Manuscript appeared in Structural Safety
Version September 2020

Combination line sampling for structural reliability analysis

Iason Papaioannou*, Daniel Straub

Engineering Risk Analysis Group, Technische Universität München, 80290 Munich, Germany

*Corresponding author. Email: iason.papaioannou@tum.de

Abstract

Line sampling is a method for efficient estimation of the failure (or rare event) probability. The method generates a set of samples on a hyperplane perpendicular to an important direction that points towards the failure domain, and estimates the probability of failure as a sample mean of one-dimensional probability integrals. The performance of the method strongly depends on the quality of the chosen important direction. Recently, an adaptive approach for adjusting the important direction during the simulation has been proposed, termed advanced line sampling (ALS). This contribution revisits the ALS method and shows that the ALS estimator can be viewed as a combination of estimators, each one corresponding to a direction in the adaptive sequence. We show that the combination implied by the original ALS is suboptimal and propose an alternative combination of estimators. The resulting method is termed combination line sampling (CLS). We demonstrate through three numerical examples that CLS outperforms the ALS estimator, in particular if the initially selected important direction is poor.

Keywords

Reliability analysis, rare event simulation, simulation method, line sampling, linear dynamic systems.

1. Introduction

In reliability analysis, one evaluates the probability of failure of an engineering system or structure. Let $\mathbf{X} = [X_1; X_2; \dots; X_n]$ denote a vector of random variables with joint probability density function (PDF) $f_{\mathbf{X}}(\boldsymbol{x})$, which models randomness and uncertainty in the inputs that influence the system response. The failure event is defined as $F = \{\boldsymbol{x} \in \mathbb{R}^n : g(\boldsymbol{x}) \leq 0\}$, where the function $g(\boldsymbol{x})$ is the so-called limit-state function (LSF) and depends on the outcome of a model of the system response. The probability of failure p_F is defined as:

$$p_F := \Pr(F) = \int_{g(\boldsymbol{x}) \leq 0} f_{\mathbf{X}}(\boldsymbol{x}) d\boldsymbol{x} \,. \tag{1}$$

Evaluation of the integral in Eq. 1 is nontrivial due to the typically low magnitude of p_F and the dependence of $g(\boldsymbol{x})$ on the outcome of an – often computationally intensive – model. A number of tailored approaches have been developed, which are known as structural reliability methods (e.g., Lemaire 2013, Melchers & Beck 2018). These methods can be categorized into approximation and simulation methods. The latter category consists of methods that estimate p_F based on Monte Carlo (MC) samples. Crude MC, which employs samples from $f_{\mathbf{X}}(\boldsymbol{x})$, is inefficient for estimating rare events. This is because its coefficient of variation, measuring the accuracy of the estimator, is approximately $\delta_{MC} \approx (N p_F)^{-0.5}$, implying that for small p_F a large number of samples N is required to obtain a small δ_{MC}. Hence, a number of advanced sampling methods have been developed that aim at reducing the variance of crude MC for a fixed N. These include importance sampling methods (e.g. Bucher 1988, Papaioannou et al. 2016; 2019), subset simulation (Au & Beck 2001b) and line sampling (Hohenbichler & Rackwitz 1988, Koutsourelakis et al. 2004, Schuëller et al. 2004).

Line sampling (LS) estimates p_F through sampling on a hyperplane perpendicular to an important direction pointing towards the failure domain. The method is shown to perform well in problems where the dimension n of the random variable space is large (Pradlwarter et al. 2007). However its performance strongly depends on the choice of the important direction, which is often determined through an initial sampling step or through a single

gradient evaluation of $g(\boldsymbol{x})$. Recently, an adaptive LS approach has been proposed, termed advanced LS (ALS) (de Angelis et al. 2015). This approach adapts the important direction on the fly throughout the simulation and accounts for all the intermediate directions and corresponding samples in the final estimator.

In this paper, we revisit the ALS estimator and show that it can be viewed as a combination of estimators, each one corresponding to an important direction in the adaptive sequence. This interpretation enables one to study the optimality of the ALS estimator. We show that the combination implied by the ALS estimator is suboptimal and propose an alternative combination of estimators that outperforms the ALS estimator. The resulting method is termed combination line sampling (CLS). We study its performance and compare it with ALS by means of three numerical examples; the first involves a convex LSF, the second is a reliability problem whose LSF contains a high-frequency noise term and the third consists of the estimation of the first passage probability of a linear dynamic system with parameter uncertainties.

2. Line sampling (LS)

The LS method samples on the hyperplane perpendicular to an important direction pointing towards the failure region. It was originally developed by Hohenbichler & Rackwitz (1988) for obtaining a correction factor to estimates obtained by the first-order reliability method (FORM) and was generalized by Koutsourelakis et al. (2004) to a stand-alone simulation method. The method is usually applied in a transformed random variable space $\mathbf{U}$ consisting of independent standard normal random variables. It is $\mathbf{U} = \mathbf{T}(\mathbf{X})$ where $\mathbf{T}: \mathbb{R}^n \to \mathbb{R}^n$ is a transformation operator (Hohenbichler & Rackwitz 1981). The probability of failure can be expressed in the $\mathbf{U}$-space as follows

$$p_F = \int_{G(\boldsymbol{u})\leq 0} \varphi_n(\boldsymbol{u}) d\boldsymbol{u}\,, \tag{2}$$

where $G(\boldsymbol{u}) = g[\mathbf{T}^{-1}(\boldsymbol{u})]$ is the LSF in the $\mathbf{U}$-space and $\varphi_n(\boldsymbol{u})$ is the n-dimensional independent standard normal PDF. Let $\boldsymbol{\alpha}$ be a unit row vector that points towards the failure domain. Consider the linear mapping $\mathbf{V} = \mathbf{RU}$, where $\mathbf{R} \in \mathbb{R}^{n\times n}$, with $\mathbf{R}^{\mathrm{T}}\mathbf{R} = \mathbf{I}$, is a suitable rotation matrix whose first row is the unit direction $\boldsymbol{\alpha}$. Due to the rotational

symmetry of the standard normal PDF, any orthogonal transformation of $\mathbf{U}$ will also be independent standard normal. The first component of $\mathbf{V}$, V_1, is a standard normal random variable that is parallel to $\boldsymbol{\alpha}$, whereas the remaining components, $\mathbf{V}_{2:n}$, form an $(n-1)$-dimensional standard normal random vector with outcome space the hyperplane perpendicular to $\boldsymbol{\alpha}$. The probability of failure can be expressed in the rotated space as:

$$\begin{aligned} p_F &= \int_{G_{\mathbf{R}}(\boldsymbol{v})\leq 0} \varphi_n(\boldsymbol{v}) d\boldsymbol{v} \\ &= \int_{\mathbb{R}^{n-1}} \int_{G_{\mathbf{R}}([v_1;\boldsymbol{v}_{2:n}])\leq 0} \varphi(v_1)\, dv_1 \varphi_{n-1}(\boldsymbol{v}_{2:n}) d\boldsymbol{v}_{2:n}\,, \end{aligned} \tag{3}$$

where $G_{\mathbf{R}}(\boldsymbol{v}) = G(\mathbf{R}^{\mathrm{T}}\boldsymbol{v})$ is the LSF in the rotated space. For ease of presentation, we make the following assumption.

Assumption 1. For all $\boldsymbol{v}_{2:n} \in \mathbb{R}^{n-1}$, $d(\boldsymbol{v}_{2:n})$ is the unique solution of $G_{\mathbf{R}}([d; \boldsymbol{v}_{2:n}]) = 0$.

Under Assumption 1, the inner integral in Eq. 3 can be evaluated as $\Phi(-d(\boldsymbol{v}_{2:n}))$, with $\Phi(\cdot)$ denoting the standard normal cumulative distribution function (CDF), such that

$$p_F = \int_{\mathbb{R}^{n-1}} \Phi\big(-d(\boldsymbol{v}_{2:n})\big) \varphi_{n-1}(\boldsymbol{v}) d\boldsymbol{v}_{2:n}\,. \tag{4}$$

Based on Eq. 4, the LS method generates a set of samples $\{\boldsymbol{v}_{2:n}^i, i = 1, \dots, N\}$ from $\mathbf{V}_{2:n}$ and estimates the probability of failure as:

$$p_F \approx p_{LS} = \frac{1}{N}\sum_{i=1}^{N} \Phi\big(-d^i\big). \tag{5}$$

d^i is the solution of $G_{\mathbf{R}}\big([d; \boldsymbol{v}_{2:n}^i]\big) = 0$. Hence, the estimator of Eq. 5 requires solving a line search for every sample $\boldsymbol{v}_{2:n}^i$ to determine d^i, which can be done, e.g., by application of the Newton method or based on a quadratic approximation of $G_{\mathbf{R}}\big([d; \boldsymbol{v}_{2:n}^i]\big)$. This step usually requires only a handful of LSF evaluations. The LS method is illustrated in Fig. 1.

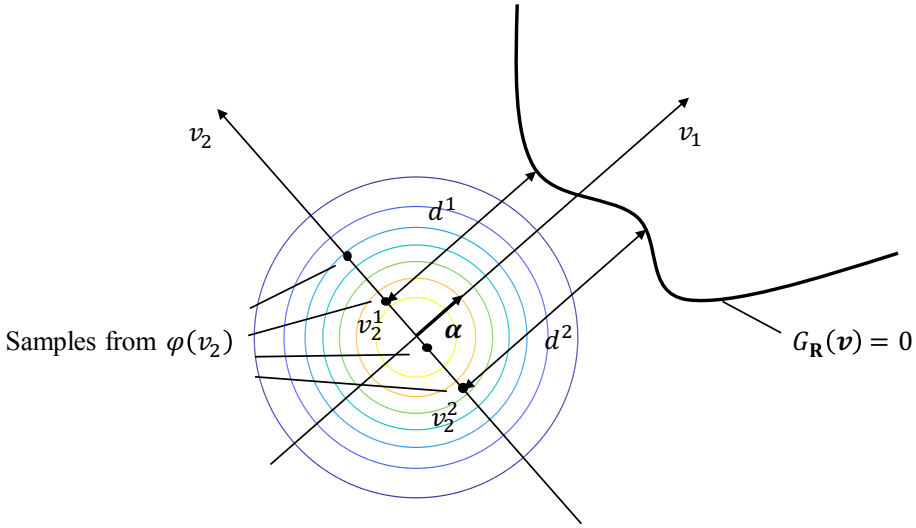

Figure 1: Illustration of the line sampling method.

The LS estimator of Eq. 5 is unbiased, while it can be shown that its variance is smaller or equal to the one of the crude MC estimator with the same number of samples (Koutsourelakis et al. 2004). The variance of the LS estimator strongly depends on the choice of the direction $\boldsymbol{\alpha}$. In (Hohenbichler & Rackwitz 1988) $\boldsymbol{\alpha}$ is chosen as the direction pointing to the FORM design point, aka the most likely failure point. This is the point of the failure domain in **U**-space with smallest distance to the origin. In (Koutsourelakis et al. 2004), different approaches for determining $\boldsymbol{\alpha}$ are investigated, including an MC sampling-based approach. Often the direction is chosen as the gradient of the LSF evaluated at a suitable point in the **U**-space (Pradlwarter et al. 2005, 2007).

Remark 2.1. The rotated space **V** can also be expressed as

$$\mathbf{V} = \mathbf{U}_{\boldsymbol{\alpha}}^{\perp} + V_1 \boldsymbol{\alpha}^{\mathrm{T}}, \tag{6}$$

where $\mathbf{U}_{\boldsymbol{\alpha}}^{\perp}$ is an n-dimensional random vector that defines the projection of **U** on the hyperplane $v_1 = 0$ and can be expressed as

$$\mathbf{U}_{\boldsymbol{\alpha}}^{\perp} = \mathbf{U} - (\mathbf{U}\boldsymbol{\alpha})\boldsymbol{\alpha}^{\mathrm{T}}. \tag{7}$$

Eqs. 6 and 7 offer an alternative means of implementing the LS estimator that avoids evaluation of the projection operator **R**. Through Eq. 7, samples $\boldsymbol{u}_{\boldsymbol{\alpha}}^{\perp,i}$ from $\mathbf{U}_{\boldsymbol{\alpha}}^{\perp}$ can be

generated based on samples from the independent standard normal vector **U** and the corresponding contribution to the LS estimator can be evaluated through solving $G\left(\boldsymbol{u}_{\boldsymbol{\alpha}}^{\perp,i} + d\boldsymbol{\alpha}^{\mathrm{T}}\right) = 0$.

Remark 2.2. If Assumption 1 does not hold, i.e., if $G_{\mathbf{R}}\left(\left[d; \boldsymbol{v}_{2:n}^{i}\right]\right) = 0$ has multiple roots, then the contribution of each sample in Eq. 5 needs to be modified to account for all roots.

Remark 2.3. The presentation of the LS method and its variants in this paper focuses on the case where the failure domain is concentrated in a distinct region of the outcome space, as is the case for component problems with mildly nonlinear LSFs. The efficient treatment of problems with multiple failure modes often requires to account for several important directions, see (Schuëller et al. 2004).

3. Combination line sampling (CLS)

This section introduces the CLS method for reliability analysis. The method builds upon the ALS method, which is described first. We then introduce the CLS estimator as a generalization of the ALS estimator. The CLS estimator is a weighted combination of LS estimators and the optimal choice of the weights is discussed next. The optimal weights cannot be computed in practice, therefore we propose a heuristic selection of the weights.

3.1 Advanced line sampling (ALS)

The ALS method, developed by de Angelis et al. (2015), consists of two enhancements of the standard LS method. The first is to process the samples $\boldsymbol{u}_{\boldsymbol{\alpha}}^{\perp,i}$ such that each new sample is near the current sample and employ the distance d^i as the starting point for the line search algorithm to determine d^{i+1}. This can result in faster convergence of the line search algorithm, which reduces the total number of LSF evaluations. The second enhancement is an adaptation of the important direction whenever a new point on the limit-state (failure) surface is found that is closer to the origin in **U**-space (i.e. a failure point with higher likelihood). This idea allows improving a poor initial choice of the important direction. In the following, we investigate this second enhancement and propose a new estimator, which – based on the same LSF evaluations as the original ALS – results in improved accuracy.

The ALS method applies the following procedure to adapt the important direction. It first chooses an initial direction $\boldsymbol{\alpha}_1$ and evaluates the distance from the origin to the failure surface along this direction, $d_{\boldsymbol{\alpha}_1}$, through solving $G\left(d_{\boldsymbol{\alpha}_1}\boldsymbol{\alpha}_1^{\mathrm{T}}\right) = 0$. At each new sample on the hyperplane perpendicular to the current direction $\boldsymbol{\alpha}_k$, with k denoting a direction counter, it evaluates the distance from the origin to the intersection of the line parallel to $\boldsymbol{\alpha}_k$ with the failure surface, $\left\|\boldsymbol{u}_{\boldsymbol{\alpha}_k}^{\perp,i} + d^i\boldsymbol{\alpha}_k^{\mathrm{T}}\right\|$. If this distance is shorter than $d_{\boldsymbol{\alpha}_k}$, i.e., if $\left\|\boldsymbol{u}_{\boldsymbol{\alpha}_k}^{\perp,i} + d^i\boldsymbol{\alpha}_k^{\mathrm{T}}\right\| < d_{\boldsymbol{\alpha}_k}$, it updates the direction through setting

$$\boldsymbol{\alpha}_{k+1}^{\mathrm{T}} = \frac{\boldsymbol{u}_{\boldsymbol{\alpha}_k}^{\perp,i} + d^i\boldsymbol{\alpha}_k^{\mathrm{T}}}{d_{\boldsymbol{\alpha}_{k+1}}} \text{ with } d_{\boldsymbol{\alpha}_{k+1}} = \left\|\boldsymbol{u}_{\boldsymbol{\alpha}_k}^{\perp,i} + d^i\boldsymbol{\alpha}_k^{\mathrm{T}}\right\|. \tag{8}$$

The sampling then proceeds with this new direction. The final estimator of the probability of failure is given by:

$$p_{ALS} = \frac{1}{N}\sum_{i=1}^{N}\Phi\left(-d^i\right), \tag{9}$$

where d^i is obtained through the solution of $G\left(\boldsymbol{u}_{\boldsymbol{\alpha}_k}^{\perp,i} + d\boldsymbol{\alpha}_k^{\mathrm{T}}\right) = 0$ with $\boldsymbol{\alpha}_k$ denoting the current direction at the time of generation of the i-th sample, as evaluated by the above procedure.

Remark 3.1. The ALS approach can be effective in low- to moderate-dimensional spaces, where the chances of finding a better direction throughout the simulation are good even in cases where the initial direction is poor. As the dimension of the sampling space increases, it becomes more difficult to achieve a substantial improvement. Hence, although the LS method with an appropriately chosen important direction is known to perform well in high dimensions, the potential of ALS (or CLS presented herein) is limited to problems in moderate dimensions.

3.2 ALS as a combination of estimators

The ALS estimator of Eq. 9 can be written as a combination of LS estimators each one corresponding to an intermediate direction. Let p_{LS_k} be the LS estimator corresponding to direction $\boldsymbol{\alpha}_k$. It is

$$p_{LS_k} = \frac{1}{N_k}\sum_{i=1}^{N_k}\Phi\left(-d^i\right), \tag{10}$$

where N_k is the number of samples generated on the hyperplane perpendicular to $\boldsymbol{\alpha}_k$, with $\sum_{k=1}^{K} N_k = N$; K denotes the total number of directions evaluated throughout the adaptive simulation process. The ALS estimator of Eq. 9 is:

$$p_{ALS} = \sum_{k=1}^{K}\frac{N_k}{N}p_{LS_k}. \tag{11}$$

Eq. 11 shows that the original ALS estimator is a linear combination of the estimators p_{LS_k}, each weighted with the number of samples N_k drawn parallel to the corresponding direction $\boldsymbol{\alpha}_k$. This interpretation of the ALS estimator motivates a more general estimator that combines the estimators of Eq. 10 as follows:

$$p_{CLS} = \sum_{k=1}^{K} w_k p_{LS_k}, \tag{12}$$

where $\{w_k, k = 1, \dots, K\}$ are weights satisfying $w_k \geq 0$ and $\sum_{k=1}^{K} w_k = 1$. We term this estimator combination LS (CLS) estimator. Such combination estimators have been studied in the context of adaptive importance sampling (e.g., Owen & Zhou 1999, Owen & Zhou 2019). The estimator p_{ALS} can be retrieved from the general CLS estimator of Eq. 12 by setting $w_k = N_k/N$. However, as we show in the following, this is not the optimal choice of w_k.

3.3 Best linear unbiased combination

The estimator p_{CLS} is unbiased, since

$$\mathrm{E}[p_{CLS}] = \sum_{k=1}^{K} w_k \mathrm{E}\left[p_{LS_k}\right] = \sum_{k=1}^{K} w_k p_F = p_F, \tag{13}$$

where $\mathrm{E}\left[p_{LS_k}\right] = p_F$ because the standard LS estimator with fixed direction is unbiased and we have used the fact that $\sum_{k=1}^{K} w_k = 1$. Moreover, its variance is

$$\mathrm{Var}[p_{CLS}] = \sum_{k=1}^{K} w_k^2 \mathrm{Var}\left[p_{LS_k}\right], \tag{14}$$

where

$$\begin{aligned} \mathrm{Var}\left[p_{LS_k}\right] &= \frac{1}{N_k^2}\sum_{i=1}^{N_k} \mathrm{Var}\left[\Phi\left(-d^i\right)\right] = \frac{N_k \mathrm{Var}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)\right]}{N_k^2} \\ &= \frac{\mathrm{Var}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)\right]}{N_k}, \end{aligned} \tag{15}$$

with $d_{\boldsymbol{\alpha}_k}(\mathbf{U})$ denoting the random variable representing the distance to the failure surface from the hyperplane perpendicular to direction $\boldsymbol{\alpha}_k$. One can show that the optimal weights, i.e., those that minimize the variance of Eq. 14 while ensuring an unbiased estimator p_{CLS}, are given as follows

$$w_k = \frac{\mathrm{Var}\left[p_{LS_k}\right]^{-1}}{\sum_{k=1}^{K} \mathrm{Var}\left[p_{LS_k}\right]^{-1}}. \tag{16}$$

A simple proof of Eq. 16 is given in Appendix A. Inserting Eq. 15 into Eq. 16, the optimal weights are given as

$$w_k = \frac{N_k \mathrm{Var}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)\right]^{-1}}{\sum_{k=1}^{K} N_k \mathrm{Var}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)\right]^{-1}}. \tag{17}$$

Eq. 17 shows that the optimal weight of estimator p_{LS_k} increases with increase of the number of samples N_k and decreases with increase of the variance of $\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)$.

To illustrate the relation between the direction $\boldsymbol{\alpha}_k$ and the variance of $\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)$, we consider the case where the LSF in **U**-space is linear, i.e., it can be expressed as:

$$G(\boldsymbol{u}) = -\boldsymbol{\alpha}\boldsymbol{u} + \beta, \tag{18}$$

where $\boldsymbol{\alpha} \in \mathbb{R}^n$ is a row vector and we assume without loss of generality that $\|\boldsymbol{\alpha}\| = 1$. In this case, it is shown in Appendix B that the variance terms in Eq. 17 can be evaluated through the following expression:

$$\mathrm{Var}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)\right] = \int_0^{1-\boldsymbol{\alpha}\cdot\boldsymbol{\alpha}_k^{\mathrm{T}}} \frac{1}{2\pi\sqrt{1-r^2}} \exp\left(-\frac{\beta^2}{1+r}\right) dr\,. \tag{19}$$

Fig. 2 plots the variance of Eq. 19 as a function of the angle between the direction $\boldsymbol{\alpha}$, which is the optimal direction for the linear problem, and the sampling direction $\boldsymbol{\alpha}_k$. For this problem, choosing $\boldsymbol{\alpha}_k = \boldsymbol{\alpha}$ would result in zero variance and, hence, an infinite weight according to Eq. 17. This is expected, since for the linear problem of Eq. 18, the contribution of the direction $\boldsymbol{\alpha}$ results in the exact probability of failure. As the angle $\cos^{-1}\left(\boldsymbol{\alpha}\cdot\boldsymbol{\alpha}_k^{\mathrm{T}}\right)$ increases, the variance term also increases, which implies that poor choices of the sampling direction should be associated with smaller weights. That is, larger weights should be given to those p_{LS_k} that are associated with better directions $\boldsymbol{\alpha}_k$, which are the latter directions in the adaptive process. In particular, the initial directions can be quite poor and hence only a smaller w_k should be assigned to them. The weights in the original ALS method are chosen proportional to N_k and are suboptimal, since they do not account for the influence of direction $\boldsymbol{\alpha}_k$ on the quality of each estimator p_{LS_k}.

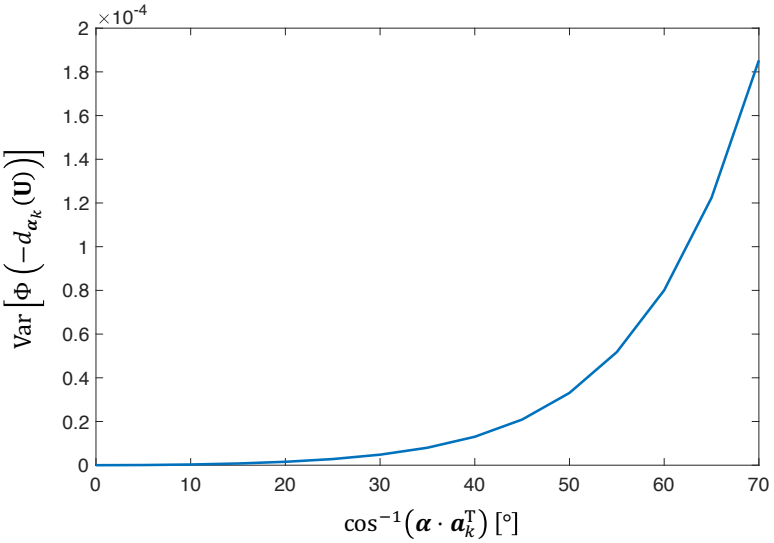

Figure 2: Influence of the sampling direction $\boldsymbol{\alpha}_k$ on the variance term $\mathrm{Var}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)\right]$ entering the optimal weights of Eq. 17, evaluated for a linear LSF in $\mathbf{U}$-space with $\beta = 3$ ($p_F = 1.3 \times 10^{-3}$).

Unfortunately, the weights of Eq. 17 cannot be used in practice because in the general case $\mathrm{Var}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)\right]$ is unknown. One possibility would be to use estimators of $\mathrm{Var}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)\right]$ based on the samples d^i of $d_{\boldsymbol{\alpha}_k}(\mathbf{U})$ generated throughout the simulation. However, this is not advisable due to several reasons; see also the discussion in the general context of linear combinations of estimators in (Owen & Zhou 2019). The main reason is that the number of samples for each estimator N_k is often quite small, since the algorithm proceeds to the next direction once it finds a more likely point on the failure surface. In this case, the sample estimator of $\mathrm{Var}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)\right]$ will evidently be poor. Moreover, the estimator p_{LS_k} and the estimator of $\mathrm{Var}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)\right]$ are correlated; smaller values of p_{LS_k} are associated with smaller values of the $\mathrm{Var}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)\right]$ estimator. As a result, p_{LS_k} and the weights w_k of Eq. 17 are negatively correlated, since these weights are proportional to $\mathrm{Var}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)\right]^{-1}$. This implies that small p_{LS_k} will tend to take large weights while large p_{LS_k} will tend to take small weights. This will lead on average to an underestimation of the probability of failure; in other words, to a negatively biased estimator. To avoid this, we introduce a practical expression for the weights that does not require estimation of a statistical quantity in the following.

3.4 A heuristic definition of the weights

The variance terms entering the optimal weights of Eq. 17 will be large if the corresponding direction differs significantly from the optimal direction (cf. Fig. 2). This will be the case for estimators based on directions $\boldsymbol{\alpha}_k$ along which the distance to the failure surface $d_{\boldsymbol{\alpha}_k}$ is large and the corresponding failure integral $\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{0})\right) = \Phi\left(-d_{\boldsymbol{\alpha}_k}\right)$ is small. In other words, the theoretically optimal weight of the estimator p_{LS_k} will be larger when the probability integral $\Phi\left(-d_{\boldsymbol{\alpha}_k}\right)$ is larger. Based on this rationale, we define the following heuristic weights:

$$w_k = \frac{N_k \Phi\left(-d_{\boldsymbol{\alpha}_k}\right)}{\sum_{k=1}^{K} N_k \Phi\left(-d_{\boldsymbol{\alpha}_k}\right)}. \tag{20}$$

The weights of Eq. 20 ensure that among estimators with the same number of samples N_k, the ones corresponding to directions that point to failure points with higher likelihood (smaller distance to the origin) will have higher contribution to the probability of failure. We remark that a similar approach to Eq. 20 has been used in (Schuëller & Stix 1987) for choosing the weights in the context of mixture importance sampling for reliability problems with multiple failure modes.

The weights of Eq. 20 can be evaluated based on the same LSF evaluations as for the ALS method. In fact, the adaptive sampling process of ALS does not need to be modified. The only difference between CLS and ALS is the final estimate of the probability failure, which in CLS is based on the combination estimator of Eq. 12 with weights according to Eq. 20. To compute these weights, the CLS algorithm needs to keep track of the intermediate directions, which is not necessary in ALS. We summarize the algorithm in Algorithm 1.

Remark 3.2. The first enhancement of the ALS method, which can improve the convergence of the line search required for each sample, can be included in the CLS algorithm through modifying the sequence of processing the samples; see (de Angelis et al. 2015) for details.

4. Numerical examples

We investigate the performance of the CLS estimator with the proposed weights of Eq. 20 with three numerical examples. The first consists of a convex LSF and serves as an illustration on the influence of the choice of the initial direction. The second is a reliability problem whose LSF contains a high-frequency noise term. The third example is an application of the approach to the estimation of the first passage probability of linear systems with parameter uncertainties.

Algorithm 1: CLS algorithm

Input: Initial direction $\boldsymbol{\alpha}_1$; sample size N

Initialization: Solve $G\left(d_{\boldsymbol{\alpha}_1}\boldsymbol{\alpha}_1^{\mathrm{T}}\right)=0$ for $d_{\boldsymbol{\alpha}_1}$; set $k \leftarrow 1$

Set $i \leftarrow 0$

for $l \leftarrow 1, N$ **do**

$i \leftarrow i+1$

$\boldsymbol{u} \sim \varphi_n(\boldsymbol{u})$

$\boldsymbol{u}_{\boldsymbol{\alpha}_k}^{\perp,i} \leftarrow \boldsymbol{u}-(\boldsymbol{u}\boldsymbol{\alpha}_k)\boldsymbol{\alpha}_k^{\mathrm{T}}$

Solve $G\left(\boldsymbol{u}_{\boldsymbol{\alpha}_k}^{\perp,i}+d^i\boldsymbol{\alpha}_k^{\mathrm{T}}\right)=0$ for d^i

$c \leftarrow \left\|\boldsymbol{u}_{\boldsymbol{\alpha}_k}^{\perp,i}+d^i\boldsymbol{\alpha}_k^{\mathrm{T}}\right\|$

if $c < d_{\boldsymbol{\alpha}_k}$ **or** $l = N$ **then**

$N_k \leftarrow i$

Evaluate p_{LS_k} with Eq. 10

if $l = N$ **then**

$K \leftarrow k$

else

$\boldsymbol{\alpha}_{k+1}^{\mathrm{T}} \leftarrow \left(\boldsymbol{u}_{\boldsymbol{\alpha}_k}^{\perp,i}+d^i\boldsymbol{\alpha}_k^{\mathrm{T}}\right)/c$

$d_{\boldsymbol{\alpha}_{k+1}} \leftarrow c$

$k \leftarrow k+1$

$i \leftarrow 0$

end if

end if

end for

Evaluate the weights $\{w_k, k=1,\ldots,K\}$ with Eq. 20

Evaluate the estimator p_{CLS} with Eq. 12

return p_{CLS}

4.1 Convex limit-state function

The LSF of the first example is given by the following quadratic function of two standard normal random variables:

$$G(\boldsymbol{u}) = 0.1(u_1 - u_2)^2 - \frac{1}{\sqrt{2}}(u_1 + u_2) + 2.5\,. \tag{21}$$

The reference probability of failure is $p_F = 4.21 \times 10^{-3}$. The design point of this problem lies on the $45°$ line, such that the optimal direction for application of LS is $\boldsymbol{\alpha} = [1/\sqrt{2}\,; 1/\sqrt{2}]$.

To illustrate the performance of the ALS and CLS estimators, we purposely choose suboptimal initial directions. We study two initial directions: $\boldsymbol{\alpha}_1^{(1)} = [0; 1]$ and $\boldsymbol{\alpha}_1^{(2)} \propto [0.5; 1]$, shown in Fig. 3. Table 1 shows the coefficient of variation of the two estimators obtained with 500 repeated simulation runs for the two choices of the initial directions and different sample sizes N. We see that for direction $\boldsymbol{\alpha}_1^{(1)}$, which differs significantly from the optimal one, the CLS estimator outperforms the standard ALS estimator. This is because the LS estimators based on directions closer to the optimal one, found throughout the course of the simulation, receive higher weights in the final probability estimate in CLS estimator than in ALS. For $\boldsymbol{\alpha}_1^{(2)}$**,** we observe only a small improvement by the CLS over the ALS estimator, since $\boldsymbol{\alpha}_1^{(2)}$is close to the optimal direction and, hence, the influence of the modified weights is small.

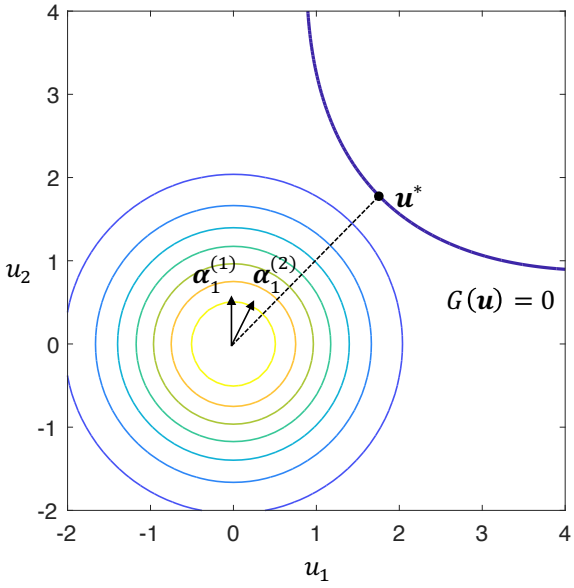

Figure 3: Convex limit-state surface and chosen initial directions $\boldsymbol{\alpha}_1^{(1)}$ and $\boldsymbol{\alpha}_1^{(2)}$. $\boldsymbol{u}^*$ denotes the design point (most likely failure point).

Table 1. Coefficient of variation of the probability estimate for the convex LSF example obtained with the ALS and CLS estimators with different initial directions.

Sample size N	$\boldsymbol{\alpha}_1^{(1)} = [0; 1]$		$\boldsymbol{\alpha}_1^{(2)} \propto [0.5; 1]$	
	ALS	CLS	ALS	CLS
20	0.46	0.26	0.14	0.12
30	0.30	0.13	0.10	0.09
50	0.21	0.07	0.07	0.06

4.2 Noisy limit-state function

We consider a LSF with a high-frequency noise term:

$$g(\boldsymbol{x}) = x_1 + 2x_2 + 3x_3 + x_4 - 5x_5 - 5x_6 + 0.1\sum_{i=1}^{4} sin(100x_i). \qquad (22)$$

The random variables in this function are statistically independent and have lognormal marginal distributions. The variables X_1 to X_4 have means 120 and standard deviations 8; X_5 has mean 50 and standard deviation 10 and X_6 has mean 40 and standard deviation 8.

This example is modified from an example in (Liu & Der Kiureghian 1991): the standard deviation of the variables are reduced to result in a smaller failure probability and the noise is amplified, while it is a function of only variables X_1 to X_4. The latter variables could represent the capacity of a system and the noise could be due to the dependence of the capacity on the outcome of a nonlinear numerical solver. The reference probability of failure is $p_F = 5.29 \times 10^{-4}$. We choose the initial direction as the negative normalized gradient of the LSF evaluated at the origin in standard normal space, which gives $\boldsymbol{\alpha}_1 = [0.35; 0.24; 0.24; 0.35; 0.63; 0.50]$. We see that the noise term changes the signs of the first four components of $\boldsymbol{\alpha}_1$ – the signs are positive although variables X_1 to X_4 are of the capacity type, and amplifies their contribution; this implies that the choice of $\boldsymbol{\alpha}_1$ is far from optimal.

Table 2 shows the coefficient of variation of the estimators obtained with 500 repeated simulation runs with the standard LS method, the ALS method and the proposed CLS estimator with weights according to Eq. 20 for different sample sizes N. As expected, both ALS and CLS significantly outperform the basic LS method with fixed direction $\boldsymbol{\alpha}_1$, as they both adapt the direction on the fly – we note that the final choice of the direction of a run with $N = 150$ is $\boldsymbol{\alpha} = [-0.01; -0.03; -0.23; -0.10; 0.86; 0.44]$, which differs significantly from $\boldsymbol{\alpha}_1$. The CLS estimator results in lower coefficient of variation than the ALS estimator, because it assigns higher weights to the samples corresponding to the later directions in the sequence.

Table 2. Coefficient of variation of the probability estimate for the noisy LSF example obtained with the LS, ALS and CLS estimators.

Sample size N	LS	ALS	CLS
80	1.02	0.50	0.21
100	0.96	0.32	0.19
150	0.84	0.21	0.14

4.3 SDOF oscillator

We consider a single degree of freedom (SDOF) oscillator subjected to Gaussian white noise excitation:

$$\ddot{x}(t) + 2\eta\omega\dot{x}(t) + \omega^2 x(t) = f(t)\,, \tag{23}$$

where ω is the natural frequency and η the damping ratio. The system is assumed to start at rest, i.e., $x(0) = 0$ and $\dot{x}(0) = 0$. The parameters $\boldsymbol{\Theta} = [\omega, \eta]$ are modeled as independent lognormal random variables; ω has mean 2π rad/s and standard deviation 0.2π rad/s and η has mean 0.05 and standard deviation 0.005. The stochastic excitation $f(t)$ has duration $T = 20$s and intensity $I = 1\,\mathrm{m}^2/\mathrm{s}^3$, and is discretized at time intervals of $\Delta t = 0.01$s with $n_t = 2001$ independent standard normal random variables gathered in the vector $\mathbf{U}_f = [U_{f,1}; \ldots; U_{f,n_t}]$, such that $f(t_i) = \sqrt{I/\Delta t}\, U_{f,i}$ for $t_i = (i-1)\Delta t$. We are interested in evaluating the probability of the first passage failure event, defined as:

$$F = \left\{ [\boldsymbol{\Theta}, \mathbf{U}_f] \in \mathbb{R}^{n_t+2} : \max_{i=1,\ldots,n_t} x(t_i, \boldsymbol{\Theta}, \mathbf{U}_f) \geq x^* \right\}. \tag{24}$$

The first passage probability can be expressed as

$$\Pr(F) = \int_{\boldsymbol{\theta}\in\mathbb{R}^2} p_{F|\boldsymbol{\Theta}}(\boldsymbol{\theta}) f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta}\,, \tag{25}$$

where $p_{F|\boldsymbol{\Theta}}(\boldsymbol{\theta}) = \Pr(F|\boldsymbol{\Theta} = \boldsymbol{\theta})$ is the conditional probability of F given a particular outcome $\boldsymbol{\theta}$ of the system parameters and $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$ is the joint PDF of $\boldsymbol{\Theta}$. The probability $p_{F|\boldsymbol{\Theta}}(\boldsymbol{\theta})$ consists of a first passage probability of a deterministic linear system subjected to Gaussian process excitation and can be evaluated efficiently by application of the importance sampling approach of (Au & Beck 2001a). Here, we apply the ALS and CLS methods to evaluate the integral of Eq. 25 with $p_{F|\boldsymbol{\Theta}}(\boldsymbol{\theta})$ evaluated with the method of (Au & Beck 2001a). To apply this approach, we need to transform Eq. 25 to the form of the integral of Eq. 2. This can be done by defining the following LSF (Wen and Chen 1987):

$$g(\boldsymbol{\theta}, z) = z - p_{F|\boldsymbol{\Theta}}(\boldsymbol{\theta})\,, \tag{26}$$

where z is the outcome of an auxiliary standard uniform random variable Z. Eq. 26 can be expressed in the standard normal space $\mathbf{U} = [\mathbf{U}_{\boldsymbol{\theta}}, U_z]$ as:

$$G(\boldsymbol{u}_{\boldsymbol{\theta}}, u_z) = u_z - \Phi^{-1}\left[p_{F|\boldsymbol{\Theta}}\left(\mathbf{T}^{-1}(\boldsymbol{u}_{\boldsymbol{\theta}})\right)\right]. \quad (27)$$

For an application of LS to this type of problems that does not require a transformation of Eq. 25, we refer to (Pradlwarter & Schuëller 2010). The threshold in Eq. 24 is taken as $x^* = 0.9$, which results in a probability of 4.08×10^{-5}, as evaluated by crude MC with 10^8 samples. The number of samples for evaluating the conditional probability $p_{F|\boldsymbol{\Theta}}(\boldsymbol{\theta})$ is taken as 20, which is reported to be an adequate sample size for obtaining a coefficient of variation in the order of 0.2 (Au and Beck 2001a). The initial direction is chosen based on a gradient evaluation of the LSF of Eq. 27 at the origin of the **U**-space, whereby the gradient of $p_{F|\boldsymbol{\Theta}}(\boldsymbol{\theta})$ is estimated with a forward difference scheme with a step of 0.05, which gives $\boldsymbol{\alpha}_1 = [-0.573; 0.397; -0.717]$. Since the estimator of $p_{F|\boldsymbol{\Theta}}(\boldsymbol{\theta})$ is noisy, the direction evaluated with this scheme is likely to be suboptimal. The line search required for each sample on the hyperplane is solved using a quadratic fit of the LSF at the line perpendicular to the sample using 5 LSF evaluations; this allows coping with the noise of the $p_{F|\boldsymbol{\Theta}}(\boldsymbol{\theta})$ estimator.

Fig. 4 shows the probability estimates obtained with the ALS and CLS estimators for 50 independent simulation runs with two different sample sizes N, where the two estimators are computed using the same samples and corresponding sequence of directions. It is shown that the CLS estimator with the weights of Eq. 20 has significantly lower variability than the standard ALS estimator. In some cases, the ALS estimator severely overestimates the first-passage probability. The reason is that the initial direction differs significantly from the optimal one and thus the initial LS estimators in the combination have large variance. The ALS estimator weighs the LS estimators according to their respective number of samples. Hence, when the total number of samples is small, all LS estimators receive significant weights. Overestimation occurs if a sample drawn perpendicular to a suboptimal direction results in small d^i and, hence, large $\Phi\left(-d^i\right)$. The CLS estimator penalizes the directions that point to failure points with large distance to the origin. Hence, the contribution of the initial LS estimators to the final probability estimate is small in this case. In essence, the weights of Eq. 20 act as a correction to the standard ALS estimator as is illustrated in Fig. 4.

Table 3 shows the 10- and 90-percentiles of the first-passage probability estimates obtained from 50 independent runs with the two estimators and different sample sizes N. For all N, the intervals are significantly tighter for the CLS estimator. We note that both estimators are based on the same samples and corresponding LSF evaluation and thus have the same computational cost.

Table 3. 10- and 90-percentiles of the probability estimates for the SDOF oscillator obtained from 50 independent simulation runs with the ALS and CLS estimators. The reference value of the probability of failure is 4.08×10^{-5}

Sample size N	ALS	CLS
	$[p_{10}, p_{90}]$	$[p_{10}, p_{90}]$
20	$[2.09,5.70] \times 10^{-5}$	$[2.54,5.21] \times 10^{-5}$
30	$[2.91,5.19] \times 10^{-5}$	$[3.31,4.41] \times 10^{-5}$
50	$[2.92,4.68] \times 10^{-5}$	$[3.47,4.40] \times 10^{-5}$

5. Concluding remarks

This paper revisits the advanced line sampling (ALS) method for reliability assessment. We show that the ALS estimator can be interpreted as a special case of a combination of line sampling estimators corresponding to important directions chosen adaptively throughout the simulation. We introduce an alternative combination that penalizes directions pointing to failure points with small likelihood. The proposed approach, termed combination line sampling (CLS), is based on the ALS algorithm and only modifies the final estimator, without requiring additional limit-state function evaluations. Three numerical examples demonstrate that the proposed modification improves significantly the performance of the method, especially in cases where the initial direction is poorly chosen.

In the numerical examples, the different variants of the LS method are applied in low- to moderate-dimensional parameter spaces. The standard LS method is known to perform well in high-dimensional problems, provided that an appropriate important direction is chosen. However, the potential of improving the initial direction through an adaptive

process decreases as the dimension of the space increases. Hence, the added value of both ALS and CLS is restricted to moderate dimensions.

The focus of this paper is component problems with mildly nonlinear limit states. Highly nonlinear limit-state functions and series system problems with multiple failure modes can be treated with LS using several important directions (Schuëller et al. 2004). This approach could potentially be combined with an adaptation of the important directions. The contribution of each mode to the probability of failure could then be evaluated with the CLS estimator. Extensions of the proposed method that could address such problems are left to future studies.

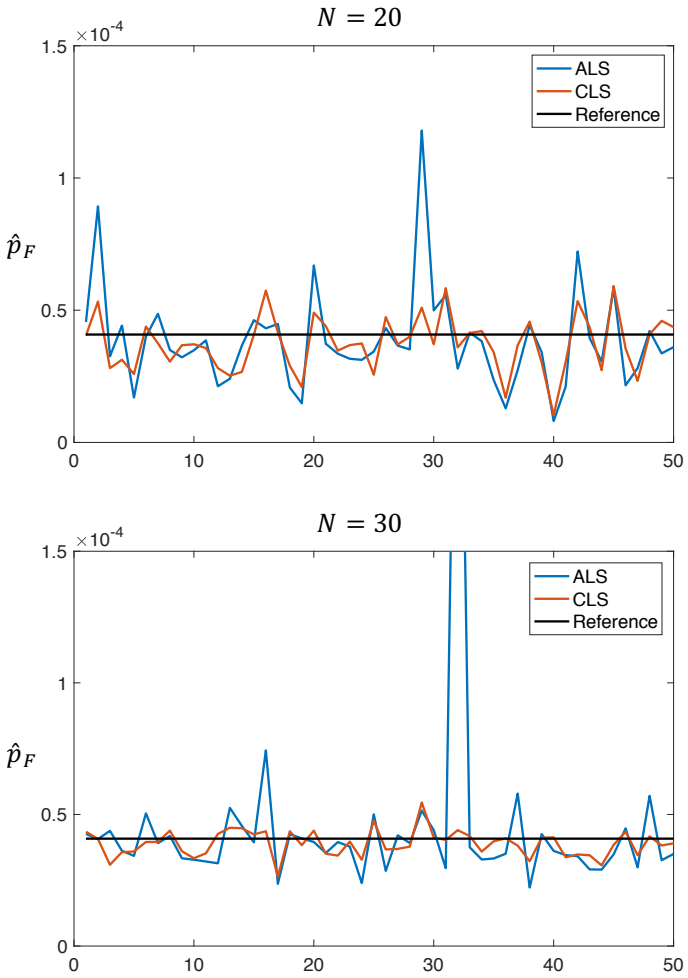

Figure 4: 50 independent estimates for the SDOF oscillator obtained with the ALS and CLS estimators using the same samples.

Appendix A

In this Appendix, we show that the optimal weights for the CLS estimator of Eq. 12, i.e. the weights that minimize its variance while ensuring an unbiased estimator, are given by the weights of Eq. 16. Through Eq. 13, it is easy to see that the necessary constraint for obtaining an unbiased estimator is that the weights sum up to one, i.e. that $\sum_{k=1}^{K} w_k = 1$. Therefore, the optimization problem can be stated as follows:

$$\begin{aligned} &\text{minimize } \sum_{k=1}^{K} w_k^2 \mathrm{Var}\left[p_{LS_k}\right] , \\ &\text{subject to: } \sum_{k=1}^{K} w_k = 1 . \end{aligned} \tag{28}$$

The program of Eq. 28 can be solved by the method of Lagrange multipliers. The Lagrangian of the problem of Eq. 28 is expressed as

$$\mathcal{L}(\boldsymbol{w}, \lambda) = \sum_{k=1}^{K} w_k^2 \mathrm{Var}\left[p_{LS_k}\right] + \lambda \left(\sum_{k=1}^{K} w_k - 1 \right), \tag{29}$$

where $\boldsymbol{w} = [w_1, \dots, w_K]$ collects the weights and λ is the Lagrange multiplier. The first-order necessary conditions read:

$$\frac{\partial \mathcal{L}(\boldsymbol{w}, \lambda)}{\partial w_k} = 0, k = 1, \dots, K; \frac{\partial \mathcal{L}(\boldsymbol{w}, \lambda)}{\partial \lambda} = 0 . \tag{30}$$

The first condition gives:

$$w_k = -\frac{\lambda}{2} \mathrm{Var}\left[p_{LS_k}\right]^{-1} , k = 1, \dots, K . \tag{31}$$

Substitution of Eq. 31 to the second condition of Eq. 30 gives

$$\sum_{k=1}^{K} -\frac{\lambda}{2} \mathrm{Var}\left[p_{LS_k}\right]^{-1} = 1 , \tag{32}$$

which results in the following expression for λ

$$\lambda = -\frac{2}{\sum_{k=1}^{K} \mathrm{Var}\left[p_{LS_k}\right]^{-1}}. \tag{33}$$

Substituting Eq. 33 to Eq. 31 gives the result of Eq. 16. Since the objective function is convex (quadratic in $\boldsymbol{w}$) and the constraint is affine, the result of Eq. 16 is the global optimum.

Appendix B

We consider the linear LSF of Eq. 18, repeated here for convenience:

$$G(\boldsymbol{u}) = -\boldsymbol{\alpha}\boldsymbol{u} + \beta\,, \tag{34}$$

with $\|\boldsymbol{\alpha}\| = 1$. The probability of failure associated with this LSF is $p_F = \Phi(-\beta)$. Assume that we want to estimate p_F with LS using direction $\boldsymbol{\alpha}_k$, i.e. based on the estimator of Eq. 10. At each sample on the hyperplane $\boldsymbol{u}_{\boldsymbol{\alpha}_k}^{\perp,i}$ the distance to the limit state surface along direction $\boldsymbol{\alpha}_k$ is found through solving $G\left(\boldsymbol{u}_{\boldsymbol{\alpha}_k}^{\perp,i} + d\boldsymbol{\alpha}_k^{\mathrm{T}}\right) = 0$, which gives $d^i = \frac{\beta - \boldsymbol{\alpha}\boldsymbol{u}_{\boldsymbol{\alpha}_k}^{\perp,i}}{\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{\mathrm{T}}}$. The sample $\boldsymbol{u}_{\boldsymbol{\alpha}_k}^{\perp,i}$ can be expressed in terms of a sample from the independent standard normal distribution through Eq. 7. Hence, the random variable representing the distance to the failure surface from the hyperplane perpendicular to direction $\boldsymbol{\alpha}_k$ can be expressed as $d_{\boldsymbol{\alpha}_k}(\mathbf{U}) = \frac{\beta - \boldsymbol{\alpha}\mathbf{U} + (\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{\mathrm{T}})\boldsymbol{\alpha}_k\mathbf{U}}{\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{\mathrm{T}}}$. The variance of the LS estimator p_{LS_k} is given by:

$$\mathrm{Var}\left[p_{LS_k}\right] = \frac{\mathrm{Var}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)\right]}{N_k}. \tag{35}$$

Expanding the numerator of Eq. 35, we get

$$\mathrm{Var}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)\right] = \mathrm{E}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)^2\right] - \mathrm{E}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)\right]^2. \tag{36}$$

It is $\mathrm{E}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)\right] = p_F = \Phi(-\beta)$. The first term of the right hand side of Eq. 36 is

$$
\begin{aligned}
&\mathrm{E}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)^2\right] = \mathrm{E}\left[\Phi\left(\frac{-\beta+\boldsymbol{\alpha}\mathbf{U}-\left(\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{\mathrm{T}}\right)\boldsymbol{\alpha}_k\mathbf{U}}{\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{\mathrm{T}}}\right)^2\right] \\
&= \mathrm{E}\left[\Pr\left(\widetilde{U}_1 \le \frac{-\beta+\boldsymbol{\alpha}\mathbf{U}-\left(\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{\mathrm{T}}\right)\boldsymbol{\alpha}_k\mathbf{U}}{\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{\mathrm{T}}}\middle|\mathbf{U}\right)\Pr\left(\widetilde{U}_2 \le \frac{-\beta+\boldsymbol{\alpha}\mathbf{U}-\left(\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{\mathrm{T}}\right)\boldsymbol{\alpha}_k\mathbf{U}}{\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{\mathrm{T}}}\middle|\mathbf{U}\right)\right] \\
&= \mathrm{E}\left[\Pr\left(\left\{\widetilde{U}_1 \le \frac{-\beta+\boldsymbol{\alpha}\mathbf{U}-\left(\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{\mathrm{T}}\right)\boldsymbol{\alpha}_k\mathbf{U}}{\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{\mathrm{T}}}\right\}\cap\left\{\widetilde{U}_2 \le \frac{-\beta+\boldsymbol{\alpha}\mathbf{U}-\left(\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{\mathrm{T}}\right)\boldsymbol{\alpha}_k\mathbf{U}}{\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{\mathrm{T}}}\right\}\middle|\mathbf{U}\right)\right] \\
&= \Pr\left(\left\{\widetilde{U}_1 \le \frac{-\beta+\boldsymbol{\alpha}\boldsymbol{U}-\left(\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{T}\right)\boldsymbol{\alpha}_k\boldsymbol{U}}{\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{T}}\right\}\cap\left\{\widetilde{U}_2 \le \frac{-\beta+\boldsymbol{\alpha}\boldsymbol{U}-\left(\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{T}\right)\boldsymbol{\alpha}_k\boldsymbol{U}}{\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{T}}\right\}\right),
\end{aligned}
\tag{37}
$$

where $\widetilde{U}_1$ and $\widetilde{U}_2$ are auxiliary independent standard normal random variables. Define the random variables $\tilde{Y}_i = \boldsymbol{\alpha}\boldsymbol{\alpha}_k^T\widetilde{U}_i - \boldsymbol{\alpha}\boldsymbol{U} + (\boldsymbol{\alpha}\boldsymbol{\alpha}_k^T)\boldsymbol{\alpha}_k\boldsymbol{U}$, $i = 1{,}2$. The variables $\tilde{Y}_i, i = 1{,}2,$ are binormal, have zero means, unit variances and correlation coefficient $\tilde{\rho}_{12} = 1 - \boldsymbol{\alpha}\boldsymbol{\alpha}_k^{\mathrm{T}}$. It is:

$$
\begin{aligned}
\mathrm{E}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)^2\right] &= \Pr\left(\{\tilde{Y}_1 \le -\beta\}\cap\{\tilde{Y}_2 \le -\beta\}\right) \\
&= \Phi_2(-\beta,-\beta,\tilde{\rho}_{12}),
\end{aligned}
\tag{38}
$$

where $\Phi_2(-\beta,-\beta,r)$ is the bivariate standard normal CDF with correlation parameter r. The bivariate normal CDF can be expressed in terms of a single-fold integral as follows (e.g. Owen 1980):

$$
\Phi_2(-\beta,-\beta,\tilde{\rho}_{12}) = \Phi(-\beta)^2 + \int_0^{\tilde{\rho}_{12}} \varphi_2(-\beta,-\beta,r)dr\,, \tag{39}
$$

where $\varphi_2(-\beta,-\beta,r)$ is the bivariate standard normal PDF. Combining Eqs. 36, 38 and 39, we get:

$$
\begin{aligned}
&\mathrm{Var}\left[\Phi\left(-d_{\boldsymbol{\alpha}_k}(\mathbf{U})\right)\right] = \int_0^{1-\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{\mathrm{T}}} \varphi_2(-\beta,-\beta,r)dr \\
&= \int_0^{1-\boldsymbol{\alpha}\boldsymbol{\alpha}_k^{\mathrm{T}}} \frac{1}{2\pi\sqrt{1-r^2}}\exp\left(-\frac{\beta^2}{1+r}\right)dr\,.
\end{aligned}
\tag{40}
$$

References

de Angelis, M., Patelli, E., & Beer, M. (2015). Advanced line sampling for efficient robust reliability analysis. *Structural Safety*, *52*, 170-182.

Au, S. K., & Beck, J. L. (2001). First excursion probabilities for linear systems by very efficient importance sampling. *Probabilistic Engineering Mechanics*, 16(3), 193-207.

Au, S. K., & Beck, J. L. (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4), 263-277.

Bucher, C. G. (1988). Adaptive sampling—an iterative fast Monte Carlo procedure. *Structural Safety*, 5(2), 119-126.

Hohenbichler, M., & Rackwitz, R. (1981). Non-normal dependent vectors in structural safety. *Journal of the Engineering Mechanics Division*, 107(6), 1227-1238.

Hohenbichler, M., & Rackwitz, R. (1988). Improvement of second-order reliability estimates by importance sampling. *Journal of Engineering Mechanics*, 114(12), 2195-2199.

Koutsourelakis, P. S., Pradlwarter, H. J., & Schuëller, G. I. (2004). Reliability of structures in high dimensions, part I: algorithms and applications. *Probabilistic Engineering Mechanics*, 19(4), 409-417.

Lemaire, M. (2013). *Structural Reliability*. John Wiley & Sons.

Liu, P. L., & Der Kiureghian, A. (1991). Optimization algorithms for structural reliability. *Structural Safety*, 9(3), 161-177.

Melchers, R. E., & Beck, A. T. (2018). *Structural Reliability Analysis and Prediction*. John Wiley & Sons.

Owen, D. B. (1980). A table of normal integrals: A table. *Communications in Statistics-Simulation and Computation*, 9(4), 389-419.

Owen, A. B., & Zhou, Y. (1999). Adaptive importance sampling by mixtures of products of beta distributions. *Technical Report*. Department of Statistics, Stanford University.

Owen, A. B., & Zhou, Y. (2019). The square root rule for adaptive importance sampling. *arXiv preprint arXiv:1901.02976.*

Papaioannou, I., Geyer, S., & Straub, D. (2019). Improved cross entropy-based importance sampling with a flexible mixture model. *Reliability Engineering & System Safety*, 191, 106564.

Papaioannou, I., Papadimitriou, C., & Straub, D. (2016). Sequential importance sampling for structural reliability analysis. *Structural Safety*, 62, 66-75.

Pradlwarter, H. J., Pellissetti, M. F., Schenk, C. A., Schueller, G. I., Kreis, A., Fransen, S., Calvi, A. & Klein, M. (2005). Realistic and efficient reliability estimation for aerospace structures. *Computer Methods in Applied Mechanics and Engineering*, 194(12-16), 1597-1617.

Pradlwarter, H. J., & Schuëller, G. I. (2010). Uncertain linear structural systems in dynamics: Efficient stochastic reliability assessment. *Computers & Structures*, 88(1-2), 74-86.

Pradlwarter, H. J., Schuëller, G. I., Koutsourelakis, P. S., & Charmpis, D. C. (2007). Application of line sampling simulation method to reliability benchmark problems. *Structural Safety*, 29(3), 208-221.

Schuëller, G. I., Pradlwarter, H. J., & Koutsourelakis, P. S. (2004). A critical appraisal of reliability estimation procedures for high dimensions. *Probabilistic Engineering Mechanics*, 19(4), 463-474.

Schuëller, G. I., & Stix, R. (1987). A critical appraisal of methods to determine failure probabilities. *Structural Safety*, 4(4), 293-309.

Wen, Y. K., & Chen, H. C. (1987). On fast integration for time variant structural reliability. *Probabilistic Engineering Mechanics*, 2(3), 156-162.