Structural Safety, 5 (1988) 119–126
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

ADAPTIVE SAMPLING—AN ITERATIVE FAST MONTE CARLO PROCEDURE

Christian G. Bucher

Institute of Engineering Mechanics, University of Innsbruck, Technikerstrasse 13, A-6020 Innsbruck (Austria)

(Received August 5, 1987; accepted December 15, 1987)

ABSTRACT

An iterative Monte-Carlo simulation procedure for structural analysis is suggested. This proposed new approach utilizes results from simulation to adapt the importance sampling density to the specific problem. Considerable reduction of the statistical error of the estimated failure probability is achieved. Most important, problems connected with optimization procedures commonly used in structural reliability are avoided. This makes the suggested procedure especially attractive for systems reliability analyses.

1. INTRODUCTION

In structural reliability analyses exact methods to determine failure probabilities become increasingly attractive. This is mainly due to the fast development of computational facilities. Although the underlying concept, i.e. the evaluation of a multidimensional integral on a largely irregular domain, is very simple, rather involved computational procedures are necessary to obtain satisfactory results with acceptable computational cost. For low dimensions, numerical integration schemes (e.g. Ref. [1]) prove to be very efficient, but for cases with a larger number of basic variables Monte Carlo simulation seems to be superior [2]. Special sampling techniques were developed to reduce the statistical error inherent in Monte Carlo methods, e.g. Importance Sampling [3], Antithetic Variates [2,4], Latin Hypercube Sampling [5,6]. The computational efficiency of importance sampling using the design point is shown in Ref. [7]. However, the importance sampling densities, which are widely used, are still far from the optimal ones. Another problem might arise, if the optimization procedure utilized to find the design point fails to find the global minimum-distance point in the transformed Gaussian space. This may happen especially for problems of systems reliability, i.e. in multiple-failure-mode problems.

The purpose of this paper is to present an efficient sampling procedure which iteratively adapts the importance sampling density according to the respective individual problem.

0167-4730/88/$03.50

2. STATEMENT OF THE PROBLEM

The basic equation of structural reliability is

$$p_f = \int_{D_f} f_X(x)\, dx \tag{1}$$

in which p_f is the failure probability, D_f the failure domain, x the vector of basic variables and $f_X(x)$ the joint density of the x-components. Using importance sampling, eqn. (2) is replaced by

$$p_f = \int_{D_f} \frac{f_X(y)}{g_Y(y)} g_Y(y)\, dy \tag{2}$$

in which the importance sampling density $g_Y(y)$ is chosen to reduce the statistical error of the estimated $\bar{p}_f$ value

$$\bar{p}_f = \frac{1}{N} \sum_{i=1}^{N} I_{D_f}(y_i) \frac{f_Y(y_i)}{g_Y(y_i)} = \frac{1}{N} \sum_{i=1}^{N} p_{f_i} \tag{3}$$

It is well known, that the statistical error of the estimate obtained by eqn. (3) reduces to zero, if $g_Y(y)$ is chosen to be the original density conditional on the failure domain [2]:

$$g_Y(y) = f_X(y \mid y \in D_f) \tag{4}$$

In practice, however, this is an impossible choice.

Using an iterative adaptive algorithm, eqn. (4) can be satisfied at least in terms of first and second moments, so that

$$E_g[y] = E_f[y \mid y \in D_f] \tag{5a}$$

$$E_g[yy^T] = E_f[yy^T \mid y \in D_f] \tag{5b}$$

is satisfied. The indices g and f refer to the respective joint densities. From an initial simulation run the right-hand sides of eqns. (5) can be estimated. These values are used to adapt $g_Y(y)$ for the next run. Since the knowledge of mean and covariances uniquely determines a jointly normal density, this type of density is chosen for $g_Y(y)$ in the following analysis.

A simple example already shows that the conditional mean value for this Iterative Fast Monte-Carlo procedure (IFM-point) can be quite far away from the design point. Let for a one-dimensional problem the failure domain be

$$D_f = \{x \mid x > \beta\} \tag{6}$$

and $f_X(x)$ standard normal.

It is easily shown that

$$E_f[x \mid x > \beta] = \bar{x}_g = \frac{\exp(-\beta^2/2)}{\sqrt{2\pi}\,\Phi(-\beta)} \tag{7a}$$

$$E_f\left[(x - \bar{x}_g)^2 \,\middle|\, x > \beta\right] = \sigma_{x_g}^2 = 1 + \beta\bar{x}_g - \bar{x}_g^2 \tag{7b}$$

The original density and a normal importance sampling density based on eqns. (7) are plotted in Fig. 1.

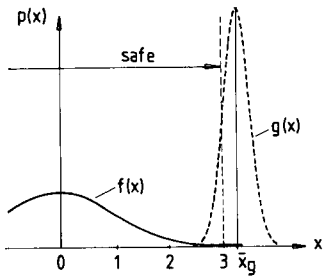

Fig. 1. IFM-point and importance sampling density for one-dimensional case.

For $\beta = 3$ the results are $\bar{x}_g = 3.29$ and $\sigma_{x_g} = 0.214$. This already shows that the optimum (Gaussian) importance sampling density has considerably different properties as compared to the original density. It is seen that almost all the samples will be concentrated in the failure domain and, moreover, in that part of the failure domain which contributes most to the total failure probability.

3. ADAPTIVE SAMPLING

3.1 General

It is known that for problems involving a large number of random variables the statistical uncertainties involved in Monte Carlo procedures become more significant. In context with adaptive sampling this means that the estimates of conditional mean and covariances must be sufficiently accurate in order to obtain satisfactory convergence of the procedure. If the limit state function is available in functional form, then the initial run may be performed as standard importance sampling [7], using the original densities shifted to the design point. There are, however, cases in which the design point cannot be found. This may be due to the existence of multiple "design point", i.e. the optimization algorithm finds a local minimum-distance point (in standard Gaussian space) which is not the global minimum. The situation is even worse, if the limit state function cannot be written in functional form, i.e. only the characteristic function of the failure domain is available. Whereas in the first case the global minimum may be found using different starting values, in the second case most optimization procedures (especially gradient methods) fail.

For those cases a starting procedure is suggested which is based on pure Monte Carlo simulation. It should be mentioned that for the sake of simplicity the following notations refer to uncorrelated variables. The analysis for correlated variables proceeds in the same way after a transformation to uncorrelated space is performed.

3.2 Starting procedure

In order to obtain a starting vector for the conditional mean, a kind of sensitivity analysis with respect to each random variable is performed. Simulation for each variable is carried out while

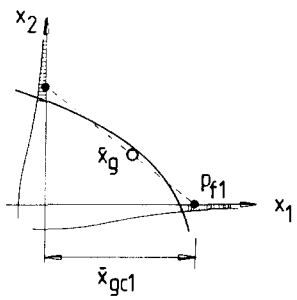

Fig. 2. Interpretation of starting vector $\bar{x}_g$ as mass center of conditional failure probabilities p_{f_i}.

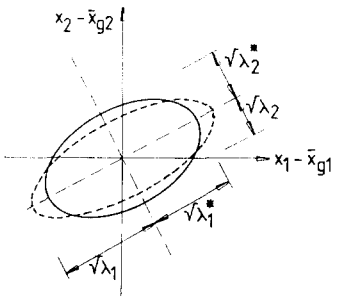

Fig. 3. Geometrical interpretation of partly averaged eigenvalues λ_i^*.

all the other variables are kept at the mean value. This results in conditional failure probabilities p_{f_i} for each variable.

Correspondingly the individual IFM-points $\bar{x}_{g_{ci}}$ are found. An approximation for the starting vector $\bar{x}_g$ is obtained by the weighted average.

$$\bar{x}_{g_i} = p_{f_i} \bar{x}_{g_{ci}} / \sum_{i=1}^{n} p_{f_i} \tag{8}$$

Equation (8) may be interpreted as the mass center of masses p_{fi} located on the axes x_i at a distance $\bar{x}_{g_{ci}}$ from the origin (cf. Fig. 2).

Depending on the problem, the starting value $\bar{x}_g$ may be relatively far in the failure domain. In this case, a safe point on the straight line between $\bar{x}_g$ and the mean vector $\bar{x}$ is used as new starting point. This strategy prevents simulation entirely in the failure domain which could lead to unconservative estimates of the failure probability.

In the next step, simulation around $\bar{x}_g$ is carried out, again varying each variable x_i at a time. From these simulations, the conditional standard deviations σ_{g_i} are estimated.

3.3 Iterative Fast Monte Carlo

Using the starting vector $\bar{x}_g$ and the initial standard deviations σ_{g_i}, importance sampling is carried out. From the failure events the conditional mean value $\bar{x}_g$ replaces the starting vector $\bar{x}_g$. The conditional covariance matrix $\boldsymbol{C}_{x_g}$ is estimated as well and diagonalized by means of an eigenvalue analysis. Since the statistical errors inherent in the estimation of $\boldsymbol{C}_{x_g}$ may distort the eigenvalues thus causing convergence problems, the eigenvalues λ_i are replaced by λ_i^*.

$$\lambda_i^* = c \sum_{i=1}^{n} \lambda_i + (1 - c)\lambda_i \tag{9}$$

In the above equation c is the coefficient of variation of the estimated failure probability. The geometrical interpretation of eqn. (9) is given in Fig. 3. From a probabilistic viewpoint eqn. (9) means that the eigenvalues are partly averaged thus reducing the statistical uncertainty. The theoretical efficiency of adaptive sampling is slightly reduced but convergence was found to be improved considerably.

4. APPLICATION TO RELIABILITY ANALYSIS

4.1 Implementation

The Iterative Fast Monte Carlo procedure as developed in Section 3 has been implemented in the general-purpose reliability code ISPUD [8]. The probabilistic environment of ISPUD allows the treatment of correlated non-Gaussian variables. For this purpose the joint density model as suggested by Nataf [9] is utilized. For further details the reader is referred to Ref. [8].

4.2 Numerical examples

The following examples show the results for importance sampling densities obtained by iteration based on eqns. (5) and the corresponding estimate p_f of the failure probability along with the respective coefficient of variation c. The iteration starts as described in Section 3.2 in all cases.

Example 1

Let the failure domain D_f be defined by

$$D_f = \{(x_1, x_2) \,|\, (4 - x_1 < 0) \vee (4 - x_2 < 0)\} \tag{10}$$

with x_1, x_2 independent standard normal. This example represents typically the series failure of two components. Obviously there are two "design points" (cf. Fig. 4) which are both equally likely. However, using optimization with only one starting value, only one of these points can be identified. Table 1 compares the results from adaptive sampling with those from importance sampling (density shifted to one of the design points). It is seen that conventional importance sampling practically neglects one of the failure modes.

Additionally, in Fig. 4 the 1-σ contours of the sampling densities for adaptive sampling and standard importance sampling are compared

It should be mentioned, that for this example the individual failure modes can be identified quite easily but this need not be true for more complicated practical situations.

Example 2

Let

$$D_f = \left\{ (x_1, x_2, \ldots, x_{10}) \,\middle|\, 3 + x_1 - \tfrac{1}{6} \sum_{i=2}^{10} x_i^2 < 0 \right\} \tag{11}$$

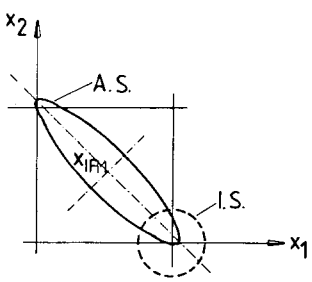

Fig. 4. Comparison of sampling densities for adaptive sampling and importance sampling (Example 1).

TABLE 1

Comparison of results for Example 1

Method	Run No. [a]	p_f	c.o.v.
Adaptive sampling	1	6.73×10^{-5}	0.15
	2	6.54×10^{-5}	0.07
	3	6.53×10^{-5}	0.06
Importance sampling	1	3.01×10^{-5}	0.05
	2	3.32×10^{-5}	0.04
	3	3.33×10^{-5}	0.04
First-order	–	3.17×10^{-5}	–
Exact	–	6.34×10^{-5}	–

[a] Each run consisted of 1000 simulations.

TABLE 2

Comparison of results for Example 2

Method	Run No.	p_f	c.o.v.
Adaptive sampling	1	0.101	0.12
	2	0.110	0.10
	3	0.116	0.07
Importance sampling	1	0.151	0.36
	2	0.402	0.60
	3	0.143	0.39
First-order	–	1.35×10^{-3}	–
Exact	–	0.111	–

with $x_i (i = 1,2,\ldots,10)$ independent standard Gaussian. This highly non-linear problem has previously been used [7] to indicate linearization errors introduced by first-order reliability estimates.

The resulting failure probabilities and coefficients of variation from 3 runs with 500 simulations each are given in Table 2. For comparative purposes the results based on standard importance sampling around the design point and the first-order approximation are given as well.

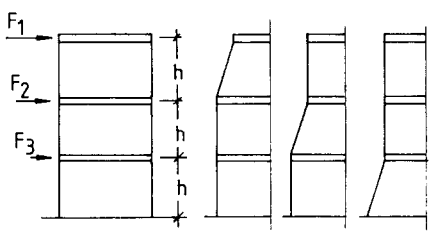

Fig. 5. Shear beam frame structure with failure modes.

TABLE 3

Probabilistic data of basic variables for Example 3

Variable	Type	Mean	Standard deviation
F_1	Gumbel	10.0	4.0
F_2	Gumbel	7.0	3.0
F_3	Gumbel	5.5	2.0
M_1/h	Lognormal	9.0	1.6
M_2/h	Lognormal	13.0	2.0
M_3/h	Lognormal	15.0	2.4

TABLE 4

Correlation matrix of basic variables for Example 3

$\rho_{x_i x_j}$	F_1	F_2	F_3	M_1/h	M_2/h	M_3/h
F_1	1.0					
F_2	0.5	1.0			sym.	
F_3	0.2	0.5	1.0			
M_1/h	0	0	0	1.0		
M_2/h	0	0	0	0.3	1.0	
M_3/h	0	0	0	0	0.3	1.0

Example 3

Consider a frame structure (shear beam) as indicated in Fig. 5. If failure is defined in terms of partial or total collapse, then there are 3 failure modes (Fig. 5). Considering the plastic moments M_i identical within each storey, the failure domain is defined by

$$D_f = \{(F_1, F_2, F_3, M_1, M_2, M_3) \mid (F_1 > 4M_1/h) \vee (F_1 + F_2 > 4M_2/h) \vee (F_1 + F_2 + F_3 > 4M_3/h)\} \quad (12)$$

The probabilistic data for the analysis are given in Table 3, while Table 4 shows the correlation matrix. The numerical results from 1000 simulations are given in Table 5 along with the coefficients of variation.

TABLE 5

Comparison of results for Example 3

Method	Run No.	p_f	c.o.v.
Adaptive sampling	1	8.8×10^{-4}	0.14
	2	2.0×10^{-3}	0.14
	3	2.3×10^{-3}	0.08
Importance sampling (one design point)	1	1.4×10^{-3}	0.31
	2	1.1×10^{-3}	0.08
	3	1.2×20^{-3}	0.15
First-order	–	6.6×10^{-4}	–
Exact	–	2.4×10^{-3}	–

5. CONCLUSIONS

The numerical results obtained from the suggested Iterative Fast Monte Carlo procedure show that structural reliability analysis may be carried out very efficiently based on simulation only without resort to analytical approximations. The procedure is especially suitable for systems reliability analysis, since multiple failure modes need not be treated separately. Moreover, the limit state can be described in terms of the indicator function only, i.e. for each simulated set of random variables only the decision safe/unsafe has to be made, regardless of the actual failure mode. Hence, problems encountered with widely used optimization algorithms are entirely avoided. The present algorithm is seen to include multiple failure modes, whereas optimization-based procedures may neglect significant contributions to the failure probability.

Finally it should be mentioned that the implementation within the code ISPUD allows the application of the present Iterative Fast Monte Carlo procedure within a wide range of reliability problems encountered in engineering.

ACKNOWLEDGEMENT

This study was partially supported by the Austrian Research Association (FWF) under Contract No. P5573 which is acknowledged. Special thanks are due to G.I. Schuëller for his encouragement and discussions as well as to U. Bourgund for discussions.

REFERENCES

1 A.C. Genz and A.A. Malik, Remarks on algorithm 006: an adaptive algorithm for numerical integration over N-dimensional rectangular region, J. Comput. Appl. Math., 6(4) (1980) 295–299.
2 R.Y. Rubinstein, Simulation and the Monte-Carlo Method, John Wiley & Sons, New York, NY, 1981.
3 A. Harbitz, Efficient and accurate probability of failure calculation by the use of importance sampling technique, in: G. Augusti et al. (Eds.), Proc. 4th Int. Conf. on Applications of Statistics and Probability in Soil and Structural Engineering (ICASP 4),Pitagora, Editrice, Bologna, Italy, 1983, pp. 825–886.
4 J.M. Hammerley and D.C. Handscomb, Monte Carlo Methods, John Wiley & Sons, New York, NY, 1964.
5 R.L. Iman and W.J. Canover, Small sampling sensitivity analysis techniques for computer models with an application to risk assessment, Commun. Stat., Theory Methods, A9(17) (1980) 1749–1842.
6 N. Hwang, M. Reich, B. Ellingwood and M. Shinozuka, Reliability assessment and probability based design of reinforced concrete containments and shear walls, Summary Report, NUREG, CR-3957, BNL-NUREG 51956, Upton, New York, 1986.
7 G.I. Schuëller and R. Stix, A critical appraisal of methods to determine failure probabilities, Structural Safety, 4(4) (1987) 293–309.
8 U. Bourgund and C.G. Bucher, Importance Sampling Procedure Using Design Points — ISPUD — a User's Manual, Inst. Eng. Mech., Univ. Innsbruck, Rep. No. 8-86, Nov. 1986.
9 A. Nataf, Détermination des distributions de probabilités dont les marges sont donnés, C. R. Acad. Sci. (Paris), 255 (1962) 42–43.