FrontUQ 2017

Frontiers of Uncertainty Quantification in Engineering
05 - 08 September 2017
Munich, Germany

Modified Cross Entropy-based Importance Sampling with a Flexible Mixture Model

Sebastian Geyer (s.geyer@tum.de), Iason Papaioannou & Daniel Straub

Engineering Risk Analysis Group
Technische Universität München

Background and Motivation

- Main objective of a reliability analysis is to estimate the failure probability (P_F) of a system. By transforming the random variable space to an equivalent space of standard normal random variables, P_F can be defined as follows:

$$P_F = \int_{\mathbb{R}^n} I(g(\boldsymbol{u}) \leq 0) \cdot \varphi_n(\boldsymbol{u}) \, \mathrm{d}\boldsymbol{u}$$

$\varphi_n(\boldsymbol{u})$ is the n-dimensional standard normal probability density function and $g(\boldsymbol{u})$ is the limit state function whose negative values define failure of the system (Figure 1). The indicator function $I(g(\boldsymbol{u}) \leq 0)$ gives 1 for $g(\boldsymbol{u}) \leq 0$ and 0 else.

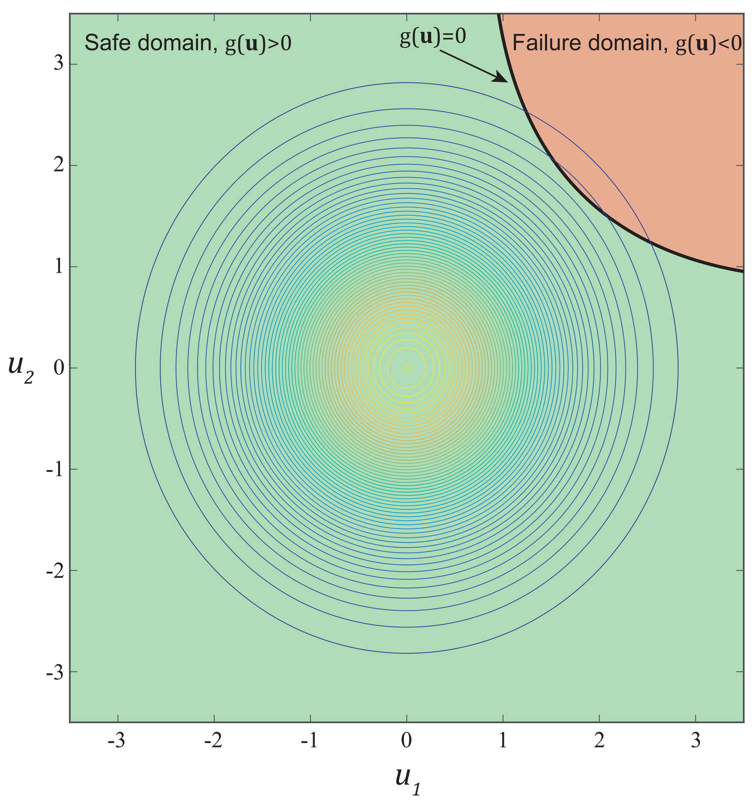

Figure 1: Reliability problem in two-dimensional standard normal space.

- Importance sampling (IS) introduces an alternative sampling density $h(\boldsymbol{u})$ in the problem formulation:

$$P_F = \int_{\mathbb{R}^n} I(g(\boldsymbol{u}) \leq 0) \cdot h(\boldsymbol{u}) \cdot \frac{\varphi_n(\boldsymbol{u})}{h(\boldsymbol{u})} \, \mathrm{d}\boldsymbol{u}$$

With n_s samples generated from $h(\boldsymbol{u})$, the IS estimate of P_F is:

$$\hat{P}_F = \frac{1}{n_s} \sum_{i=1}^{n_s} I(g(\boldsymbol{u}_i) \leq 0) \cdot \frac{\varphi_n(\boldsymbol{u}_i)}{h(\boldsymbol{u}_i)}$$

- An optimal IS density $p^*(\boldsymbol{u})$ can be defined as:

$$p^*(\boldsymbol{u}) = \frac{I(g(\boldsymbol{u}) \leq 0) \cdot \varphi_n(\boldsymbol{u})}{\int_{\mathbb{R}^n} I(g(\boldsymbol{u}) \leq 0) \cdot \varphi_n(\boldsymbol{u}) \, \mathrm{d}\boldsymbol{u}}$$

Since this expression requires knowledge of P_F, it is practically not applicable, but knowledge of the type of $p^*(\boldsymbol{u})$ can be used to identify a near-optimal IS density.

- The cross entropy (CE) method tries to find a near-optimal IS density through minimizing the Kullback-Leibler (KL) divergence between $p^*(\boldsymbol{u})$ and a parametric family of distributions $h(\boldsymbol{u};\boldsymbol{v})$, with $\boldsymbol{v}$ denoting the parameter vector:

$$\text{arg min}_{\boldsymbol{v}} \; D(p^*(\boldsymbol{u}), h(\boldsymbol{u};\boldsymbol{v}))$$

$$= \text{arg max}_{\boldsymbol{v}} \int_{\mathbb{R}^n} I(g(\boldsymbol{u}) \leq 0) \cdot \varphi_n(\boldsymbol{u}) \cdot \ln\left(h(\boldsymbol{u};\boldsymbol{v})\right) \mathrm{d}\boldsymbol{u}$$

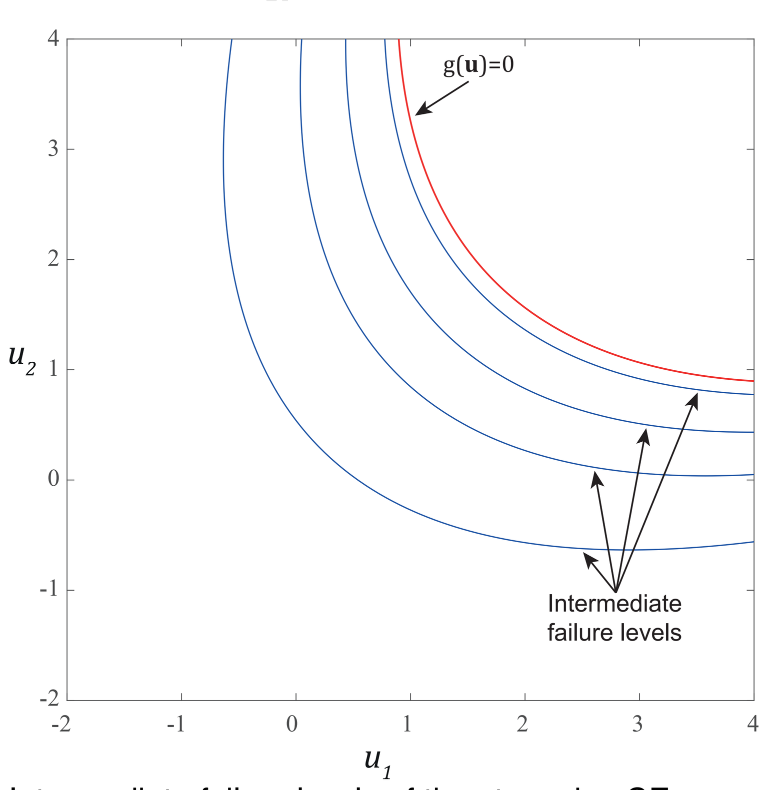

Figure 2: Intermediate failure levels of the step-wise CE approach.

- In each step, an intermediate failure domain is defined (Figure 2), such that $\rho \cdot n_s$, $\rho \in]0,1[$ of the n_s generated samples fall into the failure domain. The parameter update is done through minimizing the KL divergence between the optimal IS density of the intermediate failure domain and $h(\boldsymbol{u};\boldsymbol{v})$.

- **PROBLEM 1: The well-known distribution families are only suitable for certain types of problems (e.g. low/high dimensions, component/system reliability).**
- **PROBLEM 2: Only $\rho \cdot n_s$ samples are taken into account for the parameter update in the CE method.**

A flexible mixture model

- Transformation of the n-dimensional standard normal space to polar coordinates:

$$U = R \cdot A$$

R: Radius, follows the χ-distribution with n degrees of freedom
A: Unit directional vector, follows the uniform distribution on the n-dimensional unit hypersphere

- In high dimensions, the probability mass of the standard normal distribution concentrates around an 'important ring', with radius $R_{imp} = \sqrt{n}$.
- **The von Mises-Fisher (vMF) distribution**: A directional probability distribution on the n-dimensional hypersphere. Its parameters are the mean direction and the concentration parameter, which describes the concentration around the mean direction.
  $\rightarrow$ **Distribution of $\mathbf{A}$**
- **The Nakagami distribution**: A generalization of the χ-distribution.
  $\rightarrow$ **Distribution of $\mathbf{R}$**
- A mixture model of several of these combined distributions (vMFNM) is generated for problems with multimodal failure.
  $\rightarrow$ **Flexible mixture model to solve different problem settings in low and high dimensions.**

Modified cross entropy method

- Approximation of the indicator function by the following expression:

$$I(g(\boldsymbol{u}) \leq 0) \approx \Phi\left(\frac{-g(\boldsymbol{u})}{\sigma}\right)$$

This holds only if σ is chosen small enough. Φ is the standard normal cumulative distribution function.

- Approximate the indicator function in the optimal IS density sequentially by estimating the parameter σ_l in step l such that the coefficient of variation of the sample weights $W_l(\boldsymbol{u})$ adhere to a target value. The weights are defined as:

$$W_l(\boldsymbol{x}) = \Phi\left(\frac{-g(\boldsymbol{u})}{\sigma_l}\right) \cdot \frac{\varphi_n(\boldsymbol{u})}{h_{l-1}(\boldsymbol{u})}$$

- The initial sampling density is chosen as the standard normal density in polar coordinates.
- Fit the parameters of the vMFNM for the next step with the n_s available samples by an iterative expectation-maximization algorithm which minimizes an IS estimate of the Kullback-Leibler divergence.
  $\rightarrow$ All n_s available samples are taken into account for the parameter update.
- The algorithm exits when a substantial number of samples is located in the failure domain. The failure probability is then estimated via IS, with the density estimated at the final step as the IS density.

Numerical example

- Series system consisting of two linear limit state functions in the n-dimensional standard normal space:

$$g(\mathbf{x}) = \min \begin{Bmatrix} \beta - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i \\ \beta + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i \end{Bmatrix}$$

- $P_F = 2 \cdot \Phi(-\beta)$ independent of the dimension n
  ($\beta = 3.5 : P_F = 4.66 \cdot 10^{-4}$)

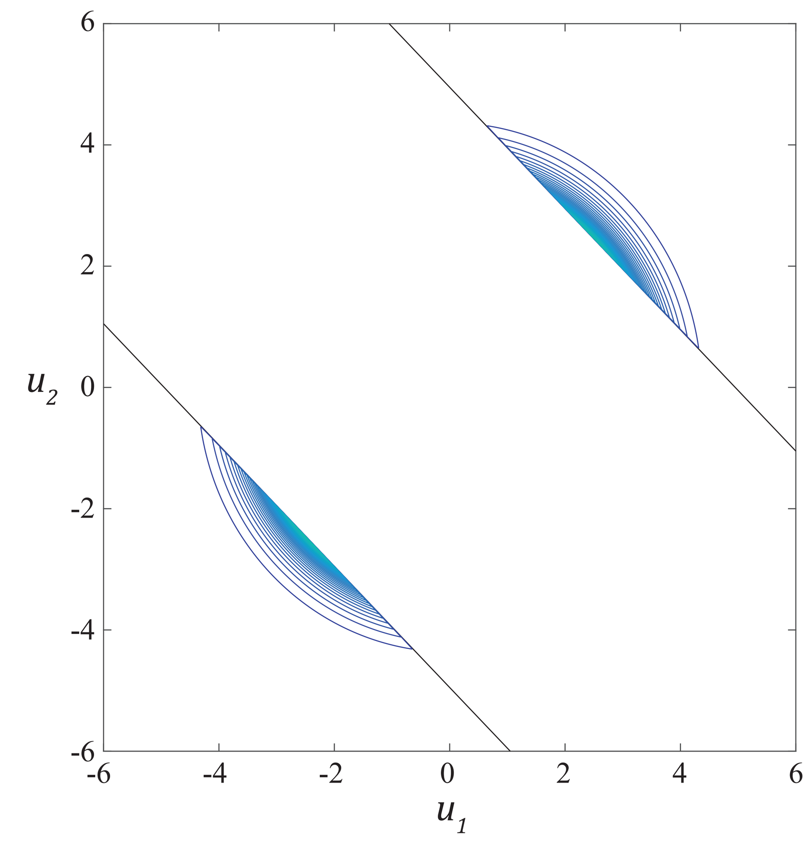

Figure 3: Optimal IS density for the example problem in two dimensions

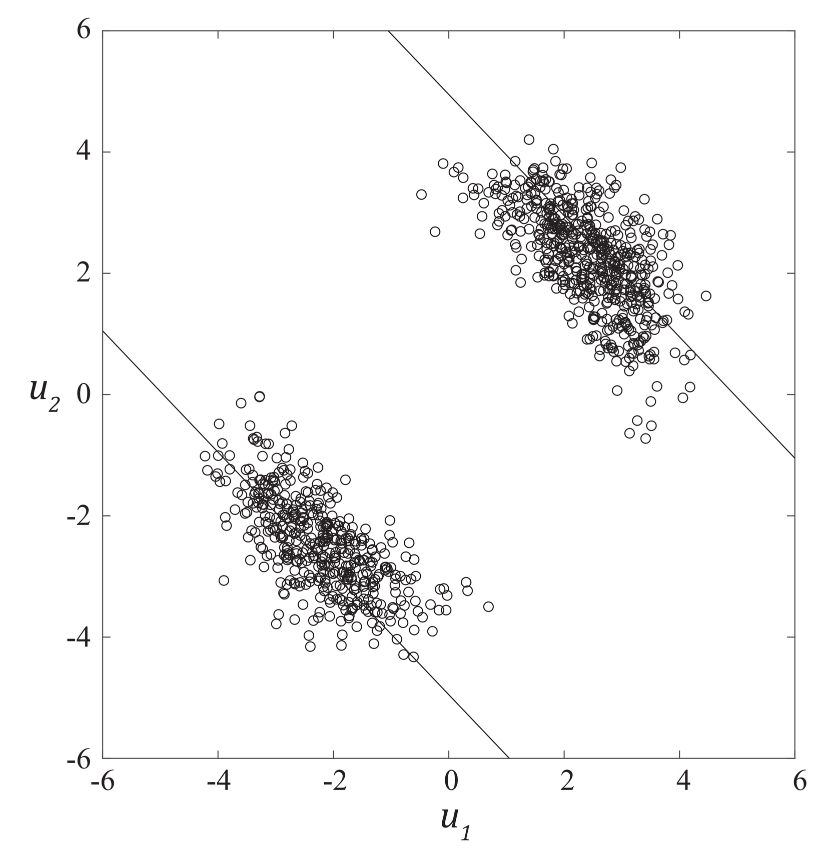

Figure 4: 1000 samples of the near-optimal IS density obtained with a vMFNM with two components. Parameters estimated with $n_s = 1000$ per level.

- The vMFNM with two components successfully detects the two failure domains (Figures 3, 4).
- The modified CE method performs better than the standard CE method in terms of bias and coefficient of variation of the estimate (Figures 5, 6).
- The modified CE method using vMFNM performs well in low to moderately high dimensions. In high dimensions, the number of samples per level has to be increased.

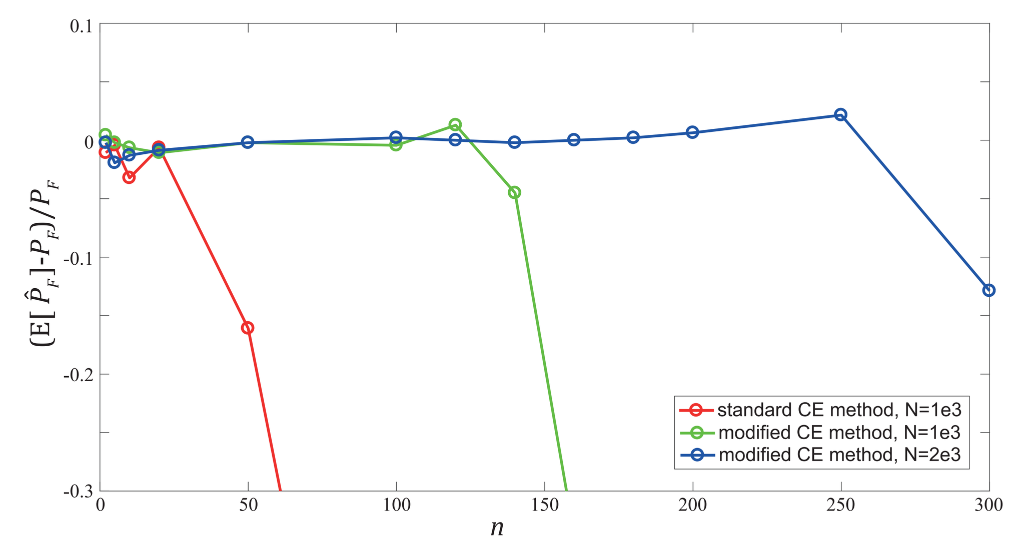

Figure 5: Relative bias of the estimated failure probability as a function of increasing dimensions; Comparison of standard CE method with $n_s = 1000$ per level, modified CE method with $n_s = 1000$ per level and modified CE method with $n_s = 2000$ per level.

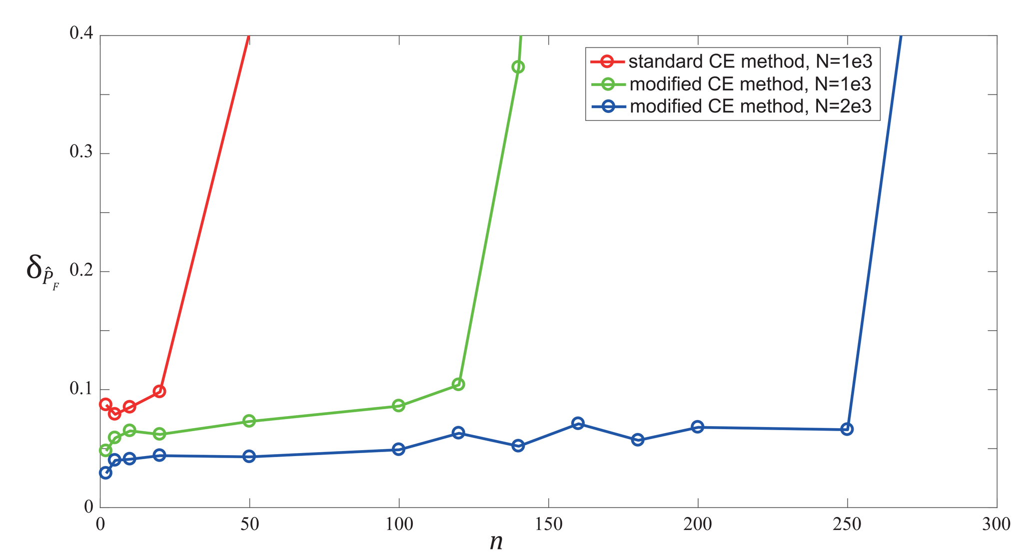

Figure 6: Coefficient of variation of the estimated failure probability as a function of increasing dimensions; Comparison of standard CE method with $n_s = 1000$ per level, modified CE method with $n_s = 1000$ per level and modified CE method with $n_s = 2000$ per level.