Sequential importance sampling for structural reliability

I. Papaioannou & D. Straub
Engineering Risk Analysis Group, Technische Universitt Mnchen, Arcisstr. 21, 80290 Mnchen, Germany

C. Papadimitriou
Department of Mechanical Engineering, University of Thessaly, Pedion Areos, 38334 Volos, Greece

ABSTRACT: This paper discusses the application of sequential importance sampling (SIS) to the estimation of the probability of failure in structural reliability. SIS was developed originally in the statistical community for exploring posterior distributions and estimating normalizing constants. The basic idea is to gradually translate samples from the prior distribution to samples from the posterior distribution through a sequential reweighting operation. In the context of structural reliability, SIS can be applied to produce samples of an approximately optimal importance sampling density, which can then be used for estimating the sought probability through importance sampling. The transition of the samples is defined through the construction of a sequence of intermediate distributions. We discuss a particular choice of the intermediate distributions and the properties of the derived algorithm. Moreover, we introduce an MCMC algorithm for application within the SIS procedure that is especially efficient for tackling high-dimensional problems.

1 INTRODUCTION

Structural reliability analysis requires the evaluation of the probability of failure, defined by the following n-fold integral:

$$P_f = \int_{g(\mathbf{x})\leq 0} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \tag{1}$$

where $\mathbf{X}$ is an n-dimensional random vector and models the system variables that are expected to present an uncertain behavior, $f_{\mathbf{X}}(\mathbf{x})$ is the joint PDF of $\mathbf{X}$ and $g(\mathbf{x}) \leq 0$ defines the failure event. The function $g(\mathbf{x})$ is usually termed limit-state function and it can include one or several distinct failure modes (Ditlevsen and Madsen 1996).

It is common to transform the random variables $\mathbf{X}$ to a probability space $\mathbf{U}$ consisting of independent standard normal random variables. This is achieved by an one-to-one transformation $\mathbf{U} = \mathbf{T}(\mathbf{X})$ (Hohenbichler and Rackwitz 1981; Der Kiureghian and Liu 1986). The probability of failure can be expressed in the transformed space as

$$P_f = \int_{G(\mathbf{u})\leq 0} \varphi_n(\mathbf{u}) d\mathbf{u} \tag{2}$$

where φ_n is the n-variate standard normal PDF and $G(\mathbf{u}) = g(\mathbf{T}^{-1}(\mathbf{u}))$ is the limit-state function in the $\mathbf{U}$-space.

The integral in Equation 2 can be evaluated by a variety of existing approaches (Ditlevsen and Madsen 1996; Lemaire 2009). Among these, simulation methods are often preferred because of their robustness in dealing with complex engineering models. The probability integral can be expressed as the expectation of the indicator function $I\left(G(\mathbf{u}) \leq 0\right)$, where $I\left(G(\mathbf{u}) \leq 0\right) = 1$ if

$G(\mathbf{u}) \leq 0$ and $I\left(G(\mathbf{u}) \leq 0\right) = 0$ otherwise. Standard Monte Carlo estimates P_f by generating n_s independent samples $\{\mathbf{u}_k, k = 1, \ldots, n_s\}$ from the PDF $\varphi_n(\mathbf{u})$ and taking the sample mean of $I\left(G(\mathbf{u}) \leq 0\right)$, i.e.

$$\hat{P}_f = \hat{\mathrm{E}}_{\varphi_n}\left[I\left(G(\mathbf{u}) \leq 0\right)\right] = \frac{1}{n_s}\sum_{k=1}^{n_s} I\left(G(\mathbf{u}_k) \leq 0\right) \tag{3}$$

The estimate of Equation 3 is unbiased and has coefficient of variation:

$$\delta_{\hat{P}_f} = \sqrt{\frac{1 - P_f}{n_s P_f}} \tag{4}$$

$\delta_{\hat{P}_f}$ is a measure of the statistical accuracy of $\hat{P}_f$. Although $\delta_{\hat{P}_f}$ does not depend on the dimension of the random variable space n, it is inversely proportional to the target probability P_f. Hence for a probability in the order of 10^{-k}, crude Monte Carlo requires approximately 10^{k+2} samples to achieve an accuracy of $\delta_{\hat{P}_f} = 10\%$.

Several methods have been proposed that aim at reducing the variance of the crude Monte Carlo estimate. These include importance sampling (IS) and its adaptive variants (Bucher 1988; Au and Beck 2003; Kurtz and Song 2013), line sampling (Hohenbichler and Rackwitz 1988;Koutsourelakis et al. 2004) and subset simulation (SubS) (Au and Beck 2001). All of the above methods are based on producing samples that explore the failure region. In this paper, we discuss a sampling method that adaptively samples the failure region, termed sequential importance sampling (SIS). SIS was developed in the statistical community for exploring posterior distributions and estimating normalizing constants (Neal 2001;Chopin 2002;Del Moral et al. 2006). Although the published variants of the method diverge in their implementation approaches, they are all based in the same principle of gradually transforming samples from a prior to samples from a posterior distribution through a sequential reweighting operation. A variant of the method was introduced in the engineering community as transitional Markov chain Monte Carlo (Ching and Chen 2007). A method based on SIS as well as on ideas found in (Neal 2005) has been proposed for structural reliability in (Katafygiotis and Zuev 2007) and for reliability-based optimization in (Beaurepaire et al. 2013).

Here, we discuss the principle of SIS for structural reliability and present an implementation that is suitable for application to high-dimensional problems. The performance of the method is demonstrated with numerical examples.

2 SEQUENTIAL IMPORTANCE SAMPLING FOR STRUCTURAL RELIABILITY

In this section, we describe SIS for structural reliability. We first review standard IS; then we describe SIS for sampling from a sequence of distributions; a particular sequence of distributions for application to structural reliability is discussed next; subsequently we introduce a Markov chain Monte Carlo (MCMC) algorithm as an important ingredient of SIS for application to high-dimensional problems; we finally draw a connection of SIS with SubS.

2.1 *Importance sampling*

Let $h(\mathbf{u})$ be a positive density referred to as the IS function. The integral in Equation 2 can be rewritten as:

$$P_f = \int_{\mathbb{R}^n} I\left(G(\mathbf{u}) \leq 0\right) w(\mathbf{u}) h(\mathbf{u}) d\mathbf{u} = \mathrm{E}_h\left[I\left(G(\mathbf{u}) \leq 0\right) w(\mathbf{u})\right] \tag{5}$$

where $w(\mathbf{u}) = \frac{\varphi_n(\mathbf{u})}{h(\mathbf{u})}$ is the so-called importance weight function. An estimate of P_f can be obtained by generating samples $\{\mathbf{u}_k, k = 1, \ldots, n_s\}$ from $h(\mathbf{u})$ and taking the sample mean of $I\left(G(\mathbf{u}) \leq 0\right) w(\mathbf{u})$, i.e.

$$\hat{P}_f = \hat{\mathrm{E}}_h\left[I\left(G(\mathbf{u}) \leq 0\right) w(\mathbf{u})\right] = \frac{1}{n_s}\sum_{k=1}^{n_s} I\left(G(\mathbf{u}_k) \leq 0\right) w(\mathbf{u}_k) \tag{6}$$

The probability estimate of Equation 6 is unbiased provided that the support of $h(\mathbf{u})$ contains the failure domain $G(\mathbf{u}) \leq 0$. An appropriate choice of the IS function can significantly reduce the variance of the crude Monte Carlo estimate. The theoretically optimal IS is given by the following expression:

$$h_{\text{opt}}(\mathbf{u}) = \frac{1}{P_f} I\left(G(\mathbf{u}) \leq 0\right) \varphi_n(\mathbf{u}) \tag{7}$$

Indeed the IS function of Equation 7 leads to a variance of the IS estimate of zero. However, the optimal IS function cannot be used in practice since it requires knowledge of P_f. Alternative choices are unimodal (Schuëller and Stix 1987; Bucher 1988) or multimodal (Au and Beck 2003; Kurtz and Song 2013) densities based on initial sampling or other type of calculations. However, it has been discussed (Au and Beck 2003; Katafygiotis and Zuev 2008) that in many settings involving a large number of random variables, IS based on such densities may fail to describe the important region leading to a dramatic increase of the variance of the resulting estimate.

2.2 *Sequential importance sampling*

Consider a sequence of distributions $\{h_j(\mathbf{u}), j = 0, \ldots, M\}$, where each distribution is known up to a normalizing constant, i.e.

$$h_j(\mathbf{u}) = \frac{\eta_j(\mathbf{u})}{P_j} \tag{8}$$

where $\eta_j(\mathbf{u})$ is known pointwise and the normalizing constant P_j is unknown. We assume that $\eta_0(\mathbf{u}) = h_0(\mathbf{u})$ and hence $P_0 = 1$. We further assume that $h_0(\mathbf{u})$ is easy to sample from. We are interested in obtaining samples from $h_M(\mathbf{u})$ and estimating the normalizing constant P_M. The idea of SIS is to sample the distributions $\{h_j(\mathbf{u}), j = 0, \ldots, M\}$ in a step-wise manner and estimate each normalizing constant P_j by IS using as IS density the function $h_{j-1}(\mathbf{u})$. Assume that at step $j-1$ samples $\{\mathbf{u}_k, k = 1, \ldots, n_s\}$ from $h_{j-1}(\mathbf{u})$ are available. The constant P_j can be written as:

$$P_j = \int_{\mathbb{R}^n} \eta_j(\mathbf{u}) d\mathbf{u} = P_{j-1} \int_{\mathbb{R}^n} w_j(\mathbf{u}) h_{j-1}(\mathbf{u}) d\mathbf{u} = P_{j-1} \mathrm{E}_{h_{j-1}}\left[w_j(\mathbf{u})\right] \tag{9}$$

where $w_j(\mathbf{u}) = \frac{\eta_j(\mathbf{u})}{\eta_{j-1}(\mathbf{u})}$. An estimate of the ratio of normalizing constants $S_j = \frac{P_j}{P_{j-1}}$ is given by:

$$\hat{S}_j = \frac{\hat{P}_j}{\hat{P}_{j-1}} = \hat{\mathrm{E}}_{h_{j-1}}\left[w_j(\mathbf{u})\right] = \frac{1}{n_s} \sum_{k=1}^{n_s} w_j(\mathbf{u}_k) \tag{10}$$

To obtain an accurate estimate $\hat{S}_j$, we need to ensure that the two densities $h_{j-1}(\mathbf{u})$ and $h_j(\mathbf{u})$ do not vary significantly. This can be controlled by selecting $\eta_j(\mathbf{u})$ such that the variance of the importance weights is small. Given samples from $h_{j-1}(\mathbf{u})$, we can obtain samples from $h_j(\mathbf{u})$ applying the following resample-move scheme. First, we apply a resampling method that selects randomly with replacement samples from $\{\mathbf{u}_k, k = 1, \ldots, n_s\}$ with probability assigned to each kth sample proportional to $w_j(\mathbf{u}_k)$ (Doucet et al. 2001). We then move the resulting samples in regions of high probability mass of $h_j(\mathbf{u})$ by applying MCMC with invariant distribution $h_j(\mathbf{u})$. This procedure is repeated for each subsequent step and an estimate of P_M is obtained as:

$$\hat{P}_M = \prod_{j=1}^{M} \hat{S}_j \tag{11}$$

2.3 *Choice of intermediate distributions*

In the context of structural reliability, SIS can be applied to obtain samples from an approximation of the optimal IS density of Equation 7. The indicator function $I\left(G(\mathbf{u}) \leq 0\right)$ can be expressed by the following limit [e.g. see (Spanier and Oldham 1987)]

$$I\left(G(\mathbf{u}) \leq 0\right) = \lim_{\sigma \to 0} \Phi\left(-\frac{G(\mathbf{u})}{\sigma}\right) \tag{12}$$

where Φ is the standard normal CDF. Choosing $\sigma = \sigma_M$, with σ_M small enough, we can approximate $I\left(G(\mathbf{u}) \leq 0\right)$ by the following expression

$$I\left(G(\mathbf{u}) \leq 0\right) \approx \Phi\left(-\frac{G(\mathbf{u})}{\sigma_M}\right) \tag{13}$$

Inserting Equation (13) into Equation (7), we obtain the following approximation of the optimal IS density:

$$h_{\text{opt}}(\mathbf{u}) \approx h_M(\mathbf{u}) = \frac{1}{P_M}\Phi\left(-\frac{G(\mathbf{u})}{\sigma_M}\right)\varphi_n(\mathbf{u}) \tag{14}$$

Define the sequence of distribution $\{h_j(\mathbf{u}), j = 0, \ldots, M\}$, with

$$h_j(\mathbf{u}) = \frac{1}{P_j}\Phi\left(-\frac{G(\mathbf{u})}{\sigma_j}\right)\varphi_n(\mathbf{u}) = \frac{1}{P_j}\eta_j(\mathbf{u}) \tag{15}$$

where $\infty = \sigma_0 > \ldots > \sigma_M > 0$. This sequence, which was also used in the method of (Beaurepaire et al. 2013), allows to gradually approach the density of Equation 14 by a series of smooth approximations of the optimal IS density, as demonstrated in Figure 1. We can apply SIS to sample this distribution sequence and estimate the constant P_M. Note that $h_0(\mathbf{u}) = \varphi_n(\mathbf{u})$, which can be readily sampled from. To ensure that each pair of consecutive distributions are not too different from one another, we can select the parameters σ_j adaptively, such that the sample coefficient of variation $\hat{\delta}_{w_j}$ of the importance weights adheres a target value δ_{target}. Hence, at each step of the SIS procedure, we need to solve the following optimization problem

$$\sigma_j = \text{argmin}\|\hat{\delta}_{w_j} - \delta_{\text{target}}\| \tag{16}$$

The procedure is stopped when the coefficient of variation of the samples $\{I\left(G(\mathbf{u}_k) \leq 0\right)\frac{\varphi_n(\mathbf{u}_k)}{h_j(\mathbf{u}_k)}, k = 1, \ldots, n_s\}$ is smaller than δ_{target}, M is set to the current step j and the probability of failure is evaluated by IS with the available samples from $h_M(\mathbf{u})$.

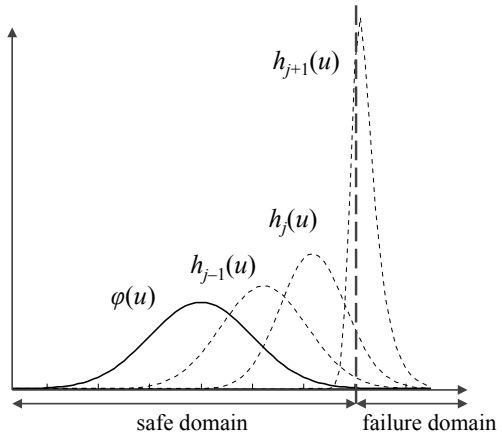

Figure 1: Sequence of intermediate distributions

2.4 *MCMC sampling*

At each step j of the SIS procedure, MCMC sampling is used to move the samples $\{\mathbf{u}_k, k = 1, \ldots, n_s\}$ obtained from the applied resampling scheme to regions of high probability density of $h_j(\mathbf{u})$. The basic idea of MCMC is to sample states of a Markov chain with stationary distribution equal to the target distribution. The most widely used MCMC method is the Metropolis-Hastings (M-H) algorithm (Hastings 1970). The transition from a state $\mathbf{u}_0$ to the next state $\mathbf{u}_1$ of the M-H algorithm for sampling from $h_j(\mathbf{u}) \propto \Phi\left(-G(\mathbf{u})/\sigma_j\right)\varphi_n(\mathbf{u})$ is as follows:

1. Generate a candidate state $\mathbf{v}$ from a proposal probability density $q(\cdot|\mathbf{u}_0)$
2. Calculate the ratio

$$r(\mathbf{u}_0, \mathbf{v}) = \frac{\Phi\left(-G(\mathbf{v})/\sigma_j\right)\varphi_n(\mathbf{v})q(\mathbf{u}_0|\mathbf{v})}{\Phi\left(-G(\mathbf{u}_0)/\sigma_j\right)\varphi_n(\mathbf{u}_0)q(\mathbf{v}|\mathbf{u}_0)} \tag{17}$$

3. Accept $\mathbf{v}$ with probability $\alpha(\mathbf{u}_0, \mathbf{v}) = \min\{1, r(\mathbf{u}_0, \mathbf{v})\}$, i.e set $\mathbf{u}_1 = \mathbf{v}$ with probability $\alpha(\mathbf{u}_0, \mathbf{v})$ and $\mathbf{u}_1 = \mathbf{u}_0$ with probability $1 - \alpha(\mathbf{u}_0, \mathbf{v})$.

The M-H update leaves the stationary distribution $h_j(\mathbf{u})$ unchanged. Hence, the Markov chain will asymptotically converge to $h_j(\mathbf{u})$, under certain restrictions on the choice of the proposal density (Hastings 1970). The transient period that is required until the Markov chain reaches its stationary state is termed burn-in period. Moreover, the generated samples will be correlated according to the correlation of the Markov process which depends on the particular choice of $q(\cdot|\mathbf{u}_0)$.

A standard choice of the proposal distribution is a multivariate symmetric distribution centered at the current state $\mathbf{u}_0$. The resulting M-H samplers become inefficient in sampling high dimensional target distributions (Au and Beck 2001; Katafygiotis and Zuev 2007). This inefficiency is due to the appearance of the ratio $\frac{\varphi_n(\mathbf{v})}{\varphi_n(\mathbf{u}_0)}$ in the acceptance probability $\alpha(\mathbf{u}_0, \mathbf{v})$. As explained in (Katafygiotis and Zuev 2007), this ratio becomes extremely small in high dimensions and hence the probability of obtaining repeated samples will be extremely high. To overcome this issue, we propose to choose $q(\cdot|\mathbf{u}_0)$ as the multivariate Gaussian density conditional on the current state $\mathbf{u}_0$, i.e.

$$q(\mathbf{v}|\mathbf{u}_0) = \varphi_n\left(\mathbf{v} - \rho\mathbf{u}_0, (1-\rho^2)\mathbf{I}\right) \tag{18}$$

The density of Equation 18 assumes that the current and candidate state are jointly Gaussian with component-wise cross-correlation coefficient ρ. We therefore term the resulting update conditional sampling M-H algorithm. Inserting this proposal to the ratio of Equation 17, we get

$$r(\mathbf{u}_0, \mathbf{v}) = \frac{\Phi\left(-G(\mathbf{v})/\sigma_j\right)}{\Phi\left(-G(\mathbf{u}_0)/\sigma_j\right)} \tag{19}$$

That is, the fraction $\frac{\varphi_n(\mathbf{v})}{\varphi_n(\mathbf{u}_0)}$ disappears and $r(\mathbf{u}_0, \mathbf{v})$ depends only on the ratio of one-dimensional distribution functions with arguments limit-state function values. This ratio depends on the proximity of the limit-state function values for the current and candidate state and not on the dimension of the random variable space. Hence, the conditional sampling M-H algorithm is suitable for application to high dimensional problems. It should be noted that the proposal density of Equation 18 is often used for sampling Gaussian process prior models (Neal 1998).

In standard SIS, to generate n_s samples from $h_j(\mathbf{u})$, one MCMC move is performed starting from each of the n_s seeds that resulted from the resampling step (Chopin 2002). Hence, the seeds follow only asymptotically the distribution $h_j(\mathbf{u})$. Therefore, when applying local MCMC transitions, a single move might not be sufficient for the chain to converge to its stationary state, even if $h_{j-1}(\mathbf{u})$ and $h_j(\mathbf{u})$ do not vary much. It is therefore beneficial to use less chains and allow them to mix properly. We propose to resample $n_c < n_s$ samples from the weighted sample approximation of $h_j(\mathbf{u})$ and run a Markov chain of length n_s/n_c starting from each of the n_c seeds. This will reduce the burn-in effect on the statistics of the probability estimate, as will be demonstrated in Section 3.

2.5 *Connection to Subset Simulation*

SubS is an adaptive simulation method proposed in (Au and Beck 2001). The method is based on expressing the failure event $F = [\mathbf{u} \in \mathbb{R}^n : G(\mathbf{u} \leq 0)]$ as an intersection of intermediate failure events that are nested; i.e. it holds $F_0 \subset F_1 \subset \ldots \subset F_M$, F_0 is the certain event and $F_M = F$. The probability of failure is expressed as:

$$P_f = \Pr(F) = \Pr\left(\bigcap_{j=0}^{M} F_j\right) = \prod_{j=1}^{M} \Pr(F_j|F_{j-1}) \tag{20}$$

That is, P_f is expressed as a product of larger conditional probabilities. The intermediate events are defined as $F_j = [\mathbf{u} \in \mathbb{R}^n : G(\mathbf{u} \leq b_j)]$, where $\infty = b_0 > b_1 > \ldots > b_M = 0$. The values of b_j are chosen adaptively, so that the estimates of the conditional probabilities correspond to a chosen value p. To this end, n_s samples are simulated conditional on each intermediate failure event F_{j-1} and b_j is set to the p-percentile of the limit-state function values. Samples conditional on the certain event F_0 are generated by crude Monte Carlo. Samples conditional on the events $\{F_j, j = 1, \ldots, M-1\}$ are computed by simulating states of Markov chains through MCMC starting from the samples conditional on F_{j-1} for which $G(\mathbf{u}) < b_j$.

SubS can be understood as a special case of the SIS method discussed in Section 2.2 if the intermediate distributions are chosen as the optimal IS density for each of the intermediate failure domains, i.e.

$$h_j(\mathbf{u}) \propto I\left(G(\mathbf{u}) \leq b_j\right) \varphi_n(\mathbf{u}) \tag{21}$$

However, SubS differs from the SIS method discussed here, because of the particular choice of the intermediate distributions of Equation 21, that is a series of multivariate normal distributions conditional on a set of nested failure events. Therefore, at each sampling step j, the available samples from $h_{j-1}(\mathbf{u})$ that fell in F_j will already be distributed according to $h_j(\mathbf{u})$. Hence, the resampling step discussed in Section 2.2 is not required. Moreover, all states of the simulated Markov chains will be distributed according to the target distribution $h_j(\mathbf{u})$. Thus, the Markov chains do not require a burn-in period to reach their stationary states.

3 EXAMPLES

In this section, we investigate the performance of the SIS method with three numerical examples. In the first two examples, we compare the coefficient of variation (CV) of the probability estimates with the one obtained by SubS. The third example demonstrates the applicability of the conditional sampling M-H (CSM-H) algorithm to a high dimensional problem. For all examples, the CV of the probability estimate is computed with 500 independent simulation runs.

3.1 *Convex limit-state function*

The first example consist of a convex limit-state function, defined at the standard normal space (**U**-space) as follows:

$$G_1(\mathbf{u}) = 0.1(u_1 - u_2)^2 - \frac{1}{\sqrt{2}}(u_1 - u_2) + 2.5 \tag{22}$$

The corresponding probability of failure is computed as 4.21×10^{-3}. Table 1 compares the CV of the probability estimates obtained by SubS and SIS for different number of samples per level, i.e. per intermediate distribution, namely $n_s = 500$, 1000 and 2000. The conditional probability for SubS is set to $p = 0.1$ and the target CV for selection of the intermediate distributions for SIS is set to $\delta_{\text{target}} = 1.5$. These values result in the same computational cost on average for the two methods. We investigate three different MCMC settings for the generation of the samples from each distribution in SIS: (a) Resample $n_c = n_s$ seeds and run n_c chains with unit length; (b) Resample $n_c = 0.1n_s$ seeds and run n_c chains with length 10; (c) Resample $n_c = 0.1n_s$ seeds, run n_c chains with length 15 and discard the first 5 samples from each chain to account for burn-in. Note that for SubS with $p = 0.1$ the total number of failure points per level and hence the total number of chains will be $0.1n_s$. MCMC sampling is performed applying the CSM-H algorithm described in Section 2.4 with correlation parameter $\rho = 0.8$. For SubS, MCMC is performed applying a conditional sampling algorithm presented in (Papaioannou et al. 2014) with correlation parameter between current and candidate state $\rho = 0.8$.

It is shown that SubS performs better than SIS for the case where $n_c = n_s$. This is due to the fact that in SIS the seeds of the Markov chains follow only asymptotically the target distribution and hence a single local move is not sufficient for the chains to converge to their stationary distribution. On the other hand, in SubS the seeds follow the target distribution, so convergence of the chains is not an issue. Using less number of chains with longer period, $n_c = 0.1n_s$, in SIS

Table 1: Coefficient of variation of probability estimate for Example 1.

Number of samples per level n_s	CV of estimate $\delta_{\hat{P}_f}$			
	SubS	SIS ($n_c = n_s$)	SIS ($n_c = 0.1n_s$)	SIS ($n_c = 0.1n_s$) Burn-in period$= 5$
500	0.27	0.37	0.27	0.25
1000	0.19	0.29	0.19	0.16
2000	0.14	0.19	0.13	0.11

gives comparable results with SubS, because the transient effect is reduced since the chains are run longer. In addition, applying an initial burn-in period gives better results than SubS at the expense of additional computational time.

3.2 *Series system reliability problem*

The second example is a series system reliability problem, defined by the following limit-state function at the **U**-space (Waarts 2000):

$$G_2(\mathbf{u}) = \min \left\{ \begin{array}{c} 0.1(u_1 - u_2)^2 - (u_1 + u_2)/\sqrt{2} + 3 \\ 0.1(u_1 - u_2)^2 + (u_1 + u_2)/\sqrt{2} + 3 \\ u_1 - u_2 + 7\sqrt{2} \\ u_2 - u_1 + 7\sqrt{2} \end{array} \right\} \tag{23}$$

The corresponding probability of failure is computed as 2.2×10^{-3}. Table 2 compares the CV of the probability estimates obtained by SIS with the different MCMC settings discussed in Section 3.1. It is shown that SubS performs better than SIS in the case where $n_c = n_s$. SIS performs slightly better than SubS for the case where $n_c = 0.1n_s$, while discarding the initial samples further improves the performance of SIS. Figure 2 shows the optimal IS density of the limit-state function of Equation 23 and samples from the approximate optimal IS density obtained with SIS for $n_s = 1000$ and $n_c = 0.1n_s$. It is demonstrated that the method succeeds to sample accurately all four important failure regions.

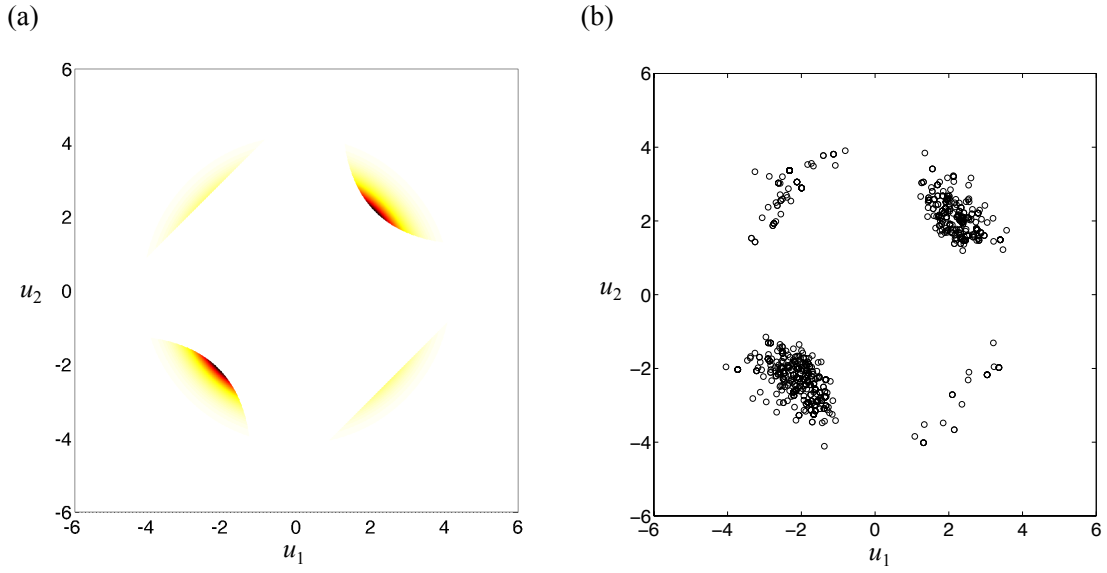

Figure 2: (a) Optimal IS density for Example 2; (b) Samples of approximate optimal IS density obtained from SIS.

Table 2: Coefficient of variation of probability estimate for Example 2.

Number of samples per level n_s	CV of estimate $\delta_{\hat{P}_f}$			
	SubS	SIS ($n_c = n_s$)	SIS ($n_c = 0.1n_s$)	SIS ($n_c = 0.1n_s$) Burn-in period$= 5$
500	0.30	0.53	0.28	0.23
1000	0.23	0.49	0.21	0.17
2000	0.16	0.25	0.14	0.11

3.3 *Linear limit-state function in high dimensions*

This example consist of limit-state function expressed as a linear function of independent standard normal random variables (Engelund and Rackwitz 1993):

$$G_3(\mathbf{u}) = -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} u_i + \beta \tag{24}$$

The probability of failure for this limit-state function is $\Phi(-\beta)$ independent of the dimension n. We choose $\beta = 3.5$ which corresponds to a probability of failure of 2.33×10^{-4}. We use this example to demonstrate the performance of the proposed CSM-H algorithm for varying random dimension n. Figure 3 compares the CV of the probability estimate obtained by SIS with CSM-H to the one by SIS with the standard M-H algorithm. For CSM-H, the correlation parameter is chosen as $\rho = 0.8$ while for M-H, the proposal PDF is chosen as the independent standard normal PDF centered at the current state. For both algorithms, we use $n_s = 1000$ samples per level and $n_c = 100$ chains with length of 10. It is shown that the CV obtained when applying the standard M-H increases with increase of the dimension n. On the other hand, the behavior of the proposed CSM-H algorithm is not influenced by the random dimension.

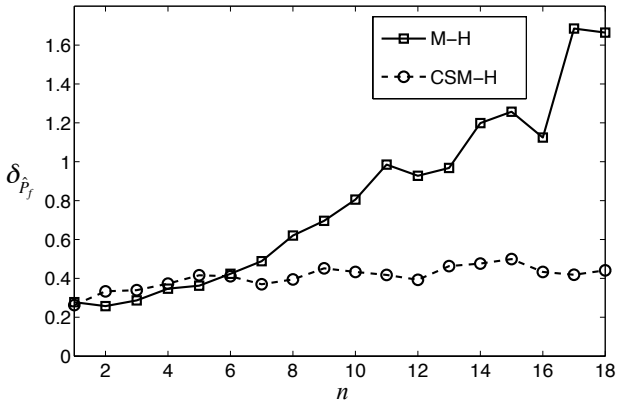

Figure 3: Coefficient of variation of probability estimate $\delta_{\hat{P}_f}$ against random dimension n for Example 3.

4 CONCLUSION

This paper presented a SIS method for structural reliability. The method is based on sampling a sequence of distributions that gradually approach the optimal IS density. Samples from each distribution are obtained through resampling weighted samples from the previous distribution in the sequence and moving the resulting samples applying MCMC. For the MCMC step, a conditional sampling Metropolis-Hastings algorithm is proposed that is suitable for application to high dimensional problems. The method is compared with SubS and it is shown that SIS performs at least as good as SubS in the case where the same number of chains is used for the MCMC sampling for both methods.

REFERENCES

Au, S. K. and J. L. Beck (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilist. Eng. Mech. 15*(4), 263–277.

Au, S. K. and J. L. Beck (2003). Importance sampling in high dimensions. *Struct. Saf. 25*, 139–163.

Beaurepaire, P., H. A. Jensen, G. I. Schuëller, and M. A. Valdebenito (2013). Reliability-based optimization using bridge importance sampling. *Probabilist. Eng. Mech. 34*, 48–57.

Bucher, C. (1988). Adaptive sampling - an iterative fast monte carlo procedure. *Struct. Saf. 5*, 119–126.

Ching, J. and Y.-C. Chen (2007). Transitional markov chain monte carlo method for bayesian model updating, model class selection, and model averaging. *J. Eng. Mech.-ASCE 133*(7), 816–832.

Chopin, N. (2002). A sequential particel filter method for static models. *Biometrika 89*(3), 539–551.

Del Moral, P., A. Doucet, and A. Jasra (2006). Sequential monte carlo samplers. *J. R. Stat. Soc. Series B Stat. Methodol. 68*(3), 411–436.

Der Kiureghian, A. and P. L. Liu (1986). Structural reliability under incomplete probability information. *J. Eng. Mech.-ASCE 112*(1), 85–104.

Ditlevsen, O. and H. O. Madsen (1996). *Structural Reliability Methods*. Chichester: Wiley.

Doucet, A., N. De Freitas, and N. J. Gordon (2001). *Sequential Monte Carlo Methods in Practice*. New York: Springer-Verlag.

Engelund, S. and R. Rackwitz (1993). A benchmark study on importance sampling techniques in structural reliability. *Struct. Saf. 12*(4), 255–276.

Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika 57*(1), 97–109.

Hohenbichler, M. and R. Rackwitz (1981). Non-normal dependent vectors in structural safety. *J. Eng. Mech. Div.-ASCE 107*(6), 1227–1238.

Hohenbichler, M. and R. Rackwitz (1988). Improvements of second-order reliability estimates by importance sampling. *J. Eng. Mech.-ASCE 114*, 2195–2199.

Katafygiotis, L. S. and K. M. Zuev (2007). Estimation of small failure probabilities in high dimensions by adaptive linked importance sampling. In M. Papadrakakis, D. C. Charmpis, N. D. Lagaros, and Y. Tsompanakis (Eds.), *Proc. ECCOMAS thematic conference on computational methods in structural dynamics and earthquake engineering (COMPDYN)*, Rethymno, Crete.

Katafygiotis, L. S. and K. M. Zuev (2008). Geometric insight into the challenges of solving high-dimensional reliability problems. *Probabilist. Eng. Mech. 23*, 208–218.

Koutsourelakis, P. S., H. J. Pradlwarter, and G. I. Schuëller (2004). Reliability of structures in high dimensions, part I: algorithms and applications. *Probabilist. Eng. Mech. 19*, 409–417.

Kurtz, N. and J. Song (2013). Cross-entropy-based adaptive importance sampling using gaussian mixture. *Struct. Saf. 42*(3), 35–44.

Lemaire, M. (2009). *Structural Reliability*. Hoboken, NJ: Wiley.

Neal, R. M. (1998). Regression and classification using gaussian process priors. In J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (Eds.), *Baysian Statistics*, Volume 6, pp. 475–501. Oxford: University Press.

Neal, R. M. (2001). Anneal importance sampling. *Stat. Comput. 11*, 125–139.

Neal, R. M. (2005). Estimating ratios of normalizing constants using linked importance sampling. Technical Report 0511, Department of Statistics, University of Toronto.

Papaioannou, I., W. Betz, K. Zwirglmaier, and D. Straub (2014). MCMC algorithms for subset simulation. Working Paper, Technische Unversität München.

Schuëller, G. I. and R. A. Stix (1987). A critical appraisal of methods to determine failure probabilities. *Struct. Saf. 4*, 293–309.

Spanier, J. and K. B. Oldham (1987). *An Atlas of Functions*. Washington, DC: Hemisphere.

Waarts, P. H. (2000). *Structural reliability using finite elements: an appraisal of DARS: Directional Adaptive Resmponse Surface Sampling*. Ph. D. thesis, TU Delft, Delft, The Netherlands.