate model and 100 test samples to determine the model parameters and to evaluate the approximation error are used. In Figure 23 the calculated relative approximation errors for two representative responses are shown depending on the filter parameters. The Figure indicate a significant reduction of the approximation errors for all surrogate models for response 1, whereby the MLP approach gives the best results. The approximation errors for response 2 can be reduced slightly only for the polynomial based approaches for a significance filter of 97% which corresponds to 27 remaining input variables. Further reduction of the variable number leads to increasing approximation errors. The SVR and MLP approximation errors are for the reduced systems always larger as for the full system. This shows, that not for all cases a significant better result can be achieved. This may be caused by several reasons, e.g. that response 2 contains a large noise fractions or that too many input variables are important for the approximation. The usage of an increased number of sampling points could possibly solve this problem.

8.2 RELIABILITY ANALYSIS

In most practical applications the characteristics of the state function and the limit state function as e.g. its shape are not known a priori. Furthermore the limit state function is often not an analytic function but derived from a numerical model. It is there-

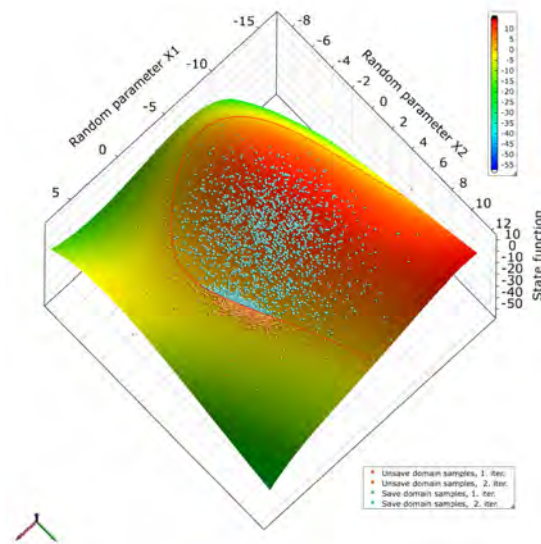


Figure 25: The limit state function $g(\mathbf{x}) = 0$ is a high-curved nonlinear one. The random parameters X_1 and X_2 are normal distributed variables with mean $\bar{X}_1 = -3.9$, and $\bar{X}_2 = 1.5$ and standard deviation $\sigma_{X_1} = \sigma_{X_2} = 1$.

fore necessary to investigate the performance of the adaptive response surface method with respect to the following criteria:

- probability or β level
- number of random variables n
- multiple β -points
- state function (high- or slow-curved, continuously differentiable, noisy)
- limit state function (high- or slow-curved, continuously differentiable, noisy)

Within this example the different results of the presented adaptive response surface method in comparison with a classic global response surface method are given. The state function

$$g(x_1, x_2) = -\left(\frac{x_1}{6} + \frac{1}{2}\right)^4 - \left(\frac{x_2}{8} - \frac{6}{5}\right)^4 - \left(\frac{2}{9}x_1 + \frac{1}{2}\right)^3 \left(\frac{x_2}{8} - \frac{6}{5}\right) + \left(\frac{x_1}{6} + \frac{1}{2}\right)^2 + \left(\frac{x_2}{8} - \frac{6}{5}\right)^2 - \frac{5}{3}x_1 - \frac{3}{4}x_2 + \frac{11}{5}$$

is a two-dimensional fourth order polynomial, as shown in Figure 24. Furthermore, the limit state function $g(\mathbf{x}) = 0$ is a high-curved nonlinear one, as shown in Figure 25. The random parameters X_1 and X_2 are normal distributed variables with mean $\bar{X}_1 = -3.9$, and $\bar{X}_2 = 1.5$ and standard deviation

$\sigma_{X_1} = \sigma_{X_2} = 1$. In order to obtain the reference failure probability a direct importance adaptive sampling is done with $N = 2000$ samples, two iterations and an initial variance multiplier $\sigma^{(i=1)} = 3$. The given failure probability is $P(\mathcal{F}) = 4.8907 \cdot 10^{-6}$ with a standard error of $\sigma_{\bar{P}(\mathcal{F})} = 1.5813 \cdot 10^{-7}$. In addition, the same adaptive sampling procedure is used to calculate the failure probability on the response surfaces.

For this example, the approximation using the global 2nd order polynomial approximation and a global central composite design of experiment leads to an error of 2268000 % in calculation the failure probability, as shown in the Figures 26 and 27. Using the global 2nd order polynomial approximation a wrong most probability failure domain is identified using importance adaptive sampling.

Applying the new adaptive response surface method to this state function, as shown in Figures 28 till 31, leads to accurate estimation of the failure probability already after the first adaption with $N = 18$ state function evaluations. In summary, using three adaptations of the central composite design with in total $N = 36$ state function evaluations the error of the given failure probability is 6% only (see Table 4 for details).

8.3 RANDOM FIELD

8.3.1 Introduction

As an example for random fields, the reliability of a cylindrical shell structure with random imperfections is studied. Within this example, the imperfections are discretized by Stochastic Finite Element methods [53, 57, 58]. It is demonstrated, how Robustness Analysis is employed in order to identify the most relevant random variables. The probability of failure is computed by Directional Sampling [85–87] as a reference, and by the Adaptive RSM in connection with Adaptive Monte Carlo [88].

8.3.2 Structure and Random Field Properties

The reliability of a geometrically imperfect shell with properties as given in Fig. 32 is analysed. The cylinder has a Navier-type support along the top and bottom edges and is loaded along the top edge with a continuous vertical unit load. The structure is modelled with a 9-node isoparametric shell finite element type within the program *Stang* [89].

A random field is applied on the structure in order to model geometrical imperfections. The random properties are coordinate deviations from the perfect structure in the cylinder's radial direction. Thus the random field is discretized at the nodes of the finite element mesh. It has zero mean and a standard deviation of $\sigma_H = 10^{-3}$ mm, which is roughly a hundredth of the radius. The orthogonal field has different correlation functions along the perimeter and height as plotted in Fig. 33. The spatial correla-

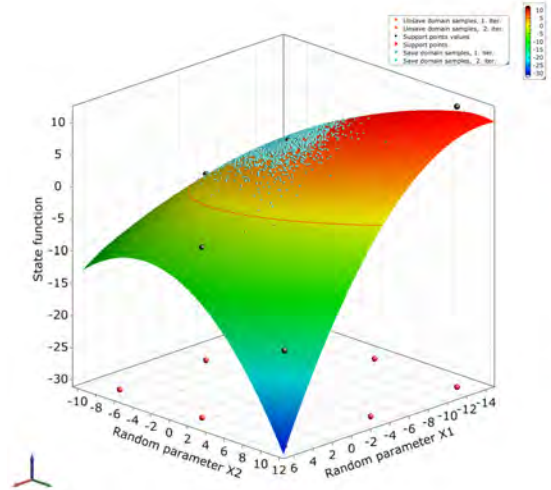


Figure 26: Approximated state function $g(\mathbf{x})$ and limit state function $g(\mathbf{x}) = 0$ using the global 2nd order polynomial approximation.

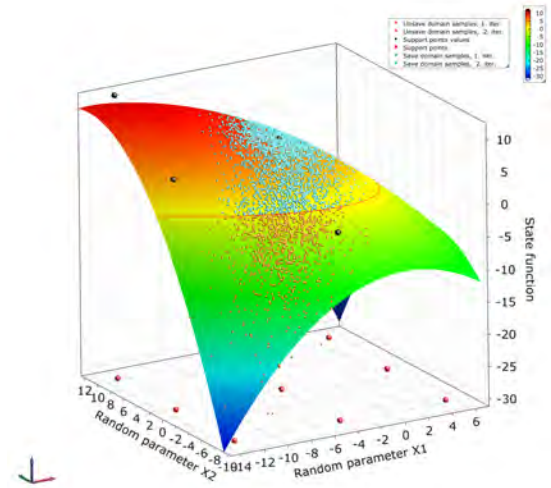


Figure 27: A wrong most probability failure domain is identified using importance adaptive sampling.

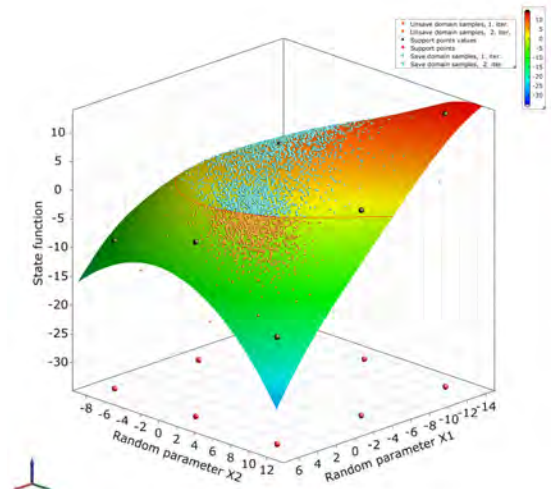


Figure 28: Approximated state function $g(\mathbf{x})$ and limit state function $g(\mathbf{x}) = 0$ using MLS.

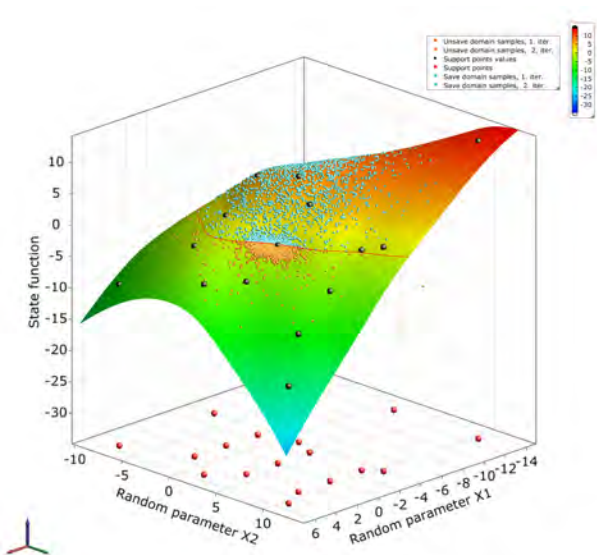


Figure 29: Adaptive moving least square approximation with first and second design of experiment.

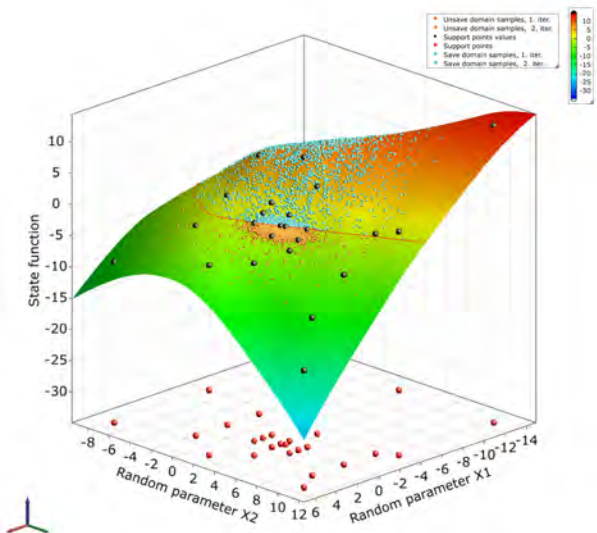


Figure 30: Adaptive moving least square approximation with first till third design of experiment.

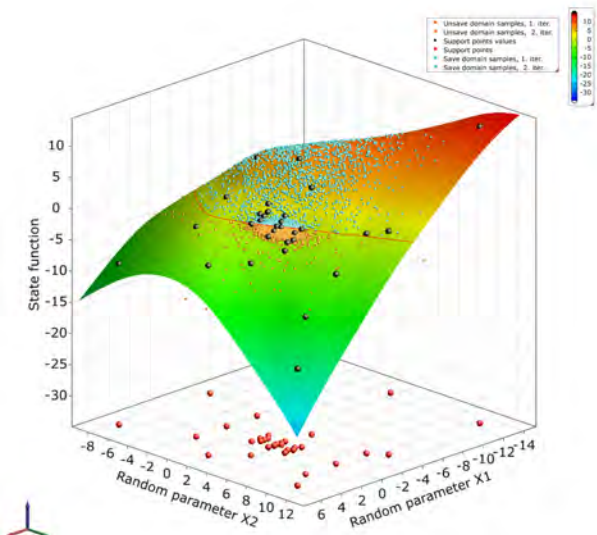


Figure 31: Adaptive moving least square approximation with first till fourth design of experiment.

Wall thickn. [mm]	0.197
Radius [mm]	101.6
Height [mm]	139.7
Young's mod. [N/mm ²]	$6.895 \cdot 10^4$

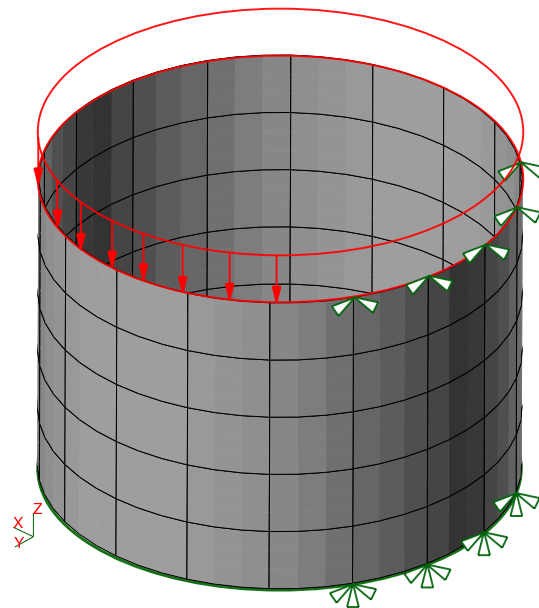


Figure 32: Finite element model of the cylindrical shell with schematic loads and supports.

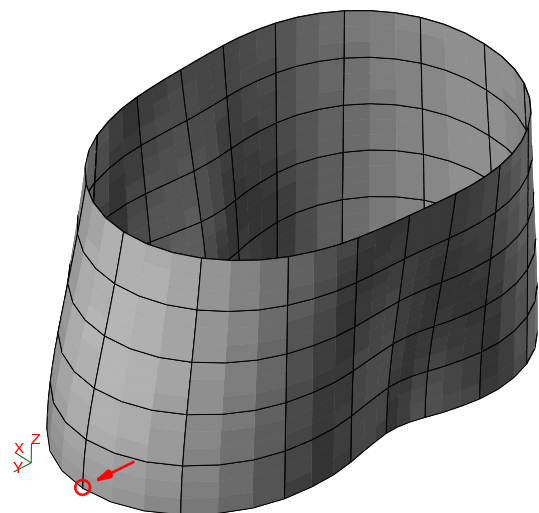
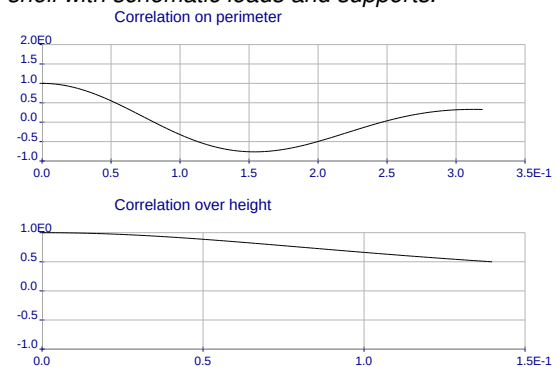


Figure 33: Correlation functions over perimeter (above) and height (below) and spatial correlation structure with respect to the marked node (right).

tion structure with respect to one node on the bottom edge is visualized, too.

The reliability of the imperfect structure towards stability failure shall be studied here. The method of analysis to be applied has to be a compromise between accuracy and computing time. For the example observed here, which shows a pre-buckling behaviour close to linear and a sudden buckling failure, the linear buckling analysis suffices. It is very fast and not prone to being trapped on a postbuckling equilibrium state.

The limit load for reliability analysis is adopted from the buckling load of the perfect structure subtracted a “safety margin”. Two case studies at different safety levels have been carried out with limit state functions as follow:

$$g(\mathbf{X}) = \begin{cases} lF_{buckling} - 34 \text{ kN} \leq 0 & \text{case 1} \\ F_{buckling} - 30 \text{ kN} \leq 0 & \text{case 2} \end{cases} \quad (57)$$

Failure is defined as the event, that the limit state function takes values of less than zero.

8.3.3 Preliminary studies

A Robustness Analysis is performed, wherein all 396 variables of the random field are involved and, among others, their relations to the critical buckling loads of the simulated imperfect structures are examined.

No significant linear correlations could be found. Instead, strong quadratic correlations are observed between the first 14 random variables and the critical load (where “first” indicates those variables with the highest variances, i.e. highest eigenvalues after decomposition of the covariance matrix, cf. sect. 6.3). For variables of order higher than 14, the quadratic correlation coefficients are close to zero. The quadratic correlation matrix is displayed in Fig. 34. The nonlinear dependency becomes obvious by a scatter plot, in Fig. 35, e.g., of input variable no. 1 vs. the critical load.

Based on the quadratic regression for the buckling load, with each random variable set in for X successively, the coefficients of determination are computed. The sum of all values is less than 100 %. That means, the variance of the critical load cannot be fully explained by a quadratic relation to the input variables. The results are sorted and plotted as bar diagram, Fig. 36. The strongest influences can easily be identified. A closer look reveals that not all of the “first 14” variables (see above) are most relevant, but a few variables assigned to higher order eigenvectors of the covariance matrix as well. The eigenvalues and eigenvectors of the covariance matrix (used as random amplitudes and shape functions in the reliability analysis) which are selected by the criterion of the “top 14” coefficients of determination are that of order 1, 2, 5, 6, 15, 21, 22, 26, 29, 30, 32, 34, 83 and 197.

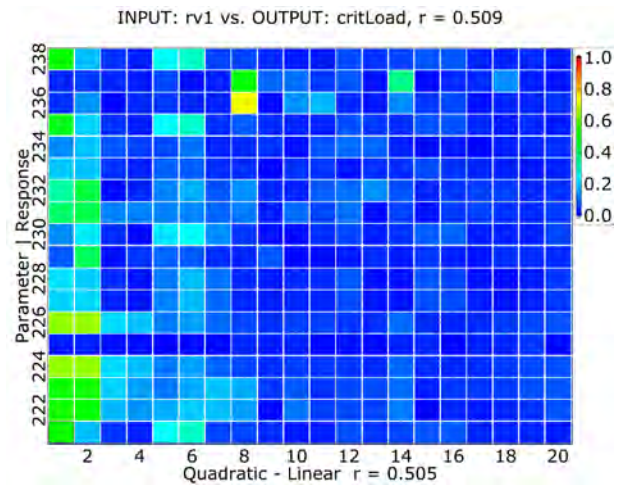


Figure 34: Matrix (part) of quadratic correlation coefficients. The lowest row shows correlations of the critical load with the first 20 random variables.

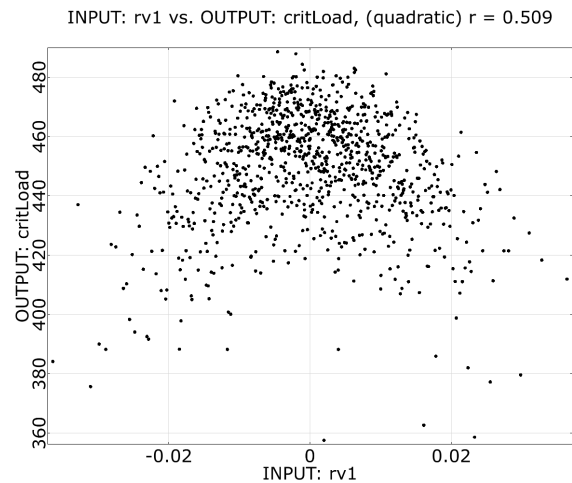


Figure 35: Anthill plot of critical load vs. first random variable.

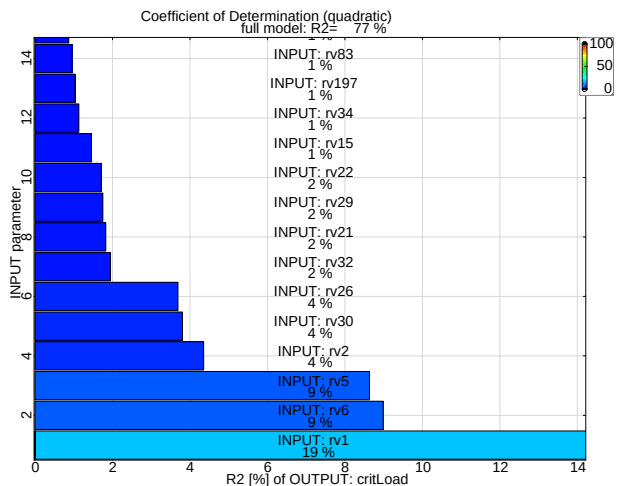


Figure 36: The top 14 coefficients of determination of critical load, quadratic regression model for random input variables.

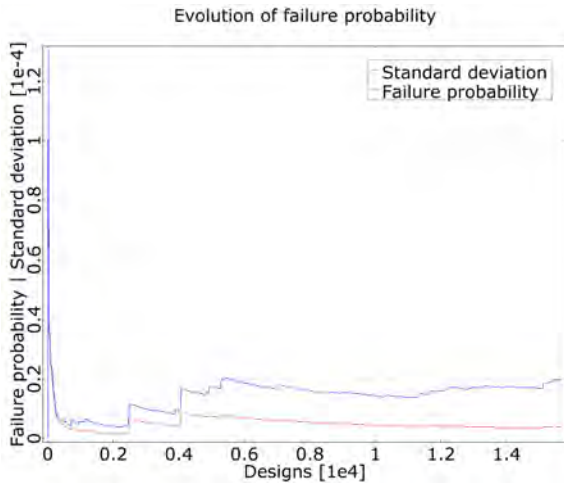


Figure 37: History of the failure probability and the standard deviation of the estimator – Directional Sampling.

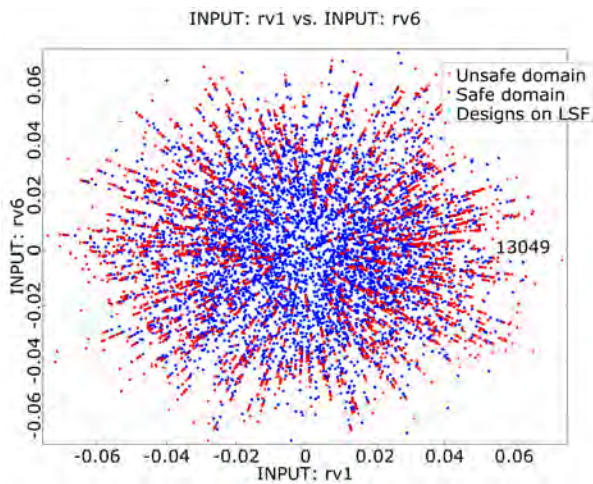


Figure 38: Two-dimensional anthill plot of the $N = 3000$ simulated directions in the subspace of the first and second random variable.

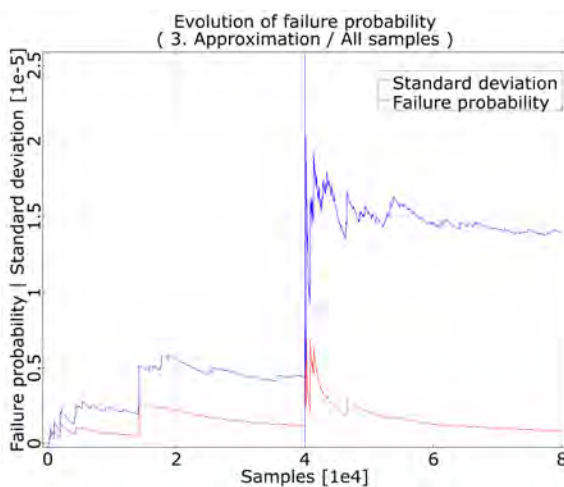


Figure 39: History of the failure probability and the standard deviation of the estimator – ARSM.

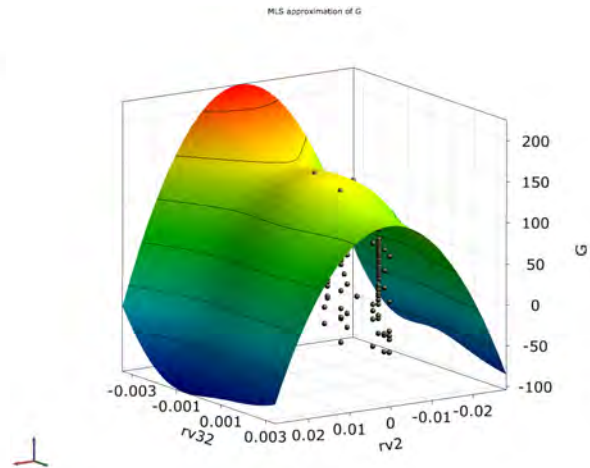


Figure 40: Moving Least Square approximation of the state function $g(x)$ in the subspace of variables 2 and 32.

8.3.4 Reliability Analysis – Case 1

For the first limit state condition as given in equation 57, the reliability of the structure is studied by means of a “plain” Monte Carlo simulation [as in 90], with the limit state function as defined by eq. (57). Three variants are computed: as a reference, the full representation of the random field, which employs 396 random variables, is used. Second, the “first 14” variables were selected with the criterion defined by eq. 50 and third, a set of random variables with the “top 14” coefficients of determination, cf. sect. 8.3.3. In each case, a sample with 36000 realizations is generated by Latin Hypercube Sampling [91–93]. No other variance reduction scheme such as Importance Sampling is applied. Because the random field defines the structural geometry and hence the structural behaviour, the limit state function cannot be programmed explicitly, but a linear buckling analysis has to be carried out for each sample.

The failure probabilities computed with the different sets of variables are listed in table 10. The so-called statistical error, i.e. the standard deviation of the estimator of the failure probability, is listed as well, in order to assess the confidence in the result. The simulation results with all variables and the “top 14” selection show a good quality. With the “first 14” set of variables, the probability of failure is underestimated by more than a magnitude. The statistical error in this case is unacceptably high, despite the huge sample size. This set of random variables is able to represent the random field well, but is not able to model the structural behaviour. The result obtained with the “top 14” selection is close to the reference, although it tends to be lower, too. Obviously, this selection criterion provides a good compromise for both modelling the stochastic and the mechanical problem.

8.3.5 Reliability Analysis – Case 2

For the second limit state condition of eq. 57, the reference result is computed by directional sampling with 3000 samples and in total $N = 15690$ finite element solver evaluations (see Figure 38) to give an estimated failure probability of $1.966 \cdot 10^{-5}$ and a standard deviation of the estimator of $3.927 \cdot 10^{-6}$. The history of the failure probability and the standard deviation of the estimator are shown in fig. 37. Applying the new adaptive response surface method to this state function, as shown in Figures 39 and 40, leads to an accurate estimation of the failure probability already after the first adaptation. ARSM starts with an initial D-optimal quadratic design of experiment with an initial axial multiplier of 2.0 and 4000 samples for each adaptive simulation on the surrogate model. For the second design of experiment only a D-optimal linear scheme is used. So in total only $N = 229$ finite element evaluations are necessary to estimate the failure probability of $1.392 \cdot 10^{-5}$.

8.4 EXAMPLE – ROBUST DESIGN OPTIMIZATION OF A DYNAMIC STRUCTURE

8.4.1 Structural system

The aim of the classical optimization problem for structural elements is to minimize the mass while observing deformation or strength restrictions. As an example for a robust design optimization the mass of this simple beam with rectangular cross section (w, h) is to be minimized subjected to a harmonic load. The central deflection w_d due to the dynamic load $F(t)$ has to be smaller than 5 mm . The objective function (i.e. the structural mass m)

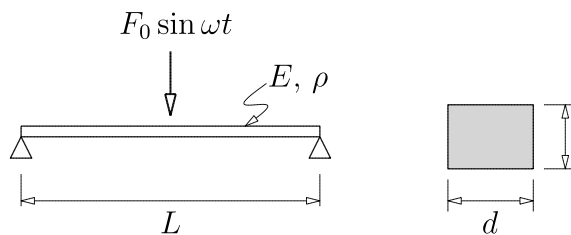


Figure 41: Beam with rectangular cross section

and the admissible area are displayed in Figure 42 for assumed values of $F_0 = 20\text{ kN}$, $\omega = 60\text{ rad/s}$, $E = 3 \cdot 10^{10}\text{ N/m}^2$, $\rho = 2500\text{ kg/m}^3$, $L = 10\text{ m}$ and $g = 9.81\text{ m/s}^2$.

Furthermore, in many application cases – especially concerning structural dynamics – the characterizing parameters are afflicted with stochastic uncertainties. In the present example it is assumed that the dynamic load amplitude F_0 and the excitation angular frequency ω are random variables with Gaussian distribution. The mean values correspond to the aforementioned nominal values, and both variational

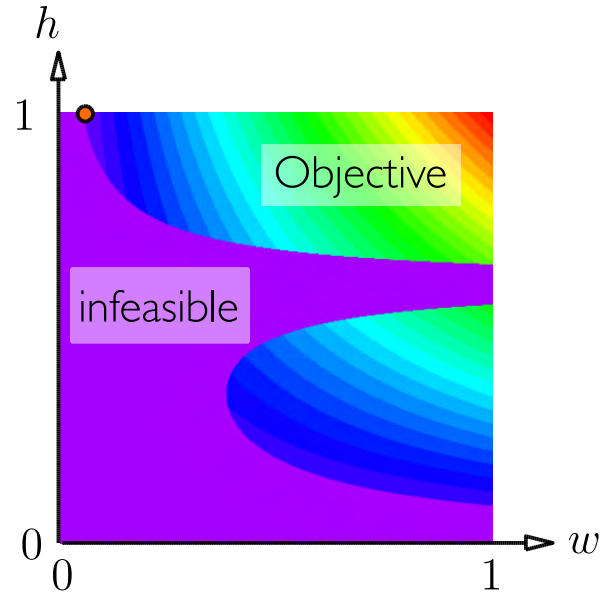


Figure 42: Deterministic objective and feasible design space. The deterministic optimal design is located in the red colored point ($w = 0.064$, $h = 1.00$).

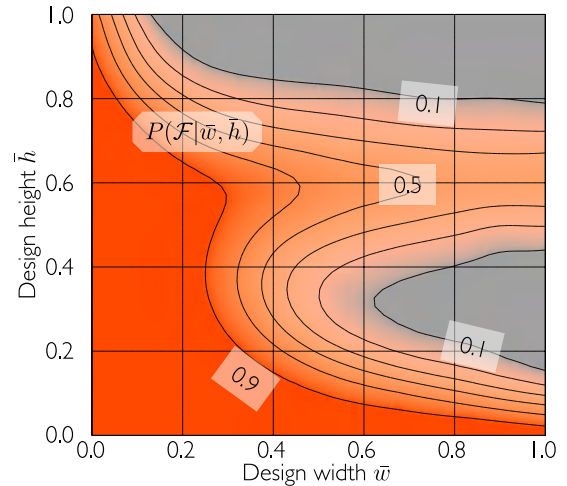


Figure 43: Compute conditional probability of violating constraint depending on h and w by using FORM.

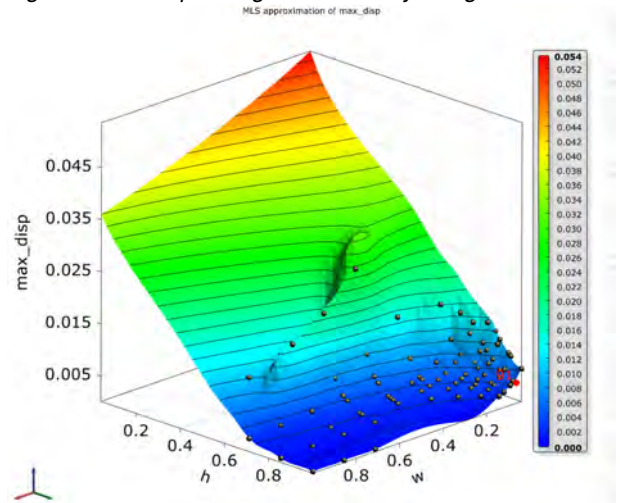


Figure 44: MLS approximation of the dynamic displacement w_d depending on h and w by using an adaptive design of experiment.

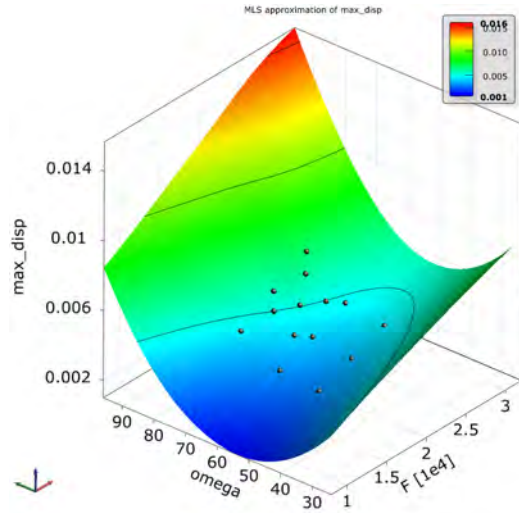


Figure 45: Adaptive design of experiment of random variable ω and F_0 during the reliability analysis. The limit state function $g(\mathbf{X})$ is defined as a limit state condition $w_d \leq 0.005$.

coefficients have been assumed to be 10%. This yields that the restriction from the dynamic load can only be met with a certain probability < 1 .

8.4.2 Deterministic optimisation

The deterministic optimization is based on an adaptive response surface method using the objective function

$$f(w, h) = h \cdot w \cdot L \cdot \rho$$

The deterministic optimal design is located in the red colored point ($m = 1600, w = 0.064, h = 1.00$). The required number of solver evaluations is $N = 91$ using linear polynomials, as shown in Fig. 49. With respect to the random load parameters the according failure probability is 11.45% and greater than the aimed value 1%. The reliability analysis is based on an ARSM with $N = 15$ design evaluations, as shown in Fig. 45 and Fig. 46.

8.4.3 Reliability-based robust design optimization

Within a reliability-based robust design optimization an additional stochastic constraint

$$P(\mathcal{F} : w_d \leq 5 \text{ mm}) - 0.01 \leq 0$$

can be used to obtain an optimal design with an acceptable failure probability. During the optimization procedure ARSM moves and shrinks the design parameter bounds, as shown in Fig. 47 and 48. Fig. 49 illustrates this adaptive design of experiments. In order to calculate the failure probability for each individual design using the ARSM proposed in a previous section additional design evaluations on the space of the random variables F_0 and ω are required. To obtain a better performance of the optimization process it is not necessary to calculate the failure probability in each nominal design

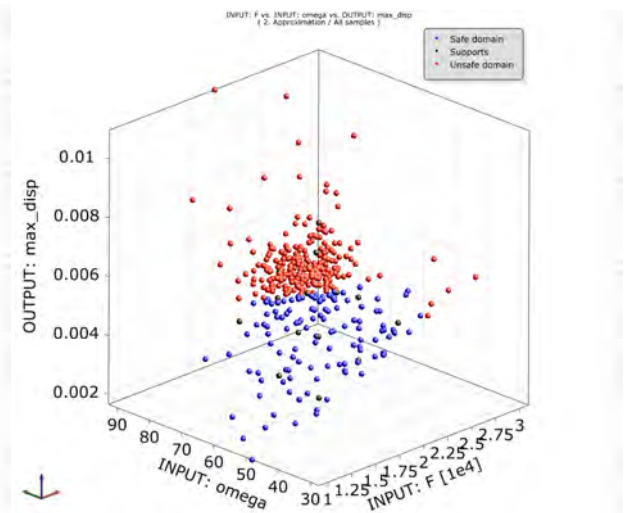


Figure 46: Adaptive sampling of random variable ω and F_0 during the reliability analysis based on the MLS approximation as shown in Fig. 45.

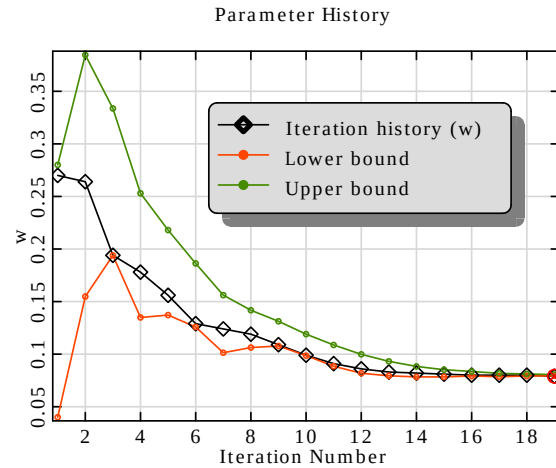


Figure 47: Adaptive modification of the design variable w and parameter bounds during the optimization.

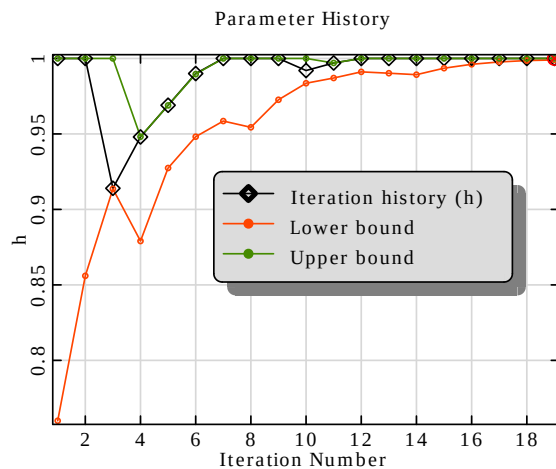


Figure 48: Adaptive modification of the design variable h and parameter bounds during the optimization.

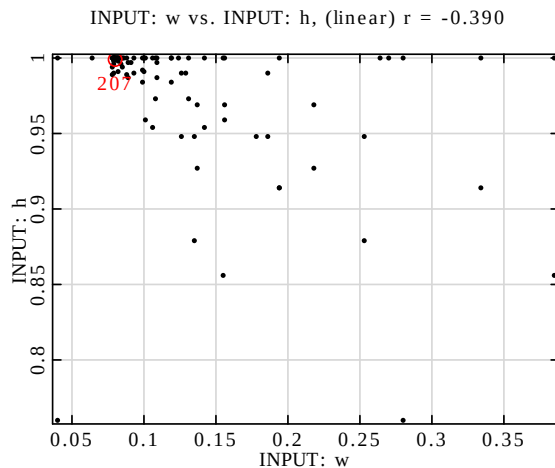


Figure 49: Adaptive design of experiment of design variable w and h during the optimization. The best robust design: $w = 0.0064$, $h = 0.220$ with a failure probability $0.98\% < 1\%$.

of the design space. Using the proposed advanced moving least square approximation a recycling of all previous calculated designs is recommended. So, in total $N = 152$ design evaluations are required, starting on the deterministic optimal design to receive a design with minimal mass with $m = 1998$ and an acceptable failure probability of $0.9\% < 1\%$.

9 CONCLUSIONS

In this paper several advanced surrogate models concerning their applicability in the framework of a robustness evaluation are investigated. Based on the results of the investigated numerical examples it can be summarized, it is not possible to define one best model type which is promising for all application. Artificial Neural Networks and Moving Least Squares work very nicely for low dimensional problems with less than 10 input variables. For higher dimensional problems the Support vector Regression method gives more accurate results especially if the number of available support points is relatively small. All types of models work similar for noisy data as for the deterministic case, if we determine the model parameters on the basis of an additional test data set. If this test data set is not available, the support point set can be subdivided in training and test data and after the model parameters have been determined the full data set can be used for the approximation. For the choice of the most proper surrogate model for a specific problem it is recommended to compare the approximation quality of the available models based on a test data set and select the best approach.

In many real applications not all input variables contribute significantly to the response functions. For this cases a selection of important variables and an approximation on the reduced variable set can increase the approximation accuracy dramatically. For

this purpose an efficient significance and important filter to identify the important input variables is developed. It can be seen that this filter combination works very reliable even for noisy response data and its application can improve the applied approximation model significantly. For real applications this stands for a dramatic reduction of the required computational costs within the robustness analysis.

In reliability and robustness analysis, imperfections of a mechanical or structural system, such as material properties or geometrical deviations, are modelled as random fields in order to account for their fluctuations over space. A random field normally comprises a huge number of random variables. The present paper proposes a method to reduce the random variables set. This reduction is performed on the basis on a robustness analysis. In this way, numerical difficulties can be avoided and the efficiency of the subsequent reliability analysis is enhanced. This so-called non-parametric structural reliability analysis is a new method to estimate the safety and reliability of finite element structures in such cases where a CAD-based parametrization is not possible or not meaningful.

In a computer application, the random field has to discretised. The discretisation points are typically nodes or integration points of the structural (finite element) model, i.e. the observed parameter (material or geometrical property) at each discretisation point is a random variable. As a consequence, the number of random variables can be very high, which inhibits the use of accurate methods of stochastic calculus, such as variance reducing Monte Carlo techniques or the Response Surface Method.

The dimension of problem shall be reduced while retaining the variability and correlation structure of the random field. The discretised random field is represented by the covariance matrix, which can be decomposed by the Karhunen - Loeve decomposition into a series of random amplitudes and an orthogonal set of deterministic shape functions. There exist criteria to choose a subset of variables which on one hand amount to a large variable fraction, but may not be relevant for the observed system performance.

A new concept utilizes statistical measures to relate the input variables with the system performance in order to identify the most important random variables and respective shape functions. The aim is to find a subset consisting of those variables, which cause the highest reaction of the system in respect of the chosen limit state criterion. It has been shown in test examples and real applications, that by this method it is possible to drastically reduce the problem dimension. The identification of relevant random field mode shapes alone already gives the designer valuable hints on critical points in the analysed structure.

Finally, a new adaptive response surface method

is introduced to analyse the design reliability with high accuracy and efficiency. Whereby the surrogate model is based on an improved moving least square approximation combined with an adaptive design of experiment. In order to obtain a fast simulation procedure on the response surface an adaptive importance sampling concept is developed. In this sense, the proposed method is very robust and efficient for $n \leq 20$ random parameters and combine the advantages of an adaptive design of experiment, adaptive sampling and efficient response surface methods.

Within the robust design optimization the design parameters can be random variables themselves and in addition the objective and the constraint functions can be random types. Using the robust design optimization we obtain optimized designs such that they are insensitive to uncertainties within a safety level of two sigma. The reliability-based optimization includes the failure probability as constraint condition or as term of the objective function themselves. So we obtain design with minimal failure probability applicable for all safety levels up to 6 sigma.

Robust design optimization can provide multiple benefits. It permits the identification of those design parameters that are critical for the achievement of a certain performance characteristic. A proper adjustment of the thus identified parameters to hit the target performance is supported. This can significantly reduce product costs.

10 ACKNOWLEDGEMENT

The author would like to express his thanks to Johannes Will, Christian Bucher, Veit Bayer, Thomas Most and Jörg Unger for their assistance of this collaborative work between the DYNARDO GmbH, the Vienna University of Technology and the Bauhaus University Weimar.

REFERENCES

- [1] A. J. Booker, Jr. Dennis, J. E., P. D. Frank, D. B. Serafini, V. Torczon, and M. Trosset. A rigorous framework for optimization of expensive functions by surrogates. *Structural Optimization*, 17(1):1 – 13, 1999.
- [2] A. Giunta and L. T. Watson. A comparison of approximation modeling technique: Polynomial versus interpolating models. In *7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis & Optimization*, pages 392 – 404. AIAA, St. Louis, MO, 1998.
- [3] T. W. Simpson, A. J. Booker, S. Ghosh, A. Giunta, Koch, P. N., and R. J. Yang. Approximation methods in multidisciplinary analysis and optimization: A panel discussion. *Structural and Multidisciplinary Optimization*, 2003.
- [4] K. K. Choi, Byeng D. Youn, and Ren-Jye Yang. Moving least square method for reliability-based design optimization. In *WCSMO-4*, Dalian, China, 2001. Center for Computer-Aided Design and Department of Mechanical Engineering.
- [5] B. D. Youn, K. K. Choi, R. J. Yang, and L. Gu. Reliability-based design optimization for crashworthiness of vehicle side impacts. *Structural and Multidisciplinary Optimization*, 26:272 – 283, 2004.
- [6] R. J. Yang and L. Gu. Experience with approximate reliability-based optimization. *Structural and Multidisciplinary Optimization*, 26:152 – 159, 2004.
- [7] M. Rais-Rohani and M. N. Singh. Comparison of global and local response surface techniques in reliability-based optimization of composite structures. *Structural and Multidisciplinary Optimization*, 26:333 – 345, 2004.
- [8] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn. Design and analysis of computer experiments. *Statistical Science*, 4(4):409 – 435, 1989.
- [9] T. W. Simpson, J. Peplinski, P. N. Koch, and J. K. Allen. Metamodels for computer-based engineering design: Survey and recommendations. *Engineering with Computers*, 17(2):129 – 150, 2001.
- [10] G. E. P. Box and N. R. Draper. *Empirical Model Building and Response Surfaces*. John Wiley and Sons, New York, USA, 1987.
- [11] R. H. Myers. *Response Surface Methodology*. Allyn and Bacon Inc., Boston USA, 1971.
- [12] R. H. Myers and D. C. Montgomery. *Response Surface Methodology - Process and Product Optimization Using Designed Experiments*. John Wiley & Sons, Inc., New York, 1995.
- [13] H. O. Madsen, Steen Krenk, and Niels C. Lind. *Methods of Structural Safety*. Prentice-Hall, Englewood Cliffs, 1986.
- [14] L. Gu, R. J. Yang, C. H. Tho, M. Makowskit, O. Faruquet, and Y. Li. Optimisation and robustness for crashworthiness of side impact. *International Journal of Vehicle Design*, 26(4):348 – 360, 2001.
- [15] L. Yu, Das P. K., and Y. Zheng. Stepwise response surface method and its application in reliability analysis of ship hull structure. *Journal of Offshore Mechanics and Arctic Engineering*, 124(4):226 – 230, 2002.
- [16] X. Liao, Q. Li, X. Yang, and W. Zhang. Multiobjective optimization for crash safety design of vehicles using stepwise regression model, 2007.
- [17] D. Shepard. A two-dimensional interpolation function for irregularly-spaced data. In *Proc.*

- ACM Natl Conf., pages 517 – 523. 1968.
- [18] C. Du. An interpolation method for grid-based terrain modelling. *THE COMPUTER JOURNAL*, 39(10):837 – 843, 1996.
 - [19] G. Matheron. Principles of geostatistics. *Economic Geology*, 58:1246 – 1266, 1963.
 - [20] J.D. Martin and T.W. Simpson. A study on the use of kriging models to approximate deterministic computer models. In *Proceedings of DETC ASME 2003 Design Engineering Technical Conferences and Computers and Information in Engineering Conference Chicago, September 2-6*. Illinois USA, 2003.
 - [21] D. R. Jones, M. Schonlau, and W. J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13:455 – 492, 1998.
 - [22] P. Lancaster and K. Salkauskas. *Curve and surface fitting; an introduction*. Academic Press, London, 1986.
 - [23] T. Most and C. G. Bucher. A moving least squares weighting function for the element-free galerkin method which almost fulfills essential boundary conditions. *Structural Engineering and Mechanics*, 21(3): 315 – 332, 2005.
 - [24] V. Vapnik and A. Chervonenkis. *Theory of Pattern Recognition*. Nauka, Moscow, 1974.
 - [25] D. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20:273–297, 1995.
 - [26] V. Vapnik. *The Nature of Statistical Learning Theory*. Springer, New York, 1995.
 - [27] B. Schoelkopf and A. J. Smola. *Learning with Kernels*. MIT Press, 2001.
 - [28] J. Platt. Fast training of support vector machines using sequential minimal optimization. In B. Schölkopf, C. Burges, and A. Smola, editors, *Advances in Kernel Methods - Support Vector Learning*. MIT Press, 1998.
 - [29] H. Drucker, C. J. C. Burges, L. Kaufman, A. Smola, and V. Vapnik. Support vector regression machines. *Advances in Neural Information Processing Systems*, 9:155–161, 1997.
 - [30] A. J. Smola and B. Schoelkopf. A tutorial on support vector regression. NeuroCOLT2 Technical Report NC2-TR-1998-030, 1998.
 - [31] A. J. Smola and B. Schoelkopf. A tutorial on support vector regression. *Statistics and Computing*, 14:199–222, 2004.
 - [32] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *Nature*, 323(6088): 533–536, 1986.
 - [33] Martin Riedmiller and Heinrich Braun. A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In *Proc. of the IEEE Intl. Conf. on Neural Networks*, pages 586–591, San Francisco, CA, 1993.
 - [34] E. M. Johansson, F. U. Dowla, and D. M. Goodman. Backpropagation learning for multilayer feed-forward neural networks using the conjugate gradient method. *International Journal of Neural Systems*, 2(4):291–301, 1992.
 - [35] Martin F. Moller. Scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, 6(4):525–533, 1993.
 - [36] Martin H. Hagan and Mohammad B. Menhaj. Training feedforward networks with the marquardt algorithm. *IEEE Transactions on Neural Networks*, 5(6):989–993, 1994.
 - [37] M. T. Hagan, H. B. Demuth, and M. Beale. *Neural Network Design*. PWS Publishing Company, 1996.
 - [38] Simon Haykin. *Neural Networks, a comprehensive foundation*, chapter 4.6, pages 181–182. Prentice Hall, second edition, 1999.
 - [39] S. Amari, N. Murata, K.-R. Müller, M. Finke, and H. Yang. Statistical theory of overtraining — is cross-validation asymptotically effective? In David S. Touretzky, Michael C. Mozer, and Michael E. Hasselmo, editors, *Advances in Neural Information Processing Systems*, volume 8, pages 176–182. The MIT Press, 1996.
 - [40] M. D. McKay, W. J. Conover, and R. J. Beckmann. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979.
 - [41] D. E. Huntington and C. S. Lyrintzis. Improvements to and limitations of latin hypercube sampling. *Probabilistic Engineering Mechanics*, 13(4):245 – 253, 1998.
 - [42] R. L. Iman and W. J. Conover. *A Distribution Free Approach to Inducing Rank Correlation Among Input Variables*. Communications in Statistics, vol. b11 edition, 1982.
 - [43] R. Rackwitz. *Response Surfaces in Structural Reliability*. Berichte zur Zuverlässigkeitstheorie der Bauwerke, Heft 67/1982, Laboratorium für den konstruktiven Ingenieurbau, Technische Universität München, München Germany, 1982.
 - [44] L. Faravelli. *Response Surface Approach for Reliability Analysis*. Pubblicazione n. 160, Dipartimento di Meccanica Strutturale Dell’ Università di Pavia, Pavia Italy, 1986.
 - [45] C. G. Bucher and U. Bourgund. *Efficient Use of Response Surface Methods*. Bericht Nr. 9 -

- 87, Institut für Mechanik, Universität Innsbruck, Innsbruck, Austria, 1987.
- [46] C. G. Bucher and U. Bourgund. A fast and efficient response surface approach for structural reliability problems. *Structural Safety*, 7(1):57–66, 1990.
- [47] S. Engelund and R. Rackwitz. Experiences with experimental design schemes for failure surface estimation and reliability. In Y. K. Lin, editor, *ASCE Specialty Conference on Probabilistic Mechanics and Structural and Geotechnical Reliability*, pages 244 – 247. Proceedings, 6th ASCE, New York, USA, 1992.
- [48] M. R. Rajashekhar and B. R. Ellingwood. A new look at the response surface approach for reliability analysis. *Structural Safety*, vol. 12, no. 1 edition, 1993.
- [49] W. Ouypornprasert and C. G. Bucher. An efficient scheme to determine response functions for reliability analyses. *Universität Innsbruck, Internal Working Report No. 30*, 30:1 – 39, 1988.
- [50] C. G. Bucher, Y. M. Chen, and G. I. Schuëller. Time variant reliability analysis utilizing response surface approach. In P. Thoft-Christensen, editor, *Proc., 2nd IFIP Working Conference on Reliability and Optimization of Structural Systems*, pages 1 – 14. Springer Verlag, Berlin, Germany, 1988.
- [51] M. J. Box and N. R. Draper. Factorial designs, the $|x^t x|$ criterion, and some related matters. *Technometrics*, Vol. 13(4):731 – 742, 1971.
- [52] O. Klingmüller and U. Bourgund. *Sicherheit und Risiko im Konstruktiven Ingenieurbau*. Vieweg, Braunschweig, Wiesbaden, Germany, 1992.
- [53] A. Der Kiureghian and J.-B. Ke. The stochastic finite element method in structural reliability. *Probabilistic Engineering Mechanics*, 3(2):135–141, 1988.
- [54] T. Hisada and S. Nakagiri. Stochastic finite element method developed for structural safety and reliability. In *Third Int. Conf. on Structural Safety and Reliability*, pages 409–417, Trondheim, 1981.
- [55] A. Nataf. *Détermination des distributions de probabilités dont les marges sont données*, volume 225 of *Comptes Rendus de l'Académie des Sciences*. 1962.
- [56] P.-L. Liu and A. Der Kiureghian. Multivariate distribution models with prescribed marginals and covariances. *Probabilistic Engineering Mechanics*, 1(2):105–112, 1986.
- [57] R. Ghanem and P.D. Spanos. *Stochastic Finite Elements - a Spectral Approach*. Springer, New York, Berlin, 1991.
- [58] C. E. Brenner. *Ein Beitrag zur Zuverlässigkeitsanalyse von Strukturen unter Berücksichtigung von Systemuntersuchungen mit Hilfe der Methode der Stochastischen Finite Elemente*. Thesis (Dr. tech.). Leopold Franzens Universität Innsbruck, Innsbruck, 1995.
- [59] Y. Schorling. *Beitrag zur Stabilitätsuntersuchung von Strukturen mit räumlich korrelierten geometrischen Imperfektionen*. Institut für Strukturmechanik, Bericht 2/98. Bauhaus - Universität Weimar, Weimar, 1998.
- [60] P. N. Koch, R.-J. Yang, and L. Gu. Design for six sigma through robust optimization. *Structural and Multidisciplinary Optimization*, 26:235 – 248, 2004.
- [61] K.-H. Hwang, K.-W. Lee, and G.-J. Park. Robust optimization of an automobile rearview mirror for vibration reduction. *Structural and Multidisciplinary Optimization*, 21:300–308, 2001.
- [62] Aharon Ben-Tal and Arkadi Nemirovski. Robust optimization – methodology and applications. *Mathematical Programming (Series B)*, 92:453 – 480, 2002.
- [63] Ioannis Doltsinis and Zhan Kang. Robust design of structures using optimization methods. *Comput. Methods Appl. Mech. Engrg.*, 193:2221 – 2237, 2004.
- [64] D.M. Byrne and S. Taguchi. The taguchi approach to parameter design. In *40th Annual Quality Congress Transactions*, pages 19 – 26. American Society for Quality Control, Milwaukee, Wisconsin, 1987.
- [65] M.S. Phadke. *Quality Engineering using Robust Design*. Prentice Hall, Englewood Cliffs, New Jersey, 1989.
- [66] S.J. Abspoel, L.F.P. Etman, J. Vervoort, R.A. van Rooij, A.J.G. Schoofs, and J.E. Rooda. Simulation based optimization of stochastic systems with integer design variables by sequential multipoint linear approximation. *Structural and Multidisciplinary Optimization*, 22:125–138, 1996.
- [67] Olivier V. Pictet, Michel M. Dacorogna, Rakhal D. Davé, Batien Chopard, Roberto Schirru, and Marco Tomassini. Genetic algorithms with collective sharing for robust optimization in financial applications. OVP.1995-02-06, http://www.olsen.ch/research/307_ga_pase95.pdf, 1996.
- [68] Juergen Branke. Creating robust solutions by means of an evolutionary algorithm. *A.E. Eiben, T. Bäck, M. Schoenhauer, and H.-P. Schwefel, editors, Parallel Problem Solving from Nature*, 1498:119–128, 1998.
- [69] R. E. Melchers. Optimality-criteria-based

- probabilistic structural design. *Structural and Multidisciplinary Optimization*, 23:34 – 39, 2001.
- [70] K. K. Choi, J. Tu, and Y. H. Park. Extensions of design potential concept for reliability-based design optimization to nonsmooth and extreme cases. *Structural and Multidisciplinary Optimization*, 22:335–350, 2001.
- [71] Matthew Allen, Michael Raulli, Kurt Maute, and Dan M. Frangopol. Reliability-based analysis and design optimization of electrostatically actuated mems. *Computers and Structures*, 82(13–14):1007–1020, May 2004.
- [72] M. Allen and K. Maute. Reliability-based design optimization of aeroelastic structures. *Structural and Multidisciplinary Optimization*, 27:228 – 242, 2004.
- [73] K. Schittkowski. Nlqpl: A fortran subroutine for solving constrained nonlinear programming problems. *Ann Oper Res*, 5:485 – 500, 1985.
- [74] G. Kharmanda, A. Mohamed, and M. Lemaire. Efficient reliability-based design optimization using a hybrid space with application to finite element analysis. *Structural and Multidisciplinary Optimization*, 24:233 – 245, 2002.
- [75] Byung D. Youn and Kyung K. Choi. A new response surface methodology for reliability-based design optimization. *Computers and Structures*, 82:241 – 256, 2004.
- [76] J. O. Royset and E. Polak. Reliability-based optimal design using sample average approximations. *Probabilistic Engineering Mechanics*, 19:331 – 343, 2004.
- [77] J. O. Royset, A. Der Kiureghian, and E. Polak. Successive approximations for the solution of optimal design problems with probabilistic objective and constraints. In Der Kiureghian, Madanat, and Pestana, editors, *Applications of Statistics and Probability in Civil Engineering*, pages 1049 – 1056. Millpress, Rotterdam, Ninth International Conference on Applications of Statistics and Probability in Civil Engineering (ICASP9), July 6-9 2003.
- [78] Wei Chen, Jin Ruichen, and Agus Sudjianto. Analytical variance-based global sensitivity analysis in simulation-based design under uncertainty. In *Proceedings of DETC'04, ASME 2004 Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Salt Lake City, Utah USA, September 28 – October 2 2004.
- [79] Benjamin Wilson, David Cappelleri, Timothy W. Simpson, and Mary Frecker. Efficient pareto frontier exploration using surrogate approximations. *Optimization and Engineering*, 2:31 – 50, 2001.
- [80] L.F.P. Etman, J.M.T.A. Adriaens, M.T.P. van Slagmaat, and A.J.G. Schoofs. Crashworthiness design optimization using multipoint sequential linear programming. *Structural Optimization*, 12:222–228, 1996.
- [81] V. V. Toropov and L.F. Alvarez. Development of mars – multipoint approximation method based on the response surface fitting. Technical report, AIAA, 1998.
- [82] N. Stander and K.J. Craig. On the robustness of a simple domain reduction scheme for simulation-based optimization. *Eng. Comput.*, 19(4):431–50, 2002.
- [83] H. Kurtaran, A. Eskandarian, D. Marzougui, and N.E. Bedewi. Crashworthiness design optimization using successive response surface approximations. *Computational Mechanics*, 29:409–421, 2002.
- [84] A. Teughels. *Inverse modelling of civil engineering structures based on operational data*. PhD thesis, Katholieke Universiteit te Leuven, 2003.
- [85] P. Bjerager. Probability integration by directional simulation. *Journal of Engineering Mechanics, ASCE*, Vol. 114(No. 8):1285 – 1302, 1988.
- [86] O. Ditlevsen and P. Bjerager. Plastic reliability analysis by directional simulation. *Journal of Engineering Mechanics, ASCE*, Vol. 115(No. 6):1347 – 1362, 1989.
- [87] R. E. Melchers. *Radial importance sampling for structural reliability*. Engrg. Mech., October, asce, 116(1) edition, 1990.
- [88] C. G. Bucher. *Adaptive Sampling - An Iterative Fast Monte Carlo Procedure*. Structural Safety, vol. 5, no. 2 edition, 1988.
- [89] C. Bucher, Y. Schorling, and W. Wall. Slang - the structural language, a tool for computational stochastic structural analysis. In S. Sture, editor, *Proc. 10th ASCE Eng. Mech. Conf., Boulder, CO, May 21–24, 1995*, pages 1123–1126. ASCE, New York, 1995.
- [90] V. Bayer. *Zur Zuverlässigkeitsbeurteilung von Baukonstruktionen unter dynamischen Einwirkungen*. Institut für Strukturmechanik, Bericht 3/99. Bauhaus - Universität Weimar, Weimar, 1999.
- [91] M. D. McKay, W. J. Conover, and R. J. Beckmann. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21:239–245, 1979.
- [92] A. Florian. An efficient sampling scheme: Updated latin hypercube sampling. *Probabilistic Engineering Mechanics*, 7:

123–130, 1992.

- [93] D. E. Huntington and C. Lyrantzis.
Improvements to and limitations of latin
hypercube sampling. *Probabilistic
Engineering Mechanics*, 13(4):245–253,
1998.

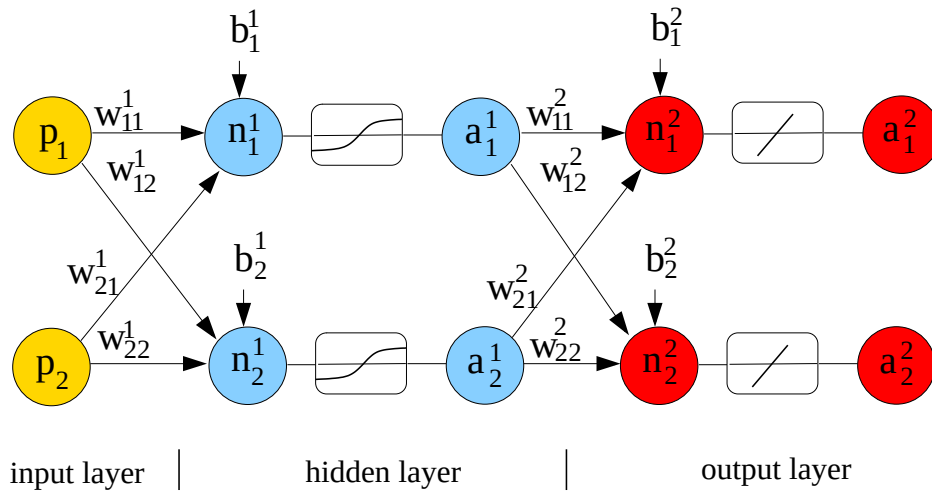


Figure 50: general layout of a multilayer perceptron

hard limit		$y = -1 \quad x < 0$ $y = +1 \quad x \geq 0$
saturating linear		$y = -1 \quad x < -1$ $y = x \quad -1 \leq x < 1$ $y = +1 \quad x \geq 1$
positive linear		$y = 0 \quad x < 0$ $y = x \quad x \geq 0$
linear		$y = x$
hyperbolic tangent sigmoid		$y = \frac{e^x - e^{-x}}{e^x + e^{-x}}$
log-sigmoid		$y = \frac{1}{1 + e^{-x}}$

Figure 51: activation functions commonly used in multilayer perceptrons

No. of random variables	396	14	14
Selection criterion	none (all)	highest $\sigma_{X_i}^2$ ("first 14")	highest R^2 ("top 14")
Prob. of failure P_f	$9.7 \cdot 10^{-3}$	$2.8 \cdot 10^{-4}$	$3.6 \cdot 10^{-3}$
Statistical error σ_{P_f} $\text{cov}(P_f)$	$5.2 \cdot 10^{-4}$ 5 %	$8.8 \cdot 10^{-5}$ 32 %	$3.2 \cdot 10^{-4}$ 9 %

Table 1: Probabilities of failure for different sets of random variables.

	Number of samples			
	20	50	100	200
Linear correlation errors				
10 random variables				
MCS	0.54548	0.35110	0.24643	0.17587
LHS	0.06885	0.02176	0.00978	0.00457
LHS-MLCE	0.03192	0.00349	0.00088	0.00025
50 random variables				
MCS	0.71477	0.47671	0.34499	0.24678
LHS	0.71788	0.62360	0.01159	0.00466
LHS-MLCE	0.56979	0.27324	0.00694	0.00217
Quadratic correlation errors				
10 random variables				
MCS	0.50023	0.31610	0.22690	0.15966
LHS	0.57650	0.37468	0.26910	0.19048
LHS-MLCE	0.58594	0.37922	0.27023	0.19145
50 random variables				
MCS	0.67240	0.44819	0.32042	0.22941
LHS	0.67659	0.46101	0.35811	0.25216
LHS-MLCE	0.67232	0.46030	0.36005	0.25330

Table 2: Maximum correlation errors for a given sampling size by using Monte Carlo Simulation (MCS), Latin Hypercube sampling (LHS) and Latin Hypercube sampling with minimized linear correlation errors (LHS-MLCE)

Number of variables	Number of samples				
	50	100	200	500	1000
10	0.09 s	0.14 s	0.33 s	1.08 s	4.23 s
50	9.6 s	37 s	3 min	17 min	63 min
100	1 min	6 min	15 min	100 min	382 min

Table 3: Computational time on a workstation with AMD Opteron 2.4 GHz for the Latin Hypercube sampling with minimized linear correlation errors

	Number of state function evaluations	Failure probability	Variance of the estimation $\sigma_{\hat{P}(\mathcal{F})}^2$	Accuracy error %
Direct adaptive importance sampling (AS)	4000	$4.8907 \cdot 10^{-6}$	$1.5813 \cdot 10^{-7}$	0
Global polynomial 2nd order approximation (RSM)	9	$2.1563 \cdot 10^{-10}$	$2.5303 \cdot 10^{-11}$	2268000
Adaptive response surface approximation (ARSM)	9	$7.9728 \cdot 10^{-11}$	$6.0763 \cdot 10^{-12}$	6134100
	18	$4.47 \cdot 10^{-6}$	$1.1713 \cdot 10^{-7}$	9
	27	$4.4096 \cdot 10^{-6}$	$1.214 \cdot 10^{-7}$	11
	36	$4.6044 \cdot 10^{-6}$	$1.1337 \cdot 10^{-7}$	6

Table 4: Results of the failure probability of the Quattro Function depending on the applied analysis method.

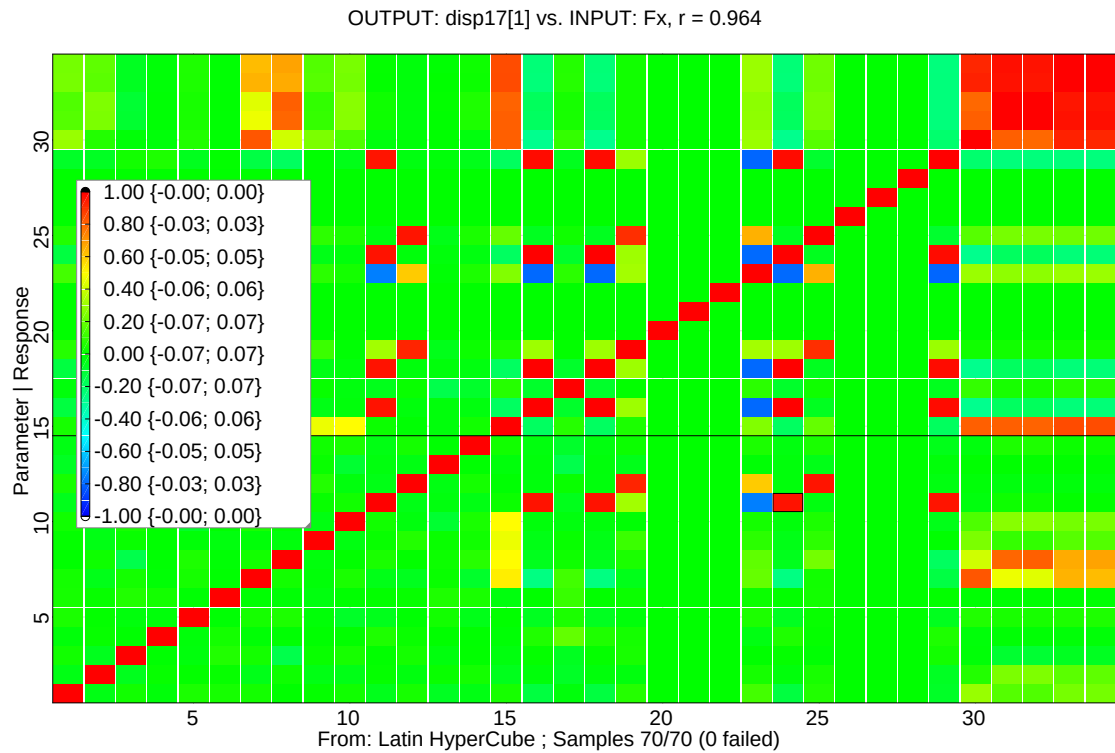


Figure 52: Matrix C_{XX} of the linear correlation coefficients with the possible lower and upper bounds according to a specified confidence interval I_p of 95%.

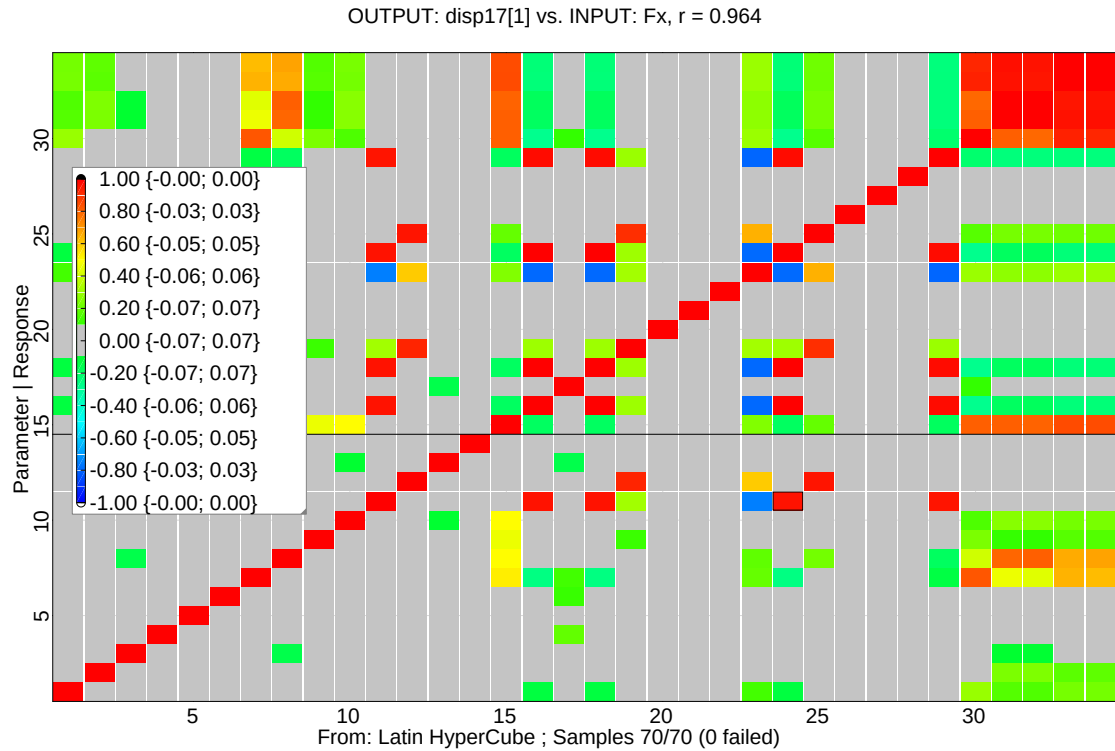


Figure 53: Matrix C_{XX} of the most significance linear correlation coefficients used for reduced polynomial regression models.

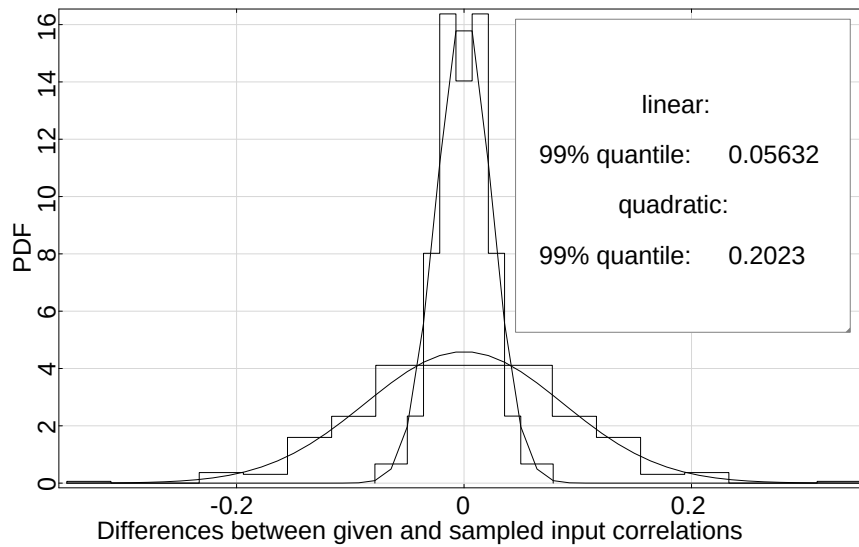


Figure 54: Histogram of the differences between the user-defined and sampled quadratic and linear input correlation matrix C_{XX} . Fitted probability density function with the 99%-fractiles of the coefficients of correlation.

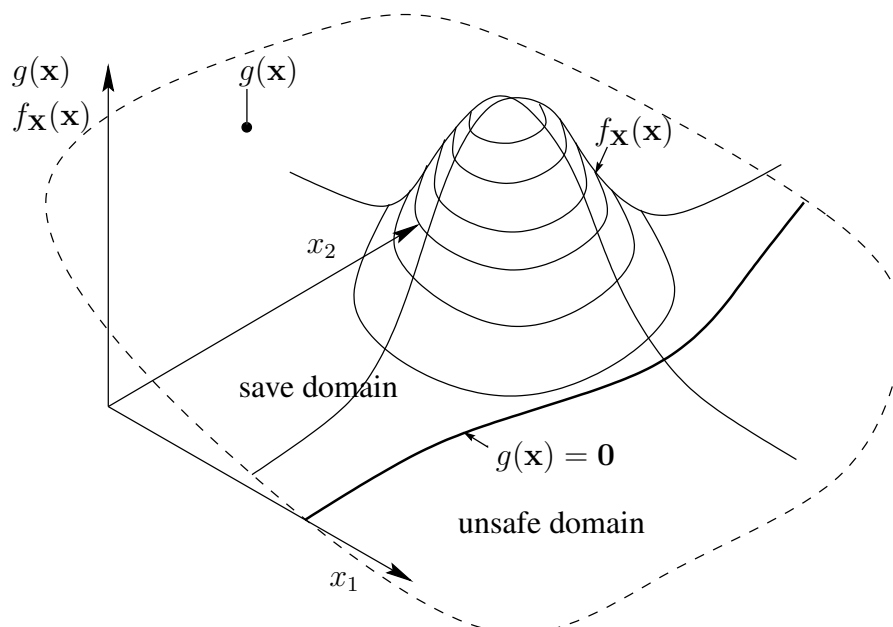


Figure 55: The state function $g(\mathbf{x})$ of a numerical model is given implicitly, e.g. is result of a finite element analysis depending on several design responses. The failure condition leads to a unknown deterministic limit state function $g(\mathbf{x}) = 0$, where $f_{\mathbf{X}}(\mathbf{x})$ is the joint probability density function.

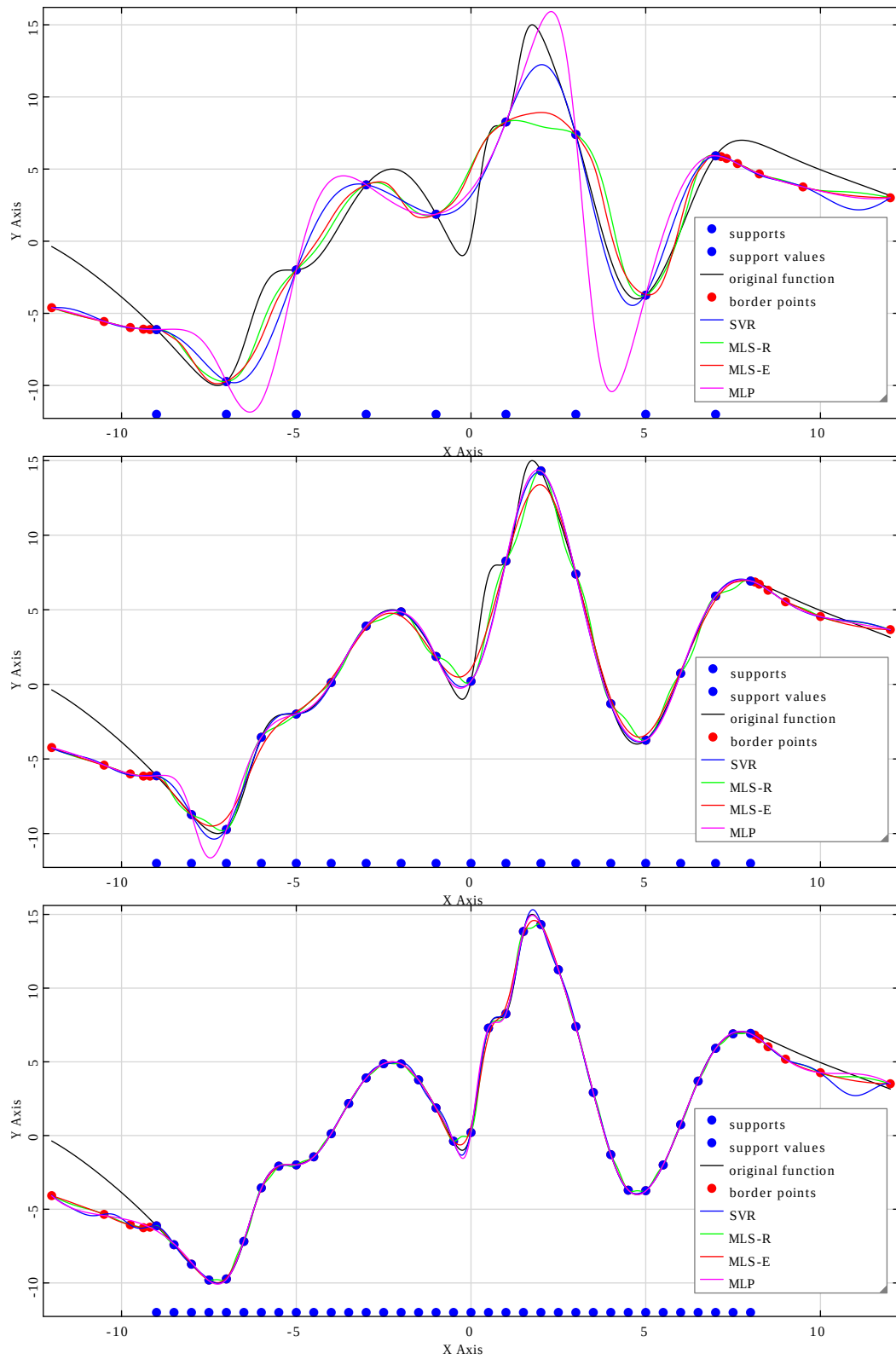


Figure 56: Deterministic 1D test function with $N = 9, 18, 35$ support points.

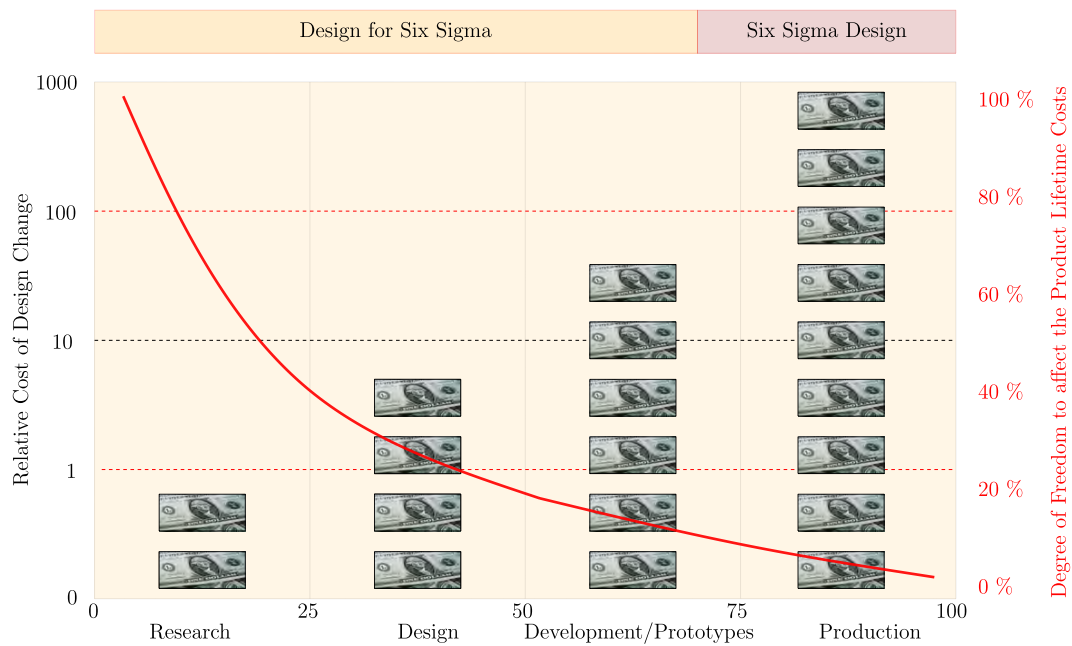


Figure 57: Product Development Phases: within the "Design for Six Sigma" the degree of freedom to affect the product lifetime cost is very high and the cost of design change is proportional moderate in contrast to "Six Sigma Design" concept to optimize the manufacturing processes only.

Sigma level	Percent variation	Probability of failure $P(\mathcal{F})$	Defects per million (short term)	Defects per million (long term)
$\pm 1\sigma$	68.26	$3.17 \cdot 10^{-1}$	317400	697700
$\pm 2\sigma$	95.46	$4.54 \cdot 10^{-2}$	45400	308733
$\pm 3\sigma$	99.73	$2.7 \cdot 10^{-3}$	2700	66803
$\pm 4\sigma$	99.9937	$6.3 \cdot 10^{-5}$	63	6200
$\pm 5\sigma$	99.999943	$5.7 \cdot 10^{-7}$	0.57	233
$\pm 6\sigma$	99.9999998	$2.0 \cdot 10^{-9}$	0.002	3.4

Table 5: Sigma level depending on the variation of the normal distribution, defects per million and associated probability of failure $P(\mathcal{F})$. A probability of 3.4 out of 1 million is achieved when the performance target is 4.5σ away from the mean value (short term). The additional 1.5σ (long term) leading to a total of 6 standard deviations are used as a safety margin to allow for "drift of the mean value" in the properties and environment which the product can see over its lifetime.