Efficiency Improvement of Stochastic Simulation by Means of Subset Sampling

Martin Liebscher, Stephan Pannier, Jan-Uwe Sickert, Wolfgang Graf
TU Dresden, Dresden, Germany

Abstract:

The design of engineering systems requires a sophisticated structural analysis close to reality. The uncertainty of structural parameters, such as loads, material parameters, and geometrical properties must be taken into account. Based on the framework of probabilistics a structural behavior may be assessed by a failure probability. The failure probability can be computed with the aid of Monte Carlo simulation. Especially in the case of complex nonlinear structures the Monte Carlo simulation meets their limits. The more advanced simulation method subset sampling, which promises to compensate this drawback, is investigated in this paper. The main idea of subset sampling is the subdivision of the failure event into a sequence of partial failure events, which are denoted as subsets. The numerical efficient sampling within the subset is realized with aid of the Markov chain Monte Carlo simulation. Subset sampling is demonstrated by means of examples.

Keywords:

subset sampling, Monte Carlo simulation, reliability analysis

1 Introduction

The design of engineering systems requires a structural analysis close to reality. In general, this is realized with the aid of a numerical computational model. For analyzing the nonlinear behavior of complex engineering structures the Finite Element method provides a sophisticated basis. Further, to obtain useful results the structural parameters, such as loads, material parameters, and geometrical properties must be specified realistically. In engineering practice, the structural parameters are frequently affected by uncertainty. This uncertainty is reflected in the structural responses and the reliability. As a consequence requirements regarding a desired structural behavior can no longer be formulated in absolute terms, but must be defined with respect to this uncertainty.

Based on the framework of probabilistics for modeling and processing uncertainty, design constraints are formulated as limits for the failure probability. The failure probability can be computed with the aid of analytical and numerical methods. For the probabilistic analysis of complex structures, analytical methods generally meet their limits, and a numerical solution is associated with a considerable computational cost. A useful basis for the solution of general problems is the Monte Carlo simulation and further developments there of. In this approach random realizations of the structural parameters are generated according to their probability distribution. The associated results are collected in a sample, which reflects the statistical properties of the structural response and enables the estimation of the failure probability. In industrial applications it can be necessary to estimate small failure probabilities. In those cases the direct Monte Carlo simulation may require a high computational effort. This applies, in particular, in association with an underlying nonlinear Finite Element analysis of large systems, which repeatedly performed with in the Monte Carlo framework. To compensate this drawback, advanced simulation methods are developed, e.g. importance sampling, line sampling, or subset sampling.

A short review of the Monte Carlo simulation given in Sec. 2 is followed by a detailed explanation of the newly introduced method subset sampling in Sec. 3. Thereby, the basic idea, algorithms, and some remarks with respect to numerical efficiency are provided. Sec. 4 includes a benchmark study and two further examples.

2 Direct Monte Carlo simulation

The direct Monte Carlo simulation is a computational algorithm, which is capable to assess the reliability of structures in terms of the failure probability. It quotes a numerical robust approximate solution of integrals which are unseizable with analytical methods. This holds for the determination of the failure probability

$$P_F = P(\underline{x} \mid g(\underline{x}) \leq 0) = \int_{\underline{x}|g(\underline{x})\leq 0} f(\underline{x}) d\underline{x} \tag{1}$$

if either the joint density function $f(\underline{x})$ is non-Gaussian or the structural performance function $g(\underline{x})$ is nonlinear. Structural analysis requires the application of an appropriate numerical computational model. In order to obtain results close to reality, nonlinear algorithms have to be applied. In this case the limit state function $g(\underline{x})$, which is the boundary between survival and failure cannot be specified in a closed form, but rather pointwise as a result of individual deterministic numerical simulations. The failure probability can be computed by means of an indicator function

$$I_F(\underline{x}) = \begin{cases} 1 & \text{if } \underline{x} \in F \\ 0 & \text{if } \underline{x} \notin F \end{cases} \qquad \text{with } F = \{\underline{x} \mid g(\underline{x}) \leq 0\} \quad . \tag{2}$$

where F denotes the failure region. The Monte Carlo simulation uses the interpretation of the integral

$$P_F = \int_{\underline{x}} I_F(\underline{x}) f(\underline{x}) d\underline{x} = E(I_F(\underline{x})) \tag{3}$$

in order to estimate the failure probability P_F. The application of the strong Law of Large Numbers yields

$$E(I_F(\underline{x})) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} I_F(\underline{x}_k) \quad . \tag{4}$$

The evaluation of Eq. (4) demands the existence of a sample $\{x_k\}_{k=1}^N$ of N elements distributed with the probability density function $f(\underline{x})$. An estimator $\hat{P}_F$ for the failure probability is derived from Eq. (4)

$$\hat{P}_F = \hat{E}(I_F(\underline{x})) = \frac{1}{N}\sum_{k=1}^{N} I_F(\underline{x}_k) \quad . \tag{5}$$

$\hat{P}_F$ converges to the failure probability P_F as the number of samples N approaches to infinity [2, 8, 10]. The expected value E and the variance Var of the estimated failure probability arise to

$$E(\hat{P}_F) = P_F \quad , \tag{6}$$

$$Var(\hat{P}_F) = \frac{(1-P_F)P_F}{N} \quad . \tag{7}$$

The coefficient of variation is determined by

$$\delta_{P_F} = \frac{\sqrt{Var\left[\hat{P}_F\right]}}{E\left[\hat{P}_F\right]} = \sqrt{\frac{1-P_F}{P_F \cdot N}} \quad . \tag{8}$$

The coefficient of variation does not depend on the dimensionality of the random vector $\underline{x}$. The determination of a small failure probability P_F with an acceptable level of accuracy demands a large number of samples N. An overview about the required number of samples for different values of the failure probability P_F and the coefficient of variation δ_{P_F} is given exemplarily in Tab. 1.

Table 1: Estimated number of samples for Monte Carlo simulation

P_F	10^{-1}	10^{-2}	10^{-3}	10^{-4}
$\delta_{P_F} = 0.3$	$1 \cdot 10^2$	$1.1 \cdot 10^3$	$1.11 \cdot 10^4$	$1.111 \cdot 10^5$
$\delta_{P_F} = 0.1$	$9 \cdot 10^2$	$1 \cdot 10^4$	$1 \cdot 10^5$	$1 \cdot 10^6$
$\delta_{P_F} = 0.01$	$9 \cdot 10^4$	$1 \cdot 10^6$	$1 \cdot 10^7$	$1 \cdot 10^8$

Tab. 1 points out, that the computational effort increases enormously for small values of failure probability P_F. In the field of stochastic analysis various methods, e.g. importance sampling [2, 8, 9, 10], line sampling [3, 6], and subset sampling [1, 6] have been developed for decreasing the computational effort in a significant manner. In this paper the subset sampling is examined and assessed in comparison to direct Monte Carlo simulation.

3 Subset sampling

3.1 Basic concept

The main principle of the presented method was first mentioned in literature as umbrella sampling [11]. Nowadays this method is referred to as subset sampling. The basic idea of subset sampling is the subdivision of the failure event into a sequence of m partial failure events (subsets) $F_1 \supset F_2 \supset \cdots \supset F_m = F$. Generally, the determination of small failure probabilities P_F with the aid of Monte Carlo simulation requires the expensive simulation of rare events. The division into subsets (subproblems) offers the possibility to transfer the simulation of rare events into a set of simulations of more frequent events. The determination of the failure regions F_i can be effected by presetting a series $\{g_i\}_{i=1}^m$ of limit values

$$F_i = \{x : g(\underline{x}) \leq g_i\} \quad . \tag{9}$$

This enables the computation of the failure probability as product of conditional probabilities $P(F_{i+1}|F_i)$ and $P(F_1)$.

$$P_F = P(F_m) = P(F_m|F_{m-1})P(F_{m-1}|F_{m-2}) \cdots P(F_2|F_1)P(F_1) \tag{10}$$

$$= P(F_1)\prod_{i=1}^{m-1} P(F_{i+1}|F_i) \tag{11}$$

Whereas m denotes the (unknown) total number of subsets. The determination of the failure regions F_i and subsequently the partial failure probabilities $P_i = P(F_{i+1} \mid F_i)$ influences the accuracy of the simulation. It is convenient to specify the limit values $\{g_i\}_{i=1}^{m}$ so that nearly equal partial failure probabilities $P_i \mid i = 2 \ldots m$ are obtained for each subset. A proposed value is $P_i = 0.1$ [1]. Unfortunately, it is difficult to specify the limit values g_i in advance according to the prescribed probability P_i. Therefore the limit values have to be determined adaptively within the simulation.

3.2 Subset sampling algorithm

In the first step the probability $P_1 = P(F_1)$ is determined by application of the direct Monte Carlo simulation

$$P(F_1) \approx \hat{P}_1 = \frac{1}{N_1} \sum_{k=1}^{N_1} I_{F_1}(\underline{x}_k^{(1)}) \quad . \tag{12}$$

To obtain conditional probabilities $P(F_{i+1}|F_i)$ the evaluation of the respective probability functions

$$f(\underline{x} \mid F_i) = \frac{I_{F_i}(\underline{x}) f(\underline{x})}{P(F_i)} \tag{13}$$

is necessary. With the application of the Markov chain Monte Carlo simulation in conjunction with the Metropolis-Hastings algorithm samples may be generated in a numerically efficient way according to $f(\underline{x} \mid F_i)$. The starting samples of the subset $i+1$ are taken out of the $\underline{x}_k^{(i)}$ $(i = 1, \ldots, m-1)$ lying in F_i. Starting from these samples, Markov chain samples can be simulated using e.g. the modified Metropolis-Hastings method.

$$P(F_{i+1}|F_i) \approx \hat{P}_{i+1} = \frac{1}{N_{i+1}} \sum_{k=1}^{N_{i+1}} I_{F_{i+1}}(\underline{x}_k^{(i+1)}) \tag{14}$$

The starting sample of the subset $i+1$ is selected randomly from the samples $\underline{x}^{(i)} \mid g_i(\underline{x}^{(i)}) \leq g_i, i = 1, \ldots, m-1$ of subset i that are located in the failure region F_i. It is valid to constitute the limit value g_i with respect to the condition $P_i \approx P_{i-1} \, \forall \, i > 1$. As already mentioned, this cannot be realized in advance. Thus, the limit value g_i of the i-th partial subset is determined adaptively during the simulation. Therefore, g_i is determined from a list of ascending sorted tuple $\{(\underline{x}_k^{(i)}, g(\underline{x}_k^{(i)}))\}_{k=1}^{N_i}$ according to the values $g(\underline{x}_k^{(i)})$. The limit value g_i is given by

$$g_i = g\left(x_j^{(i)}\right) \mid j = int(P_i \cdot N_i) \quad . \tag{15}$$

Under the condition $g_i \leq 0$ the last subset m of the simulation has been reached. The last failure probability $P(F_m \mid F_{m-1})$ can be estimated with

$$P(F_m|F_{m-1}) \approx \hat{P}_m = \frac{1}{N_m} \sum_{k=1}^{N_m} I_{F_m}(\underline{x}_k^{(m)}) \tag{16}$$

The failure probability P_F may now be computed as

$$P_F = P_1 \cdot \prod_{i=2}^{m} P_i \quad . \tag{17}$$

The subset sampling algorithm is exemplified for three subsets in Fig. 1.

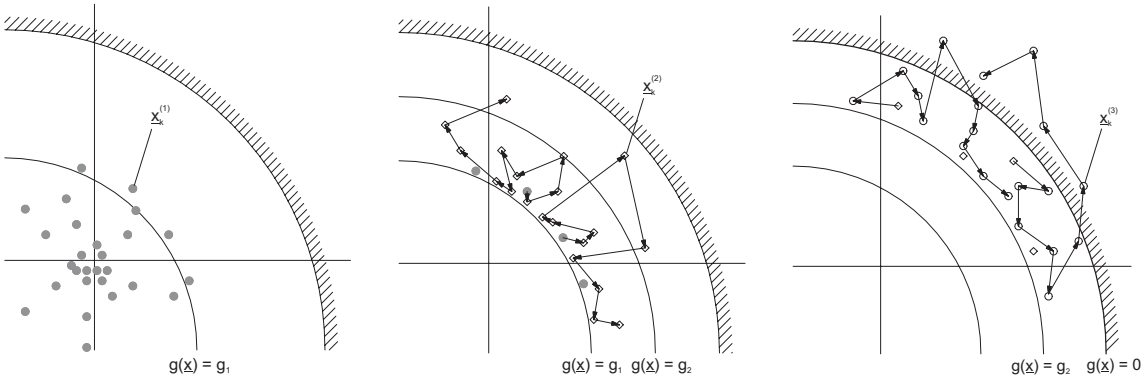


Figure 1: Subset sampling in three states

3.3 Markov chain Monte Carlo simulation

The sampling within the subsets is realized with the aid of Markov chains and the associated Markov chain Monte Carlo simulation. In the following sections the Markov chains are briefly introduced and the applied sampling algorithms are described.

3.3.1 Markov chains

A Markov chain is a sequence of random vectors $\underline{X}_0, \underline{X}_1, \ldots, \underline{X}_k$ in the state space I with the Markov property

$$P(\underline{X}_{k+1} = i_{k+1} | \underline{X}_k = i_k, \underline{X}_{k-1} = i_{k-1}, \ldots, \underline{X}_0 = i_0) = P(\underline{X}_{k+1} = i_{k+1} | \underline{X}_k = i_k) \tag{18}$$

while $i_0, \ldots, i_{k+1} \in I$. The Markov property signifies that the probability to reach an arbitrary state i_{k+1} depends only on the state i_k and is independent from all former states. The Markov chain is called homogeneous, if the probability for transition

$$P(X_{k+1} = i_{k+1} | X_k = i_k) \tag{19}$$

is independent from the state k [4].

For an arbitrary given probability density function $f(\underline{x})$ an (ergodic) Markov chain can be constructed which generates realizations approximately distributed as $f(\underline{x})$. This approach is called Markov chain Monte Carlo simulation. The expected value E for a function $h(x)$ is in the case of ergodicity of a Markov chain

$$E(h) = \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} h(x_k) \quad . \tag{20}$$

Thus the estimator $\hat{P}_F$ for the failure probability P_F can be expressed by

$$\hat{P}_F = \hat{E}(I_F(\underline{x})) = \frac{1}{N} \sum_{k=1}^{N} I_F(\underline{x}_k) \tag{21}$$

while $\{\underline{x}_k\}_{k=1}^{N}$ are realizations of the Markov chain. Eq. (21) is formal identical with Eq. (6) unlike the samples are generated from a Markov chain. They can still be used for statistical averaging as if they were independent by virtue of the Laws of Large Numbers [1]. The coefficient of variation δ_{P_F} is given by

$$\delta_{P_F} = \sqrt{\frac{1 - P_F}{P_F N} (1 + \gamma)} \tag{22}$$

In comparison to Eq. (6) the Eq. (22) includes an additional term $(1+\gamma)$ that considers the correlation of the Markov chain samples . The term $N/(1+\gamma)$ may be interpreted as the effective number of samples. Despite of the correlation of the Markov chain samples the estimator $\hat{P}_F$ is asymptotically unbiased so that the convergence is assured [1].

The implementation of Markov chain Monte Carlo simulation may be realized by various algorithms, e.g., Gibbs sampler, Metropolis-Hastings algorithm, and Barker algorithm [7]. In the following the Metropolis-Hastings algorithm is applied.

3.3.2 Metropolis-Hastings algorithm

The Metropolis-Hastings algorithm generates realizations of a Markov chain according to a given probability density function $f(\underline{x})$. Beside $f(\underline{x})$ an arbitrary n-dimensional proposal density function $q(\underline{y}|\underline{x})$ is introduced. The procedure for the transit from the state k to the state $k+1$ is given in algorithm 1.

Algorithm 1 Metropolis-Hastings algorithm [1]

Given is $\underline{x}_k$:

1. Generate a candidate $\underline{x}^*_{k+1}$:
   a) Simulate $\underline{y}$ according to $q(\underline{y} \mid \underline{x}_k)$
   b) Calculate the acceptance ratio:
   $$\rho(\underline{x}_k, \underline{y}) = \min\left\{1, \frac{f(\underline{y})}{f(\underline{x}_k)} \frac{q(\underline{x}_k \mid \underline{y})}{q(\underline{y}|\underline{x}_k)}\right\}$$
   c) Define $\underline{x}^*_{k+1}$ using the acceptance ratio:
   $$\underline{x}^*_{k+1} = \begin{cases} \underline{y} & \text{with probability} \quad \rho(\underline{x}_k, \underline{y}) \\ \underline{x}_k & \text{with probability} \quad 1-\rho(\underline{x}_k, \underline{y}) \end{cases}$$
2. Define $\underline{x}_{k+1}$ according to F:
   $$\underline{x}_{k+1} = \begin{cases} \underline{x}^*_{k+1} & \text{when} \quad \underline{x}^*_{k+1} \in F \\ \underline{x}_k & \text{when} \quad \underline{x}^*_{k+1} \notin F \\ \underline{x}_k & \text{when} \quad \underline{x}^*_{k+1} = \underline{x}_k \end{cases}$$

The vector $\underline{y}$ is always accepted if the acceptance ratio $\rho(\underline{x}_k, \underline{y})$ is 1. In any other case $\underline{y}$ is accepted as next candidate $\underline{x}^*_{k+1} = \underline{y}$ with a probability $\rho(\underline{x}_k, \underline{y})$. For a symmetric proposal density function, where $q(\underline{x}|\underline{y}) = q(\underline{y}|\underline{x})$, the acceptance ratio $\rho(\underline{x}_k, \underline{y})$ is determined by the ratio of $\frac{f(\underline{y})}{f(\underline{x}_k)}$ only .

For high dimensional probability density functions $f(\underline{x})$ the "zero-acceptance" phenomenon occurs [1]. It is characterized by generally small values for the acceptance ratio ($\rho(\underline{x}_k, \underline{y}) \simeq 0$). This yields to almost no acceptance of $\underline{y}$ as next candidate $\underline{x}^*_{k+1}$. This drawback can be dealt with the modified Metropolis-Hastings algorithm as introduced in Sec. 3.3.3.

3.3.3 Modified Metropolis-Hastings algorithm

The intention of the modified Metropolis-Hastings algorithm is to suppress the described "zeroacceptance" phenomenon. This is achieved by generating groups of components $j = 1, \ldots, n_G$ of the candidate $\underline{x}^*_{k+1} = \left(\underline{x}^{*(1)}_{k+1}, \ldots, \underline{x}^{*(n_G)}_{k+1}\right)$. Thus, the acceptance ratio $\rho(\underline{x}_k, \underline{y})$ is only influenced by the dimensionality of the groups n_G. For each group $j \mid j = 1, \ldots, n_G$ a proposal density function q_j has to be constituted. The algorithm 2 outlines the transit from the state k to the state $k+1$.

Algorithm 2 Modified Metropolis-Hastings algorithm for high dimensions [1]

Given is $\underline{x}_k$:

1. Generate a candidate $\underline{x}^*_{k+1} = \left[\underline{x}^{*(1)}_{k+1}, \ldots, \underline{x}^{*(n_G)}_{k+1}\right]$:

   a) Simulate $\underline{y}^{(j)}$ according to $q_j(\underline{y}^{(j)} \mid \underline{x}^{(j)}_k)$

   b) Calculate the acceptance ratio:

$$\rho^{(j)}(\underline{x}^{(j)}_k, \underline{y}^{(j)}) = \min\left\{1, \frac{f(\underline{y}^{(j)})}{f(\underline{x}^{(j)}_k)} \frac{q_j(\underline{x}^{(j)}_k | \underline{y}^{(j)})}{q_j(\underline{y}^{(j)} | \underline{x}^{(j)}_k)}\right\}$$

   c) Define $\underline{x}^*_{k+1}$ using the acceptance ratio:

$$\underline{x}^{*(j)}_{k+1} = \begin{cases} \underline{y}^{(j)} & \text{with probability} \quad \rho^{(j)}(\underline{x}^{(j)}_k, \underline{y}^{(j)}) \\ \underline{x}^{(j)}_k & \text{with probability} \quad 1 - \rho^{(j)}(\underline{x}^{(j)}_k, \underline{y}^{(j)}) \end{cases}$$

2. Define $\underline{x}_{k+1}$ according to F:

$$\underline{x}_{k+1} = \begin{cases} \underline{x}^*_{k+1} & \text{when} \quad \underline{x}^*_{k+1} \in F \\ \underline{x}_k & \text{when} \quad \underline{x}^*_{k+1} \notin F \\ \underline{x}_k & \text{when} \quad \underline{x}^*_{k+1} = \underline{x}_k \end{cases}$$

3.4 Numerical efficiency of subset sampling

A comparison of the numerical efficiency of subset sampling and Monte Carlo simulation is shown in Fig. 2. In contrast to Monte Carlo simulation the required sample size for subset sampling cannot be predetermined in advance. In [1] a coarse approximation of the number of samples N_T, which is sufficient to estimate $\hat{P}_F$, is given by

$$N_T \approx mN = |\log P_F|^r \times \frac{(1+\gamma)(1-p_0)}{p_0 |\log p_0|^r \delta^2} \quad . \tag{23}$$

The exponent $r \leq 3$ in Eq. (23) specifies the correlation of the P_i of the individual subsets. The coefficient of variation δ_i of $\hat{P}_i\,(i = 2, \ldots, m)$ is

$$\delta_i = \sqrt{\frac{1-P_i}{P_i N}(1+\gamma_i)} \quad . \tag{24}$$

The value of γ_i denotes the correlation among Markov chain samples and depends itself on the selected proposal density function q. The accuracy of the conditional probability estimator $\hat{P}_i$ is reduced for $\gamma_i > 0$ compared to the case of independently distributed samples ($\gamma_i = 0$). The estimator $\hat{P}_i\,(i = 2, \ldots, m)$ is unbiased, the Strong Law of Large Numbers and the Central Limit Theorem are hold. Consequently this holds for estimator $\hat{P}_F$ of the failure probability (with $N \to \infty$) (see also Sec. 3.3.1).

The coefficient of variation δ of $\hat{P}_F$ is given by

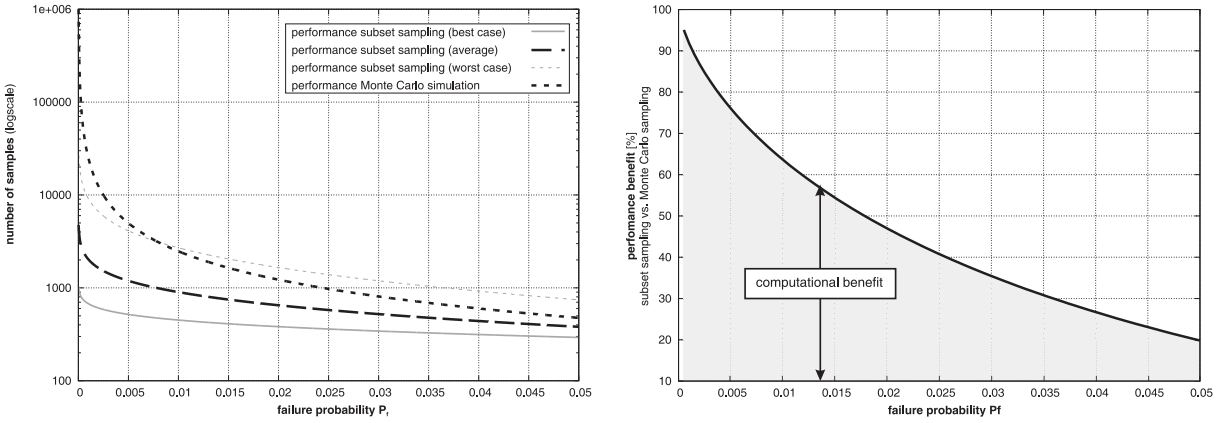
$$\delta = E\left[\frac{\hat{P}_F - P_F}{P_F}\right] \quad . \tag{25}$$

Eq. (25) is ruled by the correlation of the estimators $\hat{P}_i$. In case of uncorrelated samples the coefficient of variation δ is

$$\delta = \sqrt{\sum_{i=1}^{m} \delta_i^2} \quad . \tag{26}$$

Simulations show that despite the fact that P_i is generally estimated on the basis of correlated samples Eq. (26) is a valuable approximation. A substantial improvement in efficiency in comparison to the Monte

Carlo simulation is obtained for small failure probabilities P_F and small values for γ and r. The improvement can be noticed comparing $N \sim |\lg P_F|^r$ (subset sampling) with $N \sim 1/P_F$ (Monte Carlo simulation).



(a) Estimated number of samples for diffrent P_F

(b) Estimated computational benefit

Figure 2: Estimated computational effort of the Monte Carlo simulation vs. Subset sampling on the basis of Eq. (23)

4 Example

4.1 Benchmark study

The subset sampling method is demonstrated by means of a numerical example. The performance function $z(\underline{x})$ is given by

$$z(\underline{x}) = 8 \cdot \exp(-(x_1^2 + x_2^2)) + 2 \cdot \exp(-((x_1 - 5)^2 + (x_2 - 4)^2)) + 1 + \frac{x_1 \cdot x_2}{10}\,. \tag{27}$$

Both input variables x_1 and x_2 are assumed to be randomly distributed. Each random variable X_i is modeled with a beta distribution defined by the probability density function

$$f(x_i) = \frac{1}{B(p,q)} x_i^{p-1}(1 - x_i)^{q-1} \qquad | \; f(x_i) \geq 0 \quad . \tag{28}$$

A value of $q = p = 6$ is assigned to the distribution parameters q and p. The standardisation of the beta distribution is achieved with aid of the beta function

$$B(p,q) = \int_0^1 u^{p-1}(1 - u)^{q-1}\, du \quad . \tag{29}$$

As a result realizations are obtained in the interval $[-2, 6]$. A failure event occurs if the performance function exceeds values $z(\underline{x}) > 7.5$. Respectively, a failure is indicated by $g(\underline{x}) < 0$. The performance function $z(\underline{x})$ and limit state function $g(\underline{x})$ are shown in Fig. 3.

The sampling within the initial subset is realized with the aid of the Monte Carlo simulation. For the following subsets the subset sampling algorithm in conjunction with the Metropolis-Hastings algorithm is applied. As proposal density function a Gaussian proposal density function

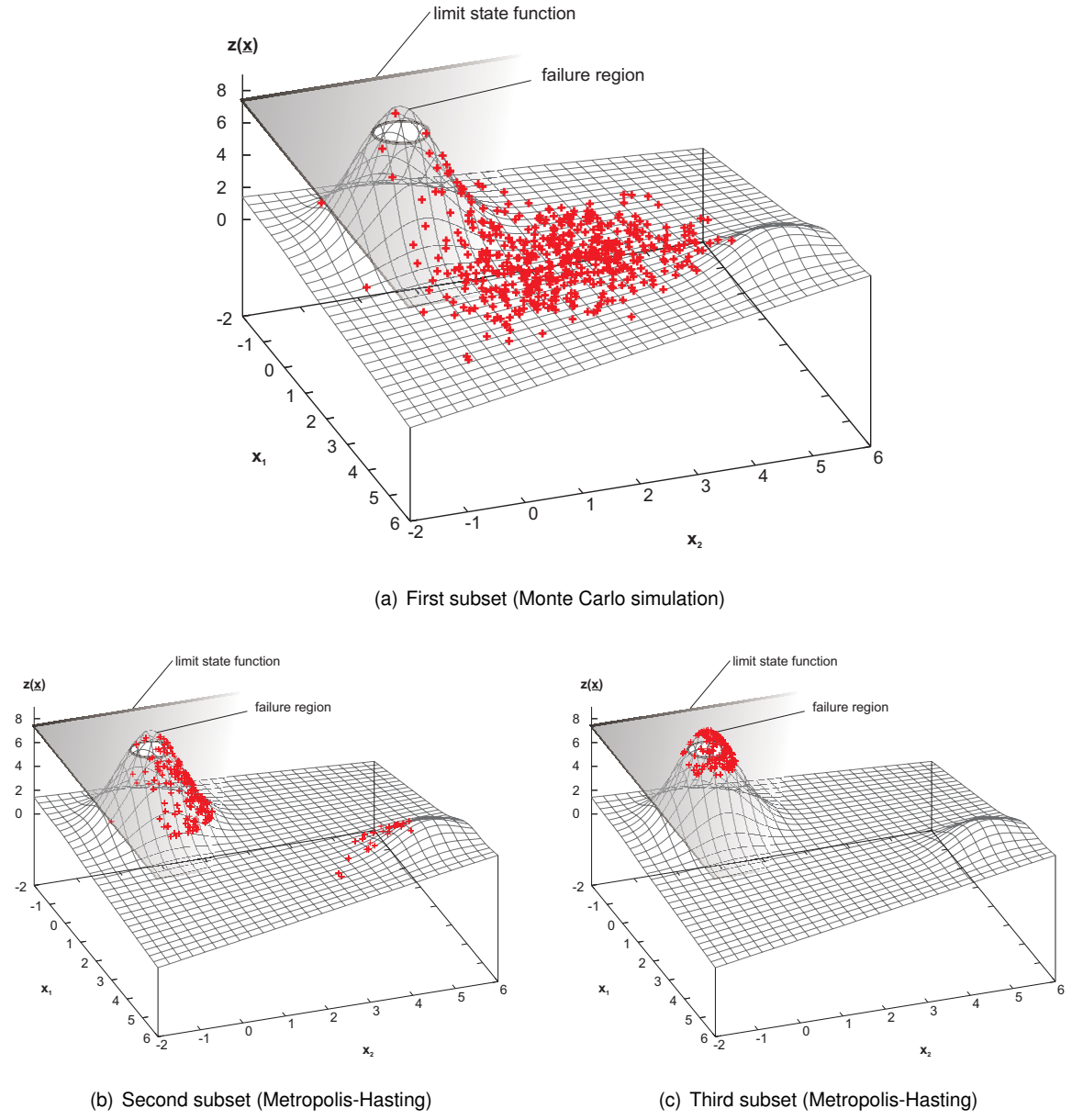
$$q(\underline{y} \mid \underline{x}_k) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\underline{y} - \underline{x}_k}{\sigma}\right)^2\right) \tag{30}$$

with a standard deviation of $\sigma = 2$ for each subset is deployed. For the determination of the failure probability three different sample sizes are investigated, see Tab. 2. The partial failure probability is predefined to $P_i = 0.1$ for each subset.

Table 2: Sample size of the different subset simulations

total number of samples	number of samples in the first subset (Monte Carlo)	number of samples in the following subsets
12000	6000	3000
8000	4000	2000
4000	2000	1000

The subset sampling algorithm is illustrated in Fig. 4. In consequence of the selected partial failure probability the simulation reaches the final failure region F after three subsets.



(a) First subset (Monte Carlo simulation)

(b) Second subset (Metropolis-Hasting)

(c) Third subset (Metropolis-Hasting)

Figure 3: Distribution of samples for each subset

In order to assess the quality of subset sampling the simulation is performed repeatedly ($1\,000$ times). The obtained results $\hat{P}_F$ are plotted histogram in Fig. 4. The determined mean values of $\hat{P}_F$ are listed in Tab. 3 for each case. In comparison to the mean value $\hat{P}_f = 4.05 \cdot 10^{-3}$, which is computed by a direct

Monte Carlo simulation with a sample size of $25\,000$, the results of the subset sampling yield a suitable approximation with a considerable small sample size.

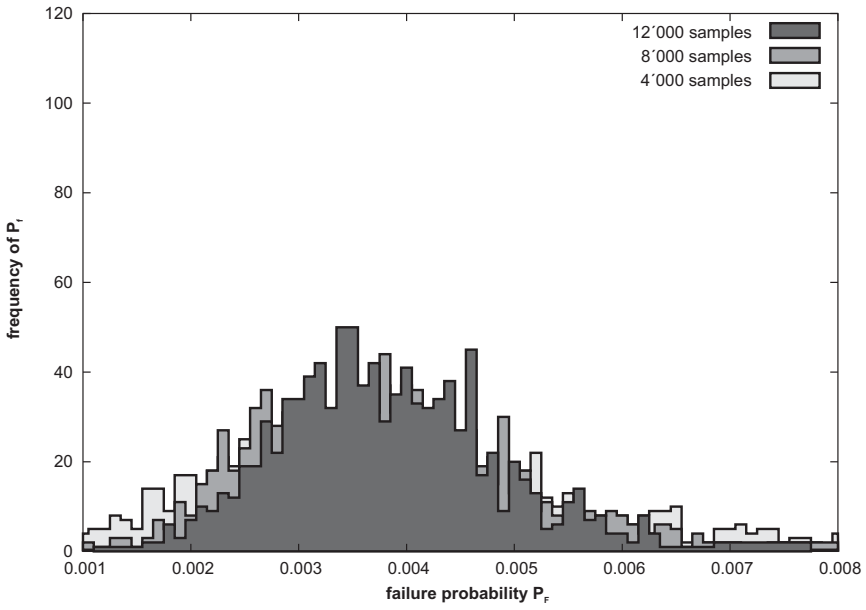


Figure 4: Histogram of the obtained failure probabilities P_F

Table 3: Resulting failure probabilities $\hat{P}_F$ for each subset simulation (mean value $\hat{P}_F$ of $1\,000$ simulations)

$N = 4000$	$N = 8000$	$N = 12000$
$\hat{P_F} = 3,8 \cdot 10^{-3}$	$\hat{P_F} = 3,77 \cdot 10^{-3}$	$\hat{P_F} = 3,88 \cdot 10^{-3}$

4.2 T-beam floor construction

A T-beam floor construction with a span of 6.0 m made of reinforced concrete (RC) is investigated with the aid of physically nonlinear analysis algorithms according to [5]. The structural reliability in relation to a certain threshold of deformation is assessed. A section of the structure with two beams is shown in Fig. 5. The section is discretized by means of 156 layered hybrid finite elements with assumed stress distribution. The beams are modeled using 12 concrete layers and the plate portions of the floor using 5 concrete layers. The steel reinforcement is specified as an uniaxial smeared layer in both cases. The physically nonlinear analysis is performed considering endochronic material laws for concrete and steel. Crack formation, tension stiffening, and steel yielding are taken into account.

The T-beam floor construction is part of an existing structure. The load prehistory includes the dead load g, the additional load g_1 and the live load p_2. Future loading includes additionally a live load p_2 and point loads P. The loads g_1, p_1, and p_2 are modeled as uniformly distributed superficial loads. The point loads P result from a suspended construction and are modeled as nodal loads of magnitude $P = 30$ kN. First, the reliability is computed, that the structure will maintain the certain threshold of deformation during operation. Thereby, $v = 30$ mm is selected as displacement threshold for the suspension point. The uncertainty of the live loads and of the material properties of the concrete are considered in the reliability assessment. The loads p_1 and p_2 are modeled as random variables with an extreme value distribution of maximum values type I (Gumbel distribution). The probability distribution function reads

$$F(p_i) = exp(-exp(-a(p_i - u))) \mid i = 1, 2 \quad . \tag{31}$$

The structural resistance depends mainly on the uncertain concrete compressive strength β. Thus, the compressive strength β is specified in this case as a Gaussian distributed stationary random field with

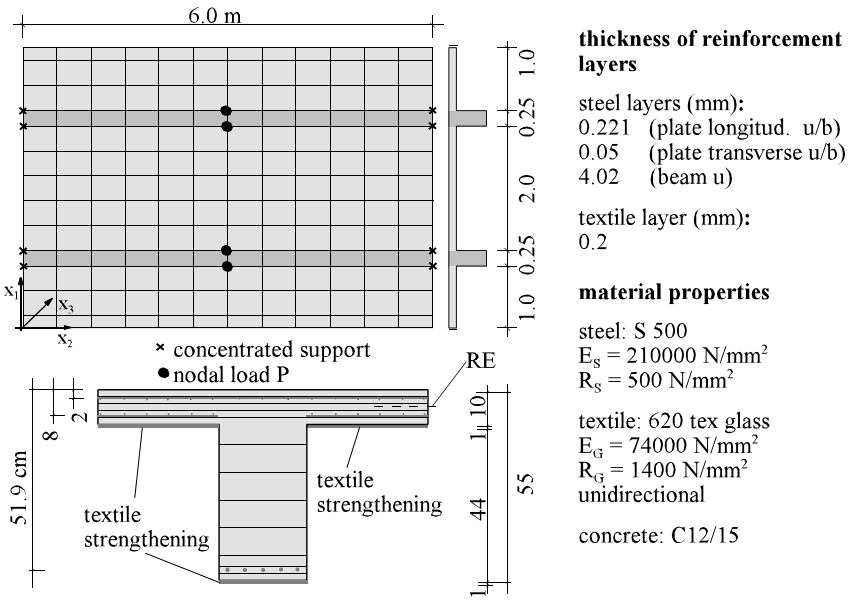


Figure 5: Geometry, material and FE model

the expected value $m_\beta = 20$ N/mm^2 and a standard deviation $\sigma_\beta = 2$ N/mm^2 . From the discretization of the random field according to the midpoint method yields 156 correlated random variables. Using an eigenvalue analysis of the correlation matrix only 34 independent random variables are determined, which are considered. Thus, together with the random loads a 36 dimensional input space results.

Two subset simulations and a direct Monte-Carlo simulation are computed. The partial failure probabilities are chosen to $P_i = 0.1$. The resulting failure probability and the number of samples are shown in Tab. 4. Subset sampling provides results in a manner that meet the results of the Monte Carlo simulation approximately, but only 30 percent of sampling points are required.

Table 4: Results of the reliability assessment

	P_F	number of MCS points	number of chains	point per chain
Subset sampling 1	0.021	400	2	200
Subset sampling 2	0.013	400	2	200
Monte-Carlo	0.025	2700		

Due to the additional loads it is expected that the reliability of the RC structure drops below a required boundary defined by $P_F = 0.005$. Therefore, the RC construction is strengthened with the aid of a textile reinforced concrete (TRC) layer. This new technology is currently under development. Initial investigations show promising results according to increase the load-bearing capacity and reliability [5].

In order to strengthen the damaged RC T-beam construction textile-reinforced fine-grade concrete layers are applied to the surface. The textile reinforcement of these layers consists of filament yarns (rovings) connected together by stitching yarn (see Fig. 6). Each roving comprises a large number of single filaments. The textile reinforcement may consist of different fiber materials, e.g., alkali resistant glass (ARG) or carbon. The strengthening of reinforced concrete (aged RC) with textile-reinforced fine-grade concrete (textile concrete) results in a composite structure (Fig. 6).

The reliability of the strengthened structure is computed. Beside the uncertainty of the live loads and the material properties of the aged RC, the uncertainty of the fine-grade concrete is considered.

The tensile strength β_t is specified in this case as a Gaussian distributed stationary random field. Again from the discretization into 156 finite elements, 156 correlated random variables result and only 34 independent random variables are considered. Thus, together with the random loads and the random field of the concrete compressive strength a 70 dimensional input space is obtained.

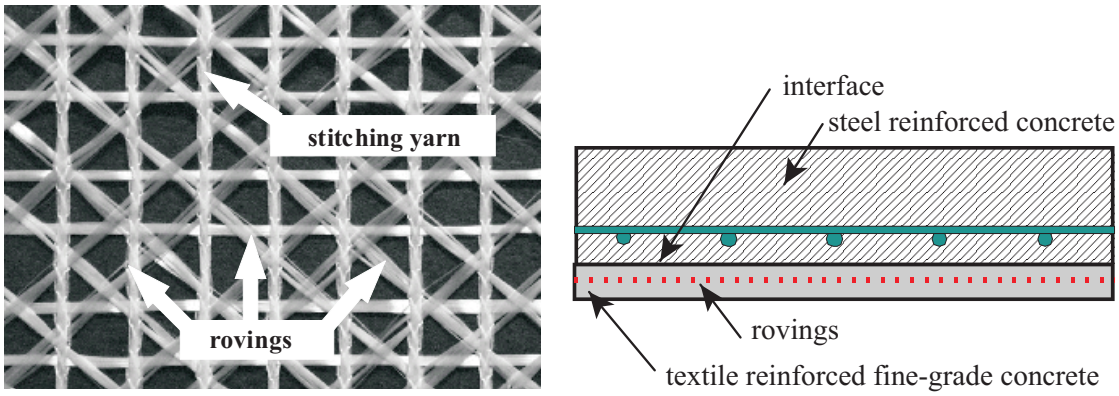


Figure 6: Example of textile reinforcement and composite material

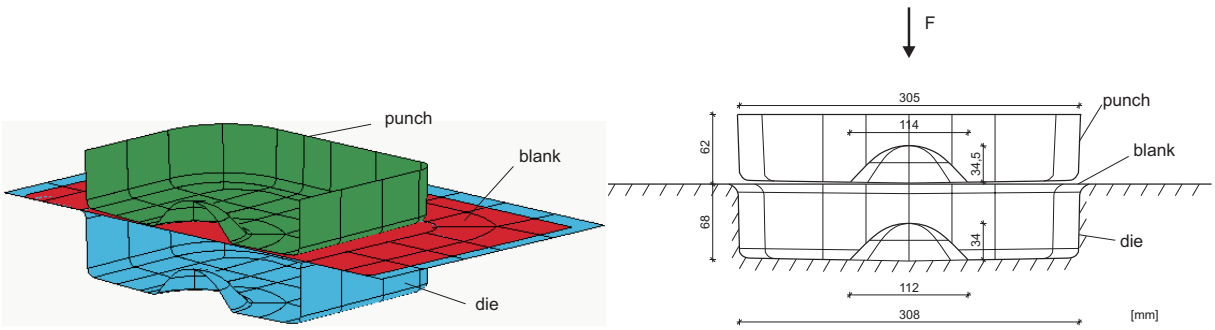
Four subset simulations and a direct Monte Carlo simulation are performed. The resulting failure probability and the number of samples are shown in Tab. 5. In this case the number of sample points decreases to 15 percent. The required reliability is preserved.

Table 5: Results of the reliability assessment of the strengthened structure

	P_F	number of MCS points	number of chains	points per chain
Subset sampling 1	0.0014	500	2	250
Subset sampling 2	0.0027	500	2	250
Subset sampling 3	0.0036	500	5	100
Subset sampling 4	0.0027	500	5	100
Monte Carlo	0.0023	10000		

4.3 Deep-drawing of a sheet metal

Deep-drawing is a sheet metal forming process. During the deep-drawing procedure the thickness of the sheet metal is reduced in some regions. If the thickness reduction exceeds a certain threshold the sheet has to be rejected. In order to predict the number of rejected sheets a reliability assessment of the process should be performed. The geometry of the investigated example is shown in Fig. 7.

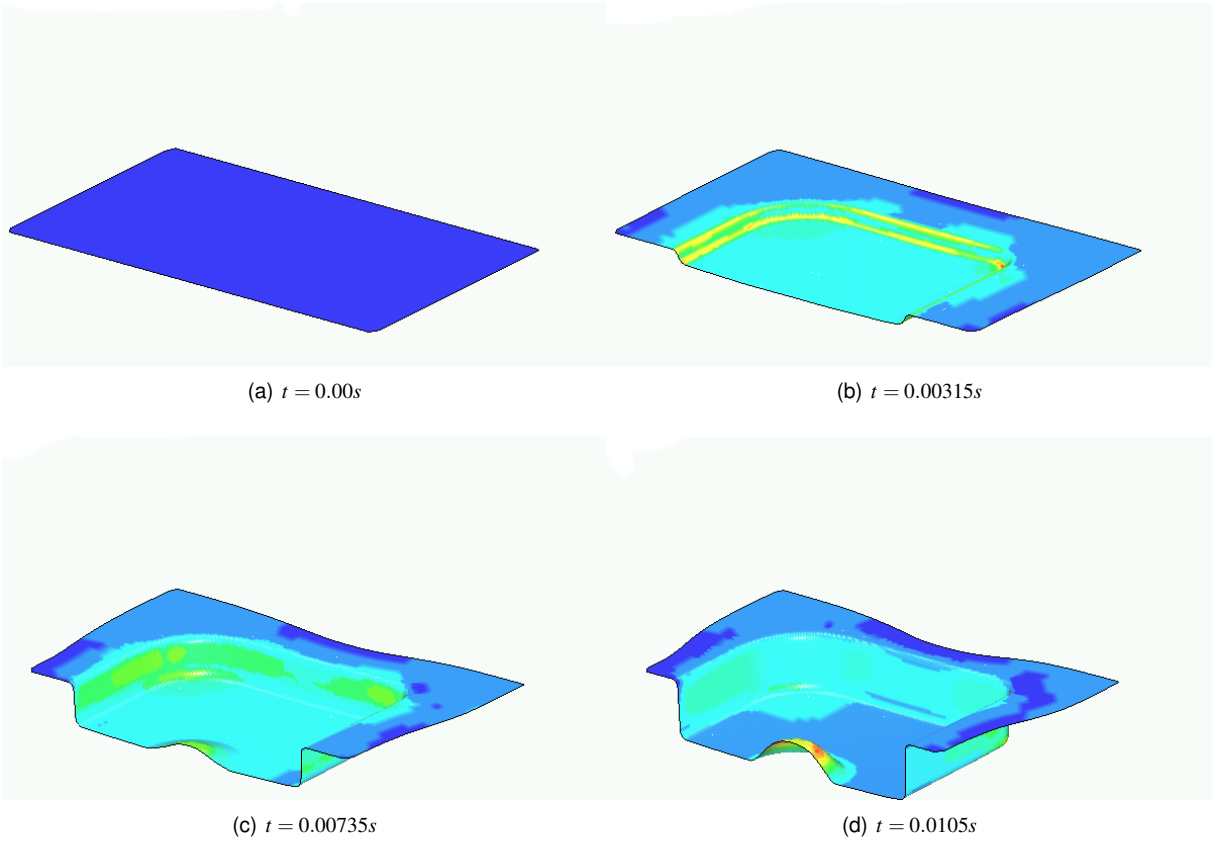


(a) 3D-Modell (b) 2D-Modell

Figure 7: Discretized model with significant measures

For the determination of the thickness reduction of the sheet metal the units shown in Fig. 7 are discretized with finite elements. The program LS-DYNA is applied for computing the thickness reduction.

A realization of the computed sheet metal forming process is depicted in Fig. 8. In the reliability assessment only the last time step $t = 0.0105\,s$ is considered.



(a) $t = 0.00s$ (b) $t = 0.00315s$

(c) $t = 0.00735s$ (d) $t = 0.0105s$

Figure 8: Deep-drawing process for one realization

Twelve uncertain input variables $X_1, \ldots, X_{12}$ are selected and modeled by means of a beta distribution. The limit state $g(\underline{x}) = 0$ is defined as a thinning of the steel sheet to 83.4 percent. Thus, a thickness reduction greater than 16.6 percent denotes a failure. The examined point is situated in the lower region of the concavity. The failure probability is determined with the aid of subset sampling. The number of samples for the first subset (Monte Carlo simulation) is chosen with 600 samples and the following subsets with 300 samples. One Markov chain for each subset is applied. The proposal density function is chosen to be a normal density function with a standard deviation $\sigma = 1$. Three subset sampling simulations are performed. A Monte Carlo simulation with a total number of 4300 samples is performed additionally to enable the validation of the results. The coefficient of variation of the Monte Carlo simulation is estimated to 20 percent using Eq. (8).

The evaluation of results shows a satisfactory conformity of results (see Tab. 6) considering that the Monte Carlo simulation itself possesses a coefficient of variation of 20 percent. Concerning the Markov chain Monte Carlo simulation the acceptance ratio is considerable high caused by the applied modified Metropolis-Hastings algorithm. Thereby, the probability of rejection of each component of the candidate $\underline{x}^*_{k+1} = \left[x^{*(1)}_{k+1}, \ldots, x^{*(12)}_{k+1}\right]$ is small.

Table 6: Results of the reliability assessment

	Monte Carlo	Subset sampling		
		1	2	3
P_F	0.006	0.0021	0.0034	0.0023
Sample size	4300		1200	

5 Conclusions

Subset sampling provides a useful method for the reliability analysis of structures. This has been demonstrated by means of examples. Compared to a direct Monte Carlo simulation the numerical efficiency is significantly improved, in particular, in the computation of small failure probabilities such as $P_F < 10^{-3} \dots 10^{-4}$. As a consequence of the reduced number of computations an increasing in the variance of the failure probability is observed. Further developments are focused on the remedy of this weakness.

Acknowledgments

The authors gratefully acknowledge the support of the German Research Foundation (DFG). Furthermore the authors acknowledge the contribution of Dipl.-Ing. J. Sieberer. The example 4.3 was provided by courtesy of the BMW group (TK 421).

References

[1] AU, S.-K.: *On the Solution of First Excursion Problems by Simulation with Applications to Probabilistic Seismic Performance Assessment*. PhD thesis, California Institute of Technology, Pasadena, 2001.

[2] HAMMERSLEY, J. and D. HANDSCOMB: *Monte Carlo Methods*. Methuen & Co Ltd, London, New York, 1964.

[3] KOUTSOURELAKIS, P. S., H. J. PRADLWARTER, and G. I. SCHUËLLER: *Reliability of structures in high dimensions, part i: algorithms and applications*. Probabilistic Engineering Mechanics, 19:409–417, 2004.

[4] KRENGEL, U.: *Einführung in die Wahrscheinlichkeitstheorie und Statistik*. Vieweg, Wiesbaden, 1990.

[5] MÖLLER, B., W. GRAF, A. HOFFMANN, and F. STEINIGEN: *Numerical simulation of rc structures with textile reinforcement*. Computers & Structures, Vol. 83(19-20):1659–1688, 2005.

[6] PELLISSETTI, M., G. SCHUËLLER, H. PRADLWARTER, A. CALVI, S. FRANSEN, and M. KLEIN: *Reliability analysis of spacecraft structures under static and dynamic loading*. Computers & Structures, 84(21):1313–1325, Aug. 2006.

[7] ROBERT, C. P. and G. CASELLA: *Monte Carlo Statistical Methods*. Springer, Berlin, 2004.

[8] RUBINSTEIN, R.: *Simulation and the Monte Carlo Method*. Wiley, New York, 1981.

[9] SICKERT, J.-U.: *Fuzzy-Zufallsfunktionen und ihre Anwendung bei der Tragwerksanalyse und Sicherheitsbeurteilung*. Heft 9, TU Dresden, 2005.

[10] SPAETHE, G.: *Die Sicherheit tragender Baukonstruktionen*. Springer, Wien, New York, 1992.

[11] TORRIE, G. M. and J. P. VALLEAU: *Nonphysical sampling distributions in monte carlo free-energy estimation: umbrella sampling*. Journal of Computational Physics, (23):187–199, 1977.

