Available online at www.sciencedirect.com

ScienceDirect

Structural Safety 29 (2007) 183–193

STRUCTURAL
SAFETY

www.elsevier.com/locate/strusafe

# Application of subset simulation methods to reliability benchmark problems

S.K. Au [a,*], J. Ching [b], J.L. Beck [c]

[a] *Department of Building and Construction, City University of Hong Kong, 83 Tat Chee Avenue, Hong Kong*
[b] *Department of Construction Engineering, National Taiwan University of Science and Technology, Taipei 106, Taiwan*
[c] *Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA 91125, USA*

Available online 1 September 2006

## Abstract

This paper presents the reliability analysis of three benchmark problems using three variants of Subset Simulation. The original version of Subset Simulation, SubSim/MCMC, employs a Markov chain Monte Carlo (MCMC) method to simulate samples conditional on intermediate failure events; it is a general method that is applicable to all the benchmark problems. SubSim/Splitting is a variant of Subset Simulation applicable to first-passage problems involving deterministic dynamical systems. It makes use of trajectory splitting for generating conditional samples. Another variant, SubSim/Hybrid, uses a combined MCMC/Splitting strategy and so it has the advantages of MCMC and splitting; it is applicable to uncertain and deterministic dynamical systems. Results show that all three Subset Simulation variants are effective in high-dimensional problems and that some computational efficiency can be gained by exploiting and incorporating system characteristics into the simulation procedure.

*Keywords:* Monte Carlo method; Reliability; Splitting; Stochastic simulation; Subset simulation

## 1. Introduction

When designing new civil, mechanical or aerospace systems that will experience dynamic excitation in their operating environment, it is desirable to quantify the predicted performance of a proposed design in terms of the reliability that it will achieve the specified design objectives. In view of the uncertainties about the modeling of systems and about the future dynamic excitation the system will experience, the design team can specify a set of possible dynamic inputs and a set of possible models of the system and then choose probability distributions over these sets to model the uncertainties. One can then evaluate the 'failure probability' of the design that measures how likely the system will achieve the desired performance over its operational lifetime, based on available information and the probability models chosen to represent the missing information

[*] Corresponding author. Present address: Formerly at School of Civil and Environmental Engineering, Nanyang Technological University, Singapore. Tel.: +852 2194 2769; fax: +852 2788 7612.
*E-mail address:* siukuiau@cityu.edu.hk (S.K. Au).

0167-4730/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.strusafe.2006.07.008

$$P(F) = \int_{R^n} I(\underline{\theta} \in F) q(\underline{\theta}) \mathrm{d}\underline{\theta} = \int_F q(\underline{\theta}) \mathrm{d}\underline{\theta} \tag{1}$$

This equation specifies the probability that $\underline{\Theta} \in R^n$ lies in the failure region $F \subset R^n$, given the probability model that specifies the PDF (probability density function) $q$ for $\underline{\Theta}$ containing all the uncertain-valued parameters needed to define completely the excitation and model of the system. Here, $I(\cdot)$ is an indicator function that is equal to 1 when its argument is true and is equal to zero otherwise. The formulation for the failure probability in (1) is general and is applicable to static and dynamic reliability problems. Over the past few decades, a number of reliability methods have been developed that are effective when the number of variables $n$ is not too large or when the failure surface has limited complexity. Excellent reviews can be found at different stages of development, e.g. [1–4].

In recent years, attention has been focused on reliability problems with complex system characteristics and with high dimensions (i.e., with a large number of uncertain or random variables) [5]. High-dimensional problems are frequently encountered in reliability problems involving stochastic processes or random fields, whose discretized representation requires a large number of i.i.d. (independent and identically distributed) variables. Ideally, the dimension of a reliability problem should be determined based on modeling reasons rather than be limited by the capability of reliability methods. Stochastic simulation methods provide an attractive means for solving high-dimensional problems, especially for complex systems where analytical results or knowledge about the dependence of the response on the excitation and modeling parameters are rarely available.

The best known and most robust stochastic simulation method is direct Monte Carlo Simulation (MCS), where the failure probability is estimated by simple statistical averaging of the indicator function

$$P(F) \approx \frac{1}{N} \sum_{k=1}^{N} I(\underline{\Theta}_k \in F) \tag{2}$$

where $\{\underline{\Theta}_k : k = 1, \ldots, N\}$ are i.i.d. samples drawn from the PDF $q$. A system analysis is required for each sample $\underline{\Theta}_k$ to check whether it corresponds to failure ($I(\underline{\Theta}_k \in F) = 1$) or not ($I(\underline{\Theta}_k \in F) = 0$). While MCS is applicable to all types of reliability problems, its computational efficiency is a practical concern when estimating small failure probabilities because information must be gained from samples that correspond to failure but these are rarely simulated. A rule of thumb for MCS is that one must generate at least 10 failure samples to get a reasonably accurate estimate of $P(F)$ from (2), so if $P(F) = 0.001$, at least 10000 system analyses must be performed when estimating $P(F)$. This has motivated recent research to develop more efficient stochastic simulation algorithms for high-dimensional dynamic reliability problems. A basic and important feature of most stochastic simulation methods is that they estimate the integral for $P(F)$ by gaining information about the system behavior through sampling $\underline{\Theta}$ from some probability distribution and then using information from these samples to account for the failure probability.

Schueller and co-workers recently organized a benchmark study on reliability estimation in higher dimensions of structural systems [6] with a goal to assess the performance of available numerical techniques. The benchmark problems comprise high-dimensional static and dynamic problems with uncertainties in system and loading properties. This paper presents the application of a stochastic simulation technique called 'Subset Simulation' to the benchmark problems. An overview of the method is presented first in the next section.

## 2. Subset Simulation

Subset Simulation [7,8] is an adaptive stochastic simulation procedure for efficiently computing small failure probabilities. Strictly speaking, it is a procedure for efficiently generating samples that correspond to specified levels of failure probabilities in a progressive manner. The method is motivated by the observation that the inefficiency of direct MCS in reliability problems arises essentially from the need to estimate small failure probability, or equivalently, from the need to generate rare failure samples. This leads to the question as to whether rare-event simulation can be avoided in finding small failure probabilities. The simple but pivotal idea behind Subset Simulation is that a small failure probability can be expressed as a product of larger conditional failure probabilities for some intermediate failure events, suggesting the possibility of converting a problem involving rare-event simulation into a sequence of problems involving more frequent events.

This idea can be expressed as follows. Let $F$ denote the target failure region in the space of the variables $\underline{\Theta}$ and let $F_1 \supset F_2 \supset \cdots \supset F_m = F$ be a chosen sequence of $m$ nested failure regions. By the definition of conditional probability

$$P(F) = P(F_m|F_{m-1})P(F_{m-1}) = \cdots = P(F_1)\prod_{i=2}^{m} P(F_i|F_{i-1}) \tag{3}$$

This equation indicates that instead of directly calculating a small $P(F)$, one can *in principle* calculate the probabilities $P(F_1)$, $P(F_i|F_{i-1})$ $(i = 2\ldots,m)$ and then take their product. The potential advantage of this alternative is that the probabilities involved can be chosen to be much greater than $P(F)$, thus avoiding simulation of rare events. Several implementation issues immediately arise, however: (1) How to select the intermediate failure events $F_i$ $(i = 1\ldots,m-1)$, and (2) how to estimate the conditional probabilities $P(F_i|F_{i-1})$ $(i = 2\ldots,m)$?

The first question regarding the choice of intermediate failure regions is addressed by specifying a general parametric form to define the failure region and to decide on the number of intermediate failure regions as well as their conditional failure probabilities. The strategy in Subset Simulation is to assume that $F$ can be generically represented as the exceedance of a critical response quantity $Y \geqslant 0$ above some specified threshold level $y$, so

$$F = \{Y > y\} \tag{4}$$

Remarkably, this representation does not lead to much loss of generality. For example, if $b_i > 0$ $(i = 1,\ldots,m)$ and $F = \bigcup_{i=1}^{m}\{Y_i > b_i\}$, then $F = \{Y > 1\}$ where $Y = \max_{i=1,\ldots,m} Y_i/b_i$. In general, $Y$ is defined using 'max' and 'min' in the same order corresponding to each occurrence of union ($\cup$) and intersection ($\cap$) in $F$, respectively. By considering $P(F) = P(Y > y)$ for different values of $y$, one can investigate the exceedance probability curve as a function of the threshold $y$, which is more informative than just a point estimate of $P(F) = P(Y > 1)$. Notice that calculating $P(F) = P(Y > y)$ as a function of $y$ is equivalent to finding the complementary CDF (cumulative distribution function) of the response quantity $Y$, where the complementary CDF = 1 – CDF, and this can be achieved down to very low probabilities in the tail of this distribution because of the efficiency of Subset Simulation.

Based on the representation of $F$ in (4), the intermediate failure regions can be parameterized as

$$F_i = \{Y > y_i\} \tag{5}$$

where $0 < y_1 < \ldots < y_m = y$ form an increasing sequence of intermediate threshold values. The question is then how to choose $\{y_i{:}i = 1,\ldots,m-1\}$. The intermediate threshold values will affect the ease of estimating the conditional probabilities, because if $y_{i-1} \ll y_i$, $P(F_i|F_{i-1})$ will be very small, and one will face a rare-event simulation problem as in direct MCS. Thus, the intermediate threshold values should be chosen close enough so that the conditional failure probabilities are not small. They should not be chosen to be too close, however, because then it would take a large total number of levels $m$ (and hence large computational effort) to progress to the target failure region of interest. A prudent choice should thus strike a balance between these two extremes.

One may think that it would be difficult to know a priori what intermediate threshold values to choose in order to induce reasonable values of the conditional probabilities, but this issue is resolved simply by adaptively choosing the intermediate threshold values so that the resulting *sample estimates* of the conditional failure probability correspond to a common specified value of $p_0 \in (0.1, 0.2)$, say. Specifying a common value for the conditional probabilities is merely for convenience of implementation but it is also a reasonable choice. Let $\{Y_{i-1,k}{:}k = 1,\ldots,N\}$ be the response values corresponding to the conditional samples $\{\underline{\Theta}_{i-1,k}{:}k = 1,\ldots,N\}$ distributed as $q(\cdot|F_{i-1})$. For a specified value of $p_0$, the intermediate threshold value $y_i$ defining $F_i$ through (5) can be simply obtained as the $[(1-p_0)N]$th largest value among $\{Y_{i-1,k}{:}k = 1,\ldots,N\}$ so that the sample estimate of $P(F_i\ |F_{i-1}) = P(Y > y_i|Y > y_{i-1})$ is by construction equal to $p_0$. Here, it is assumed that $p_0$ and $N$ are chosen so that $p_0N$ is an integer. Subset Simulation is, in fact, primarily a procedure for generating threshold values of response that correspond to some specified value of failure probability, rather than for estimating failure probabilities at specified response values. One issue with this adaptive choice of $y_i$ is that the conditional failure probability $P(F_i|F_{i-1})$ is not actually equal to $p_0$, but only its sample estimate is. In implementations, one needs to choose $N$ large enough so that the variability of $y_i$ and hence the error in

$P(F_i|F_{i-1}) \approx p_0$ is kept small. Due to the adaptive choice of intermediate failure events, the failure probability estimates for a fixed response level is only asymptotically unbiased as $N \to \infty$ and the bias is $O(1/N)$.

On the other hand, estimating the conditional probabilities $P(F_i|F_{i-1})$ may appear to be difficult to resolve because conditional probabilities are involved. Adopting a stochastic simulation approach, the conditional failure probability $P(F_i|F_{i-1})$ can be estimated by the fraction of the number of samples lying in $F_i$, where these samples are generated according to the conditional PDF $q(\cdot|F_{i-1})$. Generating these conditional samples is feasible if one uses an acceptance/rejection algorithm, but this is not a computationally *efficient* approach because a sample distributed as $q(\cdot|F_{i-1})$ is obtained by accepting a sample generated according to $q(\cdot)$ that also lies in $F_{i-1}$. This does not avoid rare-event simulation because, on average, it requires $1/P(F_{i-1})$ samples (and hence this many system analyses) to obtain one such conditional sample.

The original version of Subset Simulation [7,8] makes use of a powerful Markov Chain Monte Carlo (MCMC) technique for generating conditional samples called the Metropolis–Hastings algorithm [9–11]. Originally developed for statistical physics problems, this algorithm allows efficient generation of random samples according to an arbitrarily given probability distribution, even if it is not normalized. Due to their versatility and relevance to diverse problems, MCMC methods have found application in a wide range of fields, including astrophysics, image processing, biostatistics, phylogeny, etc [12].

Two newer variants of Subset Simulation have recently been developed that exploit the causality of dynamical systems to use trajectory splitting to generate conditional samples for the Subset Simulation approach. The first variant, Subset Simulation with Splitting (SubSim/Splitting) [13], is applicable to deterministic dynamical systems subjected to stochastic excitations. The second variant, Subset Simulation with hybrid MCMC/Splitting strategy (SubSim/Hybrid) [14], is applicable to uncertain and deterministic dynamical systems. These newer versions of Subset Simulation use the same adaptive procedure as the original version to generate conditional samples for higher threshold levels based on those from lower levels. They differ only in the way the next conditional sample is generated from the current one. Each of the three variants of Subset Simulation are now described in more detail.

### 2.1. Subset Simulation/MCMC

The original version of Subset Simulation first appeared in [8] where a component-based Metropolis–Hastings algorithm was developed to overcome the rejection problems in high dimensions [7]. The method was further generalized in [15] to incorporate knowledge about important variables into the procedure. In this paper, the original Subset Simulation method is referred to as SubSim/MCMC because it is based on a Markov Chain Monte Carlo simulation method, the M–H (Metropolis–Hastings) algorithm [9,10]. In the M-H algorithm, new random, or *candidate*, samples are generated by some *proposal* probability distribution chosen by the user and they are accepted or rejected based on a rule depending on the ratio of the desired target probability distribution to the proposal distribution. The proposal distribution governs the choice of the candidate sample, and consequently the efficiency of this MCMC procedure. It should be chosen so that the correlation among the Markov chain samples is small, as this reduces the amount of independent information available for statistical estimation through averaging and, hence, reduces efficiency. The next Markov chain sample correlates with the current one through two basic mechanisms: (1) the candidate sample is rejected and the current sample is taken as the next sample; (2) the candidate sample is accepted but it is in close proximity (due to the choice of the proposal distribution) to the current sample. A prudent choice of proposal distribution should make a balance between these two factors. Experience shows that the efficiency of Subset Simulation is not especially sensitive to the choice of proposal PDF, although improvements can often be made by tailoring it to a particular application. Details for the choice of proposal PDF in dynamic reliability problems can be found in [15].

The procedure for adaptively generating samples of $\underline{\Theta}$ conditional on $F_i$ $(i = 1, \ldots, m)$ is summarized as follows. First, $N$ samples $\{\underline{\Theta}_{0,k}: k = 1, \ldots, N\}$ are simulated by direct Monte Carlo simulation and so they are i.i.d. as the original parameter PDF $q$. The subscript '0' here denotes that the samples correspond to 'Conditional Level 0' (i.e., unconditional). The corresponding values of the critical response $\{Y_{0,k}: k = 1, \ldots, N\}$ are then computed. The value of $y_1$ is chosen as the $[(1 - p_0)N]$th value in the ascending list of $\{Y_{0,k}: k = 1, \ldots, N\}$, so that the sample estimate for $P(F_1) = P(Y > y_1)$ is always equal to $p_0$.

Due to the choice of $y_1$, there are $p_0N$ samples among $\{\underline{\Theta}_{0,k}:k=1,\ldots,N\}$ whose response $Y$ lies in $F_1=\{Y>y_1\}$. These are samples at 'Conditional Level 1' and are conditional on $F_1$. Starting from each of these samples, the M–H algorithm is used to simulate an additional $(1-p_0)N$ conditional samples so that there is a total of $N$ conditional samples at Conditional Level 1. For details of the M–H algorithm used in Subset Simulation/MCMC, the reader is referred to [7] or [15]. The value of $y_2$ is then chosen as the $[(1-p_0)N]$th value in the ascending list of $\{Y_{1,k}:k=1,\ldots,N\}$, and it defines $F_2=\{Y>y_2\}$. Note that the sample estimate for $P(F_2|F_1)=P(Y>y_2|Y>y_1)$ is automatically equal to $p_0$. Again, there are $p_0N$ samples lying in $F_2$. They are samples conditional on $F_2$ and provide 'seeds' for applying the M–H algorithm to simulate an additional $(1-p_0)N$ conditional samples so that there is a total of $N$ conditional samples at Conditional Level 2.

This procedure is repeated for higher conditional levels until the samples at Conditional Level $(m-1)$ have been generated. If the target probability of failure is specified as $P(F)=p_0^m$, then $y_m$ is chosen as the $[(1-p_0)N]$th value in the ascending list of $\{Y_{m-1,k}:k=1,\ldots,N\}$. In general, for any specified value of $y$, $P(Y>y)$ is estimated as

$$P(Y>y)\approx\begin{cases}\frac{1}{N}\sum_{k=1}^{N}I(Y_{0,k}>y) & y<y_1\\ p_0^{i-1}\frac{1}{N}\sum_{k=1}^{N}I(Y_{i-1,k}>y) & y_{i-1}<y<y_i,\ i=2,\ldots,m\\ p_0^{m-1}\frac{1}{N}\sum_{k=1}^{N}I(Y_{m-1,k}>y) & y>y_m\end{cases}\tag{6}$$

### 2.2. Subset Simulation/Splitting

SubSim/Splitting exploits a causality feature of deterministic dynamical systems subjected to additive stochastic excitations to generate the next conditional sample from the current one, namely, the distribution of the future excitation after the first-passage time is just the original distribution conditional on the past excitation. Let $\underline{U}^+$ be the future excitation after the first passage time for the $i$th intermediate failure event $F_i$ and $\underline{U}^-$ be the past excitation, i.e., $\underline{\Theta}=[\underline{U}^-,\underline{U}^+]$ is the current conditional sample. By Bayes' theorem

$$p(\underline{U}^+|F_i,\underline{U}^-)=\frac{P(F_i|\underline{U}^+,\underline{U}^-)}{P(F_i|\underline{U}^-)}P(\underline{U}^+|\underline{U}^-)=P(\underline{U}^+|\underline{U}^-)\tag{7}$$

since $P(F_i|U^+,U^-)=1$ and $P(F_i|U^-)=1$, a direct consequence of the hypothesis that the system fails when the excitation before the passage time is $U^-$. This shows that, given the excitation $U^-$ before the first-passage time, the distribution of the future excitation $U^+$ conditional on failure is just equal to its original distribution (given $U^-$). The implication is that a random sample of excitation time-history conditional on being in an intermediate failure region can be used to generate a new failure sample by keeping the current sample up to the first-passage time and then randomly generating the post-failure portion by direct Monte Carlo simulation. Due to the manner in which the new conditional sample is generated from an excitation time-history sample, the procedure is called *splitting* (e.g., [16,17]). One case where the future excitation conditioned on the past excitation can be easily sampled is a stochastic excitation defined by a windowed filtered white-noise process.

The advantage of SubSim/Splitting is that the next conditional sample is always distinct from the current one, in contrast to SubSim/MCMC where the next sample may be identical to the current one due to the rejection of the candidate sample in the M–H algorithm. The 'off-spring' conditional failure samples generated by splitting from the same 'mother' conditional failure sample have independent components after the first-passage time but the same components prior to this time. In principle, to obtain a new conditional sample by splitting, only the portion of the response time history after the first-passage point needs to be computed, as distinct from SubSim/MCMC where the full response time history needs to be computed. Thus, when the total numbers of trajectories are the same, the computational cost for SubSim/Splitting will usually be less than that for SubSim/MCMC. This, of course, assumes that provisions have been made in the dynamic analysis software to keep track of the first passage time during time-stepping so that computation of the remaining portion of the response time history can be performed in an automated fashion. The average computational effort required for producing an offspring at the $i$th conditional level is only $1 - E[T_f|F_i]/T$ of the equivalent of one sample (with full time history of duration $T$ analyzed), where $E[T_f|F_i]$ is the conditional expected first passage time. For the results presented later in Table 4, the number of samples $N_T$ reported for SubSim/Splitting represents this equivalent number of samples by taking advantage of this computational saving.

One disadvantage of SubSim/Splitting is that the generated conditional samples do not explore efficiently the whole failure region because the components of the offspring samples prior to the first-passage time are identical to the mother conditional failure sample. Another drawback is that SubSim/Splitting does not apply to problems with uncertain system parameters. This is because, for a given excitation, the location of the first-passage point may depend on the system parameters. Note that SubSim/MCMC does not have this issue since the MCMC procedure is still valid for non-causal systems. These drawbacks are resolved by SubSim/Hybrid, which combines MCMC and splitting for generating the next sample.

### 2.3. Subset Simulation/Hybrid

Subset Simulation with a hybrid strategy, referred to as SubSim/Hybrid, explores the failure region prior to first-passage times using an MCMC method while taking advantage of causality by using trajectory splitting after the first-passage time. To generate the next sample from the current one, a candidate sample of the excitation and uncertain system parameters, if any, is first generated using the M–H algorithm and the corresponding response $Y$ is computed. If it lies in the failure region, the next sample is obtained by splitting the excitation corresponding to this candidate sample at the first-passage time with the system parameters fixed at the candidate values. Otherwise, the next sample is obtained by splitting the excitation of the current sample with the system parameters fixed at the current values. In either case, the response $Y$ for the next sample must be evaluated.

SubSim/Hybrid requires more than one response evaluation for generating the next conditional sample from the current one. If the response is always computed for the whole time history, then SubSim/Hybrid requires two response evaluations per conditional sample. On the other hand, if provision is made in the software so that the response is computed only up to or after the first-passage time, it can be argued that the equivalent average number of response evaluations per conditional sample at the $i$th conditional level is roughly given by $1 + (1 - E[T_f|F_i]/T)P_R$, where $T$ and $E[T_f|F_i]$ are the total duration of study and conditional expected first-passage time, respectively, and $P_R$ is the candidate rejection probability in the MCMC procedure. For example, if $E[T_f|F_i]/T = 0.4$ and $P_R = 0.5$, then SubSim/Hybrid requires on average 1.30 evaluations per conditional sample. This equivalent number of evaluations is considered in this study when assessing the computational efficiency of the method.

Since SubSim/Hybrid is based on SubSim/Splitting, it also requires a causal system and the condition that the current excitation conditioned on the past excitation can be easily sampled. Also note that since the MCMC procedure is employed before the first-passage points for SubSim/Hybrid, it is applicable for dynamical systems with uncertain system parameters.

## 3. Benchmark reliability problems

The three Subset Simulation methods are applied to the reliability benchmark problems defined in [6]. This section briefly describes the cases under study.

### 3.1. Problem 1 – An embankment dam

In Problem 1, the reliability of a dam with random soil properties subjected to deterministic gravity load is analyzed. Three cases, corresponding to different values of $c$ (cohesive strength) and $\phi$ (friction angle), are studied and they are denoted by Prob.1.1, Prob.1.2 and Prob.1.3 in Table 1.

Table 1
Cases for Problem 1

| Ref. | $c$ (kPa) | $\phi$ |
|---|---|---|
| Prob.1.1 | 125 | 30° |
| Prob.1.2 | 225 | 22° |
| Prob.1.3 | 150 | 40° |

Table 2
Cases for Problem 2

| Uncertainty/Nonlinearity | Failure defined w.r.t. | Response threshold level (mm) | Ref. |
|---|---|---|---|
| Deterministic Nonlinear | First story | 0.047 | Prob.2.1.1.1 |
| | | 0.052 | Prob.2.1.1.2 |
| | Top story | 0.024 | Prob.2.1.2.1 |
| | | 0.028 | Prob.2.1.2.2 |
| Random Linear | First story | 0.057 | Prob.2.2.1.1 |
| | | 0.073 | Prob.2.2.1.2 |
| | Top story | 0.013 | Prob.2.2.2.1 |
| | | 0.017 | Prob.2.2.2.2 |
| Random Nonlinear | First story | 0.050 | Prob.2.3.1.1 |
| | | 0.058 | Prob.2.3.1.2 |
| | Top story | 0.024 | Prob.2.3.2.1 |
| | | 0.033 | Prob.2.3.2.2 |

### *3.2. Problem 2 – MDOF Duffing-type oscillator*

In Problem 2, the first-passage probability of the inter-story drift of a ten-degree-of-freedom Duffing-type oscillator subjected to stochastic excitation is considered. Three cases are treated, corresponding to different structural nonlinearity and uncertainties. In Case 1, the system is nonlinear with deterministic properties. In Case 2, the system is linear but its mass, stiffness and damping properties are unknown and their uncertain values are described by specified probability distributions. Case 3 is derived from Case 2 with system nonlinearity as in Case 1. These three cases are indexed by Prob.2.1, Prob.2.2 and Prob.2.3 in Table 2. Each case branches into two sub-cases that consider failure of either the first or the top story. In each sub-case, two threshold levels of failure are considered. As an example of the labeling convention, Prob.2.1.2.1 corresponds to Problem 2, Case 1 (deterministic-nonlinear system), for failure of the top story over the first specified threshold level. The cases are summarized in Table 2.

### *3.3. Problem 3 – MDOF bilinear shear building*

In Problem 3, the first-passage failure of a five-story shear building with bilinear inelastic inter-story stiffness is considered. The stochastic excitation is represented in the spectral domain by 200 random variables representing random spectral amplitudes. The system properties, including mass, damping and stiffness, are uncertain. Two cases are considered, one corresponding to failure at the first story and the other failure at the top story. In each case, failure of two threshold levels is considered. Each scenario is thus identified by three indices. For example, Prob.3.1.2 denotes Problem 3, Case 1 (failure at first story), over the second specified threshold level. The cases are summarized in Table 3.

## 4. Results

SubSim/MCMC is applied to all three benchmark problems as it is applicable to general reliability problems. SubSim/Splitting is applied to Problem 2 Case 1 only, because it is only applicable to deterministic dynamical systems. SubSim/Hybrid is applied to Problem 2 only, since it is only applicable to dynamical systems but it can handle both deterministic and uncertain systems.

Table 3
Cases for Problem 3

| Failure defined w.r.t. | Response threshold level (mm) | Ref. |
|---|---|---|
| First story | 0.032 | Prob.3.1.1 |
| | 0.039 | Prob.3.1.2 |
| Top story | 0.022 | Prob.3.2.1 |
| | 0.032 | Prob.3.2.2 |

In all the problems, Subset Simulation is applied with a conditional failure probability at each level equal to $p_0 = 0.1$ and with the number of samples set to $N = 500$ at each conditional level. The number of conditional levels is chosen to cover the required response level whose failure probability is estimated. For SubSim/MCMC and SubSim/Hybrid, the proposal PDF for each uncertain parameter is chosen as a uniform PDF centered at the current sample with width equal to twice of its standard deviation.

Table 4 summarizes the results for the three variants of Subset Simulation. For all cases, 50 independent runs have been carried out, from which the sample average value and coefficient of variation (c.o.v.) of failure probability can be obtained. The former is reported under the column titled "$P(F)$" and the latter under the column titled "$\delta$". It should be noted that $\delta$ is representative of the c.o.v. of the failure probability for any run and so the c.o.v. of $P(F)$ in the table (an average value from 50 runs) is $\delta/\sqrt{50}$. The total number of samples used for each run is denoted by $N_T$ in the table. For SubSim/MCMC, $N_T = N + N(1 - p_0)(m - 1)$ where $m$ is the number of simulation levels needed to cover the failure probability of interest; e.g., $m = 3$ for Problem 1.1 because $P(F) = 1.6 \times 10^{-3} > (0.1)^3 = p_0^m$. For SubSim/Splitting and SubSim/Hybrid, $N_T$ is the equivalent number of samples used in a single run (averaged over 50 runs) accounting for the fact that for a new conditional sample generated by splitting, only the response time history after the first-passage time needs to be computed. Since any stochastic algorithm for estimating $P(F)$ has a c.o.v. of the form $\delta = \Delta/\sqrt{N_T}$, the 'unit c.o.v.' $\Delta$ presented in Table 4 gives a measure of efficiency that is inherent to the algorithm, that is, it is in theory invariant to the accuracy achieved and the computational effort spent. Of course, smaller values of $\Delta$ correspond to higher computational efficiency. Note that, although the c.o.v. $\delta$ and the unit c.o.v. $\Delta$ are estimated empirically here by their sample values from repeated simulation runs, approximate expressions for $\delta$

Table 4
Benchmark study results

| Problem | MCS[a] | | | | SubSim/MCMC | | | | SubSim/Splitting | | | | SubSim/Hybrid | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $P(F)$ | $N_T$ | $\delta$[b] | $\Delta$ | $P(F)$[c] | $\delta$ | $N_T$ | $\Delta$ | $P(F)$[c] | $\delta$ | $N_T$ | $\Delta$ | $P(F)$[c] | $\delta$ | $N_T$ | $\Delta$ |
| 1.1 | 1.7e−3 | 107.602 | 0.07 | 24 | 1.6e−3 | 0.37 | 1400 | 14 | | | | | | | | |
| 1.2 | 2.2e−5 | 401.000 | 0.34 | 213 | 4.0e−5 | 0.85 | 2300 | 41 | | | | | | | | |
| 1.3 | 3.4e−4 | 247.375 | 0.11 | 54 | 3.1e−4 | 0.59 | 1850 | 25 | | | | | | | | |
| 2.1.1.1 | 1.5e−4 | 74.650e6 | 0.01 | 82 | 1.6e−4 | 0.48 | 1850 | 21 | 2.0e−4 | 0.41 | 1295 | 15 | 2.0e−4 | 0.33 | 2128 | 15 |
| 2.1.1.2 | 5.4e−6 | 74.650e6 | 0.05 | 430 | 6.5e−6 | 0.86 | 2750 | 41 | 8.9e−6 | 0.64 | 1925 | 28 | 9.0e−6 | 0.45 | 3163 | 25 |
| 2.1.2.1 | 4.6e−5 | 74.650e6 | 0.02 | 147 | 4.7e−5 | 0.89 | 2300 | 43 | 5.6e−5 | 0.69 | 1380 | 26 | 6.2e−5 | 0.36 | 2645 | 19 |
| 2.1.2.2 | 3.9e−6 | 74.650e6 | 0.06 | 506 | 3.7e−6 | 0.91 | 2750 | 48 | 5.3e−6 | 1.15 | 1925 | 50 | 5.7e−6 | 0.40 | 3163 | 22 |
| 2.2.1.1 | 1.1e−4 | 29.750e6 | 0.02 | 95 | 1.2e−4 | 0.77 | 1850 | 33 | | | | | 1.1e−4 | 0.41 | 2128 | 19 |
| 2.2.1.2 | 8.1e−7 | 29.750e6 | 0.21 | 1111 | 1.0e−6 | 0.99 | 2750 | 52 | | | | | 1.1e−6 | 0.77 | 3163 | 43 |
| 2.2.2.1 | 4.9e−5 | 29.750e6 | 0.03 | 143 | 6.6e−5 | 0.58 | 2300 | 28 | | | | | 5.9e−5 | 0.46 | 2645 | 24 |
| 2.2.2.2 | 2.5e−7 | 29.750e6 | 0.39 | 2000 | 4.7e−7 | 0.78 | 2750 | 41 | | | | | 3.2e−7 | 0.74 | 3680 | 45 |
| 2.3.1.1 | 9.3e−5 | 45.400e6 | 0.01 | 104 | 1.1e−4 | 0.44 | 1850 | 19 | | | | | 1.1e−4 | 0.39 | 2128 | 18 |
| 2.3.1.2 | 1.3e−6 | 45.400e6 | 0.13 | 877 | 1.5e−6 | 0.73 | 2750 | 38 | | | | | 1.8e−6 | 0.60 | 3163 | 34 |
| 2.3.2.1 | 8.9e−5 | 45.400e6 | 0.01 | 106 | 1.1e−4 | 0.59 | 1850 | 25 | | | | | 1.2e−4 | 0.36 | 2128 | 17 |
| 2.3.2.2 | 6.2e−7 | 45.400e6 | 0.17 | 1270 | 1.3e−6 | 0.88 | 2750 | 46 | | | | | 1.3e−6 | 0.50 | 3163 | 28 |
| 3.1.1 | 6.0e−5 | 122.000e6 | 0.01 | 129 | 1.7e−5 | 0.61 | 2300 | 29 | | | | | | | | |
| 3.1.2 | 1.2e−6 | 122.000e6 | 0.08 | 913 | 1.8e−7 | 1.05 | 2750 | 55 | | | | | | | | |
| 3.2.1 | 2.4e−4 | 122.000e6 | 0.01 | 65 | 2.2e−4 | 0.41 | 1850 | 18 | | | | | | | | |
| 3.2.2 | 9.0e−7 | 122.000e6 | 0.10 | 1054 | 7.2e−7 | 0.71 | 3200 | 40 | | | | | | | | |

[a] Values from [6].

[b] Based on $\delta = \sqrt{[1 - P(F)]/P(F)N_T}$.

[c] Average value from 50 independent runs, each with a c.o.v. of $\delta$.

and $\Delta$ based on information from a single run are available for SubSim/MCMC [8], SubSim/Splitting [13] and SubSim/Hybrid [14].

Results presented in Table 4 for direct Monte Carlo simulation (MCS) are from [6]. Most of the MCS results are produced with high accuracy in terms of the c.o.v. and so they serve as a benchmark for comparison. From Table 4, it is seen that the average values of failure probability produced by SubSim/MCMC mostly agree with the results of MCS to within one or two standard deviations but there are a few exceptions. It may be that the c.o.v. for SubSim/MCMC has been underestimated in these cases but the reason for these discrepancies has not been resolved.

The average values of the failure probability computed by SubSim/Splitting and SubSim/Hybrid generally agree with SubSim/MCMC to within their estimation error. Exceptions are Problems 2.1.1.1 to 2.1.2.2 and Problem 2.2.2.2. The results of SubSim/Splitting and SubSim/Hybrid generally agree, although they both tend to overestimate the failure probabilities for Problems 2.1.1.1 to 2.1.2.2. It is doubtful whether such discrepancies are due to statistical error, theoretical bias (recalling that Subset Simulation is only asymptotically biased), different parameters used, a combination of these factors, or even programming bugs in the software implementing the algorithm. This last factor cannot be totally eliminated, although some consistency checks have been done and it is noted that these discrepancies did not appear in the results for the benchmark problems as originally defined in Communication 3. For reference, the results of SubSim/MCMC were computed by the first co-author (SKA) and those of SubSim/Splitting and SubSim/Hybrid were computed by the second co-author (JC) on independent computing systems using Matlab as a computational platform.

Fig. 1 plots the unit c.o.v. Δ versus failure probability $P(F)$ for SubSim/MCMC for the different cases studied. It is seen that the unit c.o.v. for SubSim/MCMC varies roughly in a logarithmic manner, i.e. $\Delta$ is approximately proportional to $\log[1/P(F)]$, while for direct Monte Carlo Simulation (MCS), it grows drastically as $\Delta \sim 1/\sqrt{P(F)}$ for small $P(F)$. The logarithmic character is a direct result of solving a rare-event simulation problem by a series of frequent conditional failure events. Whereas the slope of the logarithmic trend may be problem dependent, the logarithmic character holds for all the problems studied, suggesting that it is robust with respect to the type of applications; in fact, a theoretical analysis shows that $\Delta$ is approximately proportional to $\log[1/P(F)]$ [15]. SubSim/MCMC appears to perform better in Problem 3 (shown with squares) than in Problem 1 (shown with circles) as the c.o.v. for the former are generally lower for similar failure probability levels.

Fig. 2 compares the results among the three variants of Subset Simulation for Problem 2. Similar to SubSim/MCMC, the unit c.o.v. for SubSim/Splitting and SubSim/Hybrid grow in a logarithmic fashion with decreasing failure probability. The splitting and hybrid algorithms lead to lower values of unit c.o.v., i.e., higher efficiency, with only one exception for SubSim/Splitting. Of course, this is at the expense of their generality in applications since SubSim/Splitting and SubSim/Hybrid are limited to first-passage reliability problems of dynamical systems.

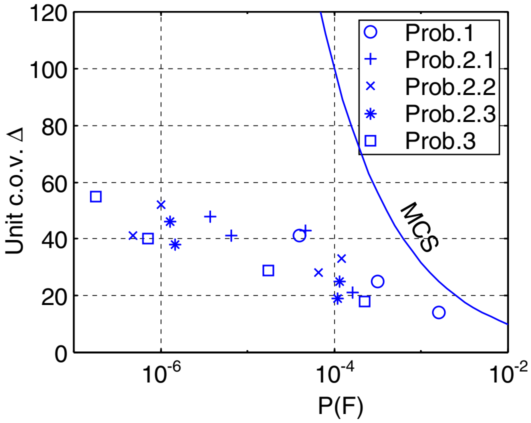

Fig. 1. Performance of SubSim/MCMC for all problems.

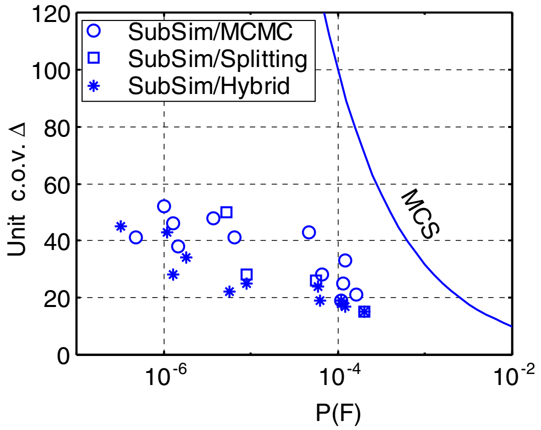

Fig. 2. Performance among variants of Subset Simulation in Problem 2.

## 5. Conclusions

Three variants of Subset Simulation are applied to three reliability benchmark problems. The original version of Subset Simulation based on the Metropolis–Hasting algorithm is applicable to all three problems. The two variants of Subset Simulation that utilize splitting, demonstrate that efficiency can be gained by incorporating knowledge about system characteristics into the simulation procedure. For all three variants, the unit c.o.v. varies logarithmically with failure probability and they therefore lead to substantial gains in computational efficiency over direct Monte Carlo simulation. This logarithmic trend is a direct consequence of converting a rare-event simulation problem into a series of problems with more frequent failure events.

## Acknowledgements

The first co-author (SKA) wishes to acknowledge financial support from Grant SU4/02 by the Nanyang Technological University, Singapore. The last two co-authors (JC, JLB) wish to acknowledge the financial support from the California Institute of Technology, especially the George W. Housner Post-doctoral Fellowship (JC) in Civil Engineering.

## References

[1] Schueller GI, Stix R. A critical appraisal of methods to determine failure probabilities. Struct Safety 1987;4:293–309.

[2] Engelund S, Rackwitz R. A benchmark study on importance sampling techniques in structural reliability. Struct Safety 1993;12:255–76.

[3] Schueller GI. Computational stochastic mechanics: recent advances. Comput Struct 2001;79:2225–34.

[4] Rackwitz R. Reliability analysis – a review and some perspectives. Struct Safety 2001;23:365–95.

[5] Schueller GI, Pradlwarter HJ, Koutsourelakis PS. A critical appraisal of reliability estimation procedures for high dimensions. Probabilistic Eng Mech 2004;19:463–74.

[6] Schueller GI, Pradlwarter HJ. Benchmark study on reliability estimation in higher dimensions of structural systems - an overview. Structural Safety 2006. See also: Available from: <http://mechanik.uibk.ac.at/Benchmark_Reliability>.

[7] Au SK. On the Solution of First Excursion Problems by Simulation with Applications to Probabilistic Seismic Performance Assessment. PhD Thesis in Civil Engineering 2001, Division of Engineering and Applied Science, California Institute of Technology, California, USA.

[8] Au SK, Beck JL. Estimation of small failure probabilities in high dimensions by Subset Simulation. Probabilistic Eng Mech 2001;16(4):263–77.

[9] Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH. Equations of state calculations by fast computing machines. J Chem Phys 1953;21(6):1087–92.

[10] Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 1970;57:97–109.

[11] Robert C, Casella G. Monte Carlo Statistical Methods. New York: Springer; 1999.

[12] Malakoff D. Bayes offers a ‘new’ way to make sense of numbers. Science 1999;286:460–1464.

[13] Ching J, Au SK, Beck JL. Reliability estimation for dynamical systems subject to stochastic excitation using Subset Simulation with splitting. Comput Meth Appl Mech Eng 2005;194:1557–79.

[14] Ching J, Beck JL, Au SK. Hybrid Subset Simulation method for reliability estimation of dynamical systems subjected to stochastic excitation. Probabilistic Eng Mech 2005;20(3):199–214.

[15] Au SK, Beck JL. Subset Simulation and its application to probabilistic seismic performance assessment. J Eng Mech ASCE 2003;129(8):1–17.

[16] Pradlwarter HJ, Schueller GI. Assessment of low probability events of dynamical systems by controlled Monte Carlo simulation. Probabilistic Eng Mech 1999;14:213–27.

[17] Melnik-Melnikov PG, Dekhtyaruk ES. Rare events probabilities estimation by Russian Roulette and Splitting simulation technique. Probabilistic Eng Mech 2000;15(2):125–9.