

Hamiltonian Monte Carlo-Subset Simulation (HMC-SS) methods for structural reliability analysis

Marco Broccardo

Senior Researcher, ETH Zurich, Zurich, Switzerland

Ziqi Wang

Assistant Professor, Guangzhou University, Guangzhou, China

Junho Song

Professor, Seoul National University, Seoul, South Korea

ABSTRACT: In this study, we carefully analyze the most recent advancements in Hamiltonian Monte Carlo methods combined with Subset Simulation (HMC-SS) in the context of structural reliability analysis. The HMC method employs Hamiltonian dynamic to sample from a target probability distribution. In contrast to the standard Markov-Chain Monte Carlo methods (e.g., Gibbs or Metropolis-Hastings techniques), HMC alleviates the burn-in phase and the random walk behavior to achieve a more effective exploration of the target probability distribution. This turns out to be important in high-dimensional spaces (e.g., when the number of random variables is high), where the bulk of probability content concentrates in the so-called typical sets. The structure of the paper is as follows. We first briefly review the Subset Simulation and the general concepts of HMC. Following, in both standard Gaussian and non-Gaussian probability spaces, we present a series of complex structural reliability problems to test in practice the validity of the method. Finally, we conclude with a series of future developments and directions.

A powerful variance-reduction Monte Carlo method which has been widely used in reliability analysis is Subset Simulation (Au & Beck 2001). The method expresses the failure domain of interest as the intersection of a sequence of nested intermediate failure domains, and the failure probability of interest is expressed as a product of conditional probabilities associated with the intermediate failure domains. Since the conditional probabilities are significantly larger than the target failure probability the computational cost of Subset Simulation is significantly lower than the crude MCS method. The challenge of the scheme, which consists of evaluating the intermediate conditional probabilities, is overcome by using efficient Markov Chain Monte Carlo (MCMC) methods.

A crucial step in Subset Simulation is to obtain random samples according to a sequence of probability distributions that are conditional on

nested intermediate failure domains. The efficiency and accuracy of Subset Simulation is directly affected by those of the MCMC algorithm used to produce random samples representing the conditional distributions in the sequence. In the current practice of Subset Simulation, the commonly used MCMC methods are based on random walks, which suffer issues of inefficiency when performing high dimensional random sampling (particularly in non-Gaussian spaces). Most recently, a non-random-walk MCMC method, namely the Hamiltonian Monte Carlo (HMC) method (Duane et.al 1987; Neal 2011), is introduced to the framework of Subset Simulation for efficient reliability analysis in both Gaussian and non-Gaussian spaces (Broccardo et.al 2018; Wang et.al 2019). The HMC method employs a deterministic mechanism inspired by Hamiltonian dynamics to propose samples for a target probability distribution. The method alleviates the

random-walk behavior to achieve a more effective and consistent exploration of the probability space compared to standard Gibbs or Metropolis-Hastings techniques.

This paper provides a review of the HMC based-Subset Simulation (HMC-SS) method in the Gaussian and non-Gaussian spaces, and the recent development of HMC-SS in conjunction with Gaussian process metamodeling. The structure of this paper is as follows. Sections 1 introduce essential elements of the original HMC-SS; Section 2 shows HMC-SS equipped with Gaussian process metamodeling; Section 3 shows a series of numerical examples; Section 4 concludes the study with a summary of the results.

1. HMC-SS

In reliability analysis, the failure probability of a system with basic random variables $\mathbf{x} \in \mathbb{R}^n$ can be expressed by an integral,

$$P_f = \int_{\mathbb{R}^n} I_{\mathcal{F}}(\mathbf{x})\pi(\mathbf{x})d\mathbf{x} \quad (1)$$

where $I_{\mathcal{F}}(\cdot)$ is a binary indicator function which gives ‘1’ if point \mathbf{x} is within the failure domain, and ‘0’ otherwise, and $\pi(\mathbf{x})$ is the joint probability density function (PDF) of \mathbf{x} .

The Subset Simulation solution of Eq. (1) involves the construction of a sequence of nested intermediate failure domains, so that the failure domain of interest, \mathcal{F} , is expressed by

$$\mathcal{F} = \bigcap_{j=1}^M \mathcal{F}_j \quad (2)$$

where $\mathcal{F}_1 \supset \mathcal{F}_2 \supset \dots \supset \mathcal{F}_M$, and $\mathcal{F} = \mathcal{F}_M$. The failure probability $P_f = \Pr(\mathbf{x} \in \mathcal{F})$ can be written as

$$\Pr(\mathbf{x} \in \mathcal{F}) = \prod_{j=1}^M \Pr(\mathbf{x} \in \mathcal{F}_j | \mathbf{x} \in \mathcal{F}_{j-1}) \quad (3)$$

where $\Pr(\mathbf{x} \in \mathcal{F}_0) = 1$. Each $\Pr(\mathbf{x} \in \mathcal{F}_j | \mathbf{x} \in \mathcal{F}_{j-1})$ in Eq. (3) can be computed using

$$\begin{aligned} \Pr(\mathbf{x} \in \mathcal{F}_j | \mathbf{x} \in \mathcal{F}_{j-1}) \\ = \int_{\mathbb{R}^n} I_{\mathcal{F}_j}(\mathbf{x})\pi(\mathbf{x} | \mathcal{F}_{j-1})d\mathbf{u} \end{aligned} \quad (4)$$

where $\pi(\mathbf{x} | \mathcal{F}_{j-1})$ is the conditional/truncated multivariate PDF. Using an MCMC technique to generate samples of $\pi(\mathbf{x} | \mathcal{F}_{j-1})$, Eq. (4) can be evaluated MCS, i.e.

$$\Pr(\mathbf{x} \in \mathcal{F}_j | \mathbf{x} \in \mathcal{F}_{j-1}) \cong \frac{1}{N} \sum_{i=1}^N I_{\mathcal{F}_j}(\mathbf{x}_i) \quad (5)$$

in which \mathbf{x}_i are samples generated from conditional PDF $\pi(\mathbf{x} | \mathcal{F}_{j-1})$. In implementations of Subset Simulation, the nested failure domains are chosen adaptively such that $\Pr(\mathbf{x} \in \mathcal{F}_j | \mathbf{x} \in \mathcal{F}_{j-1})$, $j = 1, 2, \dots, M - 1$, approximately equals to a specified percentile p_0 .

The HMC method provides an efficient way to draw random samples from the conditional distribution $\pi(\mathbf{x} | \mathcal{F}_j)$, so that Eq. (5) can be estimated effectively. In HMC method the trajectories of a Hamiltonian system are used as the proposal samples of the target distribution. Specifically, in HMC the following Hamilton’s equations need to be solved.

$$\begin{aligned} \frac{d\mathbf{q}}{dt} &= \frac{\partial H}{\partial \mathbf{p}} \\ \frac{d\mathbf{p}}{dt} &= -\frac{\partial H}{\partial \mathbf{q}} \end{aligned} \quad (6)$$

in which the Hamiltonian $H = H(\mathbf{q}, \mathbf{p})$ is a constant corresponds to the total energy of the system, and thus H is independent of time evolutions of (\mathbf{q}, \mathbf{p}) . The Hamiltonian $H(\mathbf{q}, \mathbf{p})$ can be expressed by

$$H(\mathbf{q}, \mathbf{p}) = V(\mathbf{q}) + K(\mathbf{p}) \quad (7)$$

where $V(\mathbf{q})$ is the potential energy, which is a function of the position vector \mathbf{q} alone, and $K(\mathbf{p})$ is the kinetic energy, which is a function of the momentum vector \mathbf{p} alone. Given initial values for the position and momentum, Eq.(7) completely defines the energy level for the system. Then, the solution of the Hamilton’s

equations, Eq. (6), describes an equi-Hamiltonian trajectory of the system in the phase space.

To establish a connection between the probability space of interest, and a mathematically equivalent Hamiltonian system described by Eq. (6), first the outcome \mathbf{x} is viewed as the position \mathbf{q} of a Hamiltonian system (i.e. $\mathbf{q} \equiv \mathbf{x}$). Next, a set of auxiliary random momentum variables, \mathbf{p} , which has the same dimension as \mathbf{q} , are introduced to expand the original position space, so that now one has the position-momentum phase space of a Hamiltonian system. Finally, to incorporate the probabilistic structure of $\pi(\mathbf{q})$ into the Hamiltonian system, the potential energy $V(\mathbf{q})$ is defined in terms of the target PDF $\pi(\mathbf{q})$ as

$$V(\mathbf{q}) \equiv -\log \pi(\mathbf{q}) \quad (8)$$

The form for kinetic energy $K(\mathbf{p})$ could vary with implementation, but it is typically defined as

$$K(\mathbf{p}) \equiv \frac{\mathbf{p}\mathcal{M}^{-1}\mathbf{p}}{2} \quad (9)$$

where \mathcal{M} is a positive-definite and symmetric ‘mass’ matrix. Typically, \mathcal{M} is chosen as a scalar multiple of the identity matrix. The joint PDF of (\mathbf{q}, \mathbf{p}) is defined as

$$\pi(\mathbf{q}, \mathbf{p}) \equiv \frac{1}{Z} e^{-H(\mathbf{q}, \mathbf{p})} = \frac{1}{Z} e^{-V(\mathbf{q})} e^{-K(\mathbf{p})} \quad (10)$$

where Z is a normalizing constant. Substituting Eq. (8) and Eq. (9) into Eq. (10), one obtains

$$\pi(\mathbf{q}, \mathbf{p}) = \frac{1}{Z} \pi(\mathbf{q}) e^{-\frac{\mathbf{p}\mathcal{M}^{-1}\mathbf{p}}{2}} \quad (11)$$

The above definition of the joint PDF $\pi(\mathbf{q}, \mathbf{p})$ implies: a) the position and the momentum variables are statistically independent; and b) the position is distributed following the original target distribution $\pi(\mathbf{q})$, and the momentum is distributed as a multivariate Gaussian distribution. The aforementioned two properties of $\pi(\mathbf{q}, \mathbf{p})$ implies that once it is possible sample from $\pi(\mathbf{q}, \mathbf{p})$, samples distributed as $\pi(\mathbf{q})$ are obtained by simply projecting out the momentum component of $\pi(\mathbf{q}, \mathbf{p})$ samples. sample from $\pi(\mathbf{q}, \mathbf{p})$. In

particular, HMC sampling can be divided into two main steps. In the first step, the momentum is sampled from the canonical distribution; this together with the current position completely defines an equi-Hamiltonian surface. In the second step, both position and momentum variables change within the equi-Hamiltonian surface by integrating Eq. (6) for a given time t_f . The conceptual procedure of HMC is described as follows.

Algorithm 1. Conceptual procedure of Hamiltonian Monte Carlo method

Step 1. Generate a random momentum \mathbf{p} according to PDF $e^{-K(\mathbf{p})}/Z$.

Step 2. Use the momentum \mathbf{p} and the position \mathbf{q} of a seed sample as initial conditions, propose a new state $(\mathbf{q}^*, \mathbf{p}^*)$ via solutions of the Hamilton’s equations at a time point t_f .

Step 3. Negate the proposed momentum, i.e., $\mathbf{p}^* \leftarrow -\mathbf{p}^*$

In standard Gaussian space, the Hamiltonian $H(\mathbf{u}, \mathbf{p})$ (\mathbf{u} denotes standard Gaussian random variables) can be written as

$$\begin{aligned} H(\mathbf{u}, \mathbf{p}) &= V(\mathbf{u}) + K(\mathbf{p}) \\ &= -\log \varphi(\mathbf{u}|\mathcal{F}_j) + \frac{\mathbf{p}^T \mathbf{p}}{2} \\ &= \frac{\mathbf{u}^T \mathbf{u}}{2} + \frac{\mathbf{p}^T \mathbf{p}}{2} - \log I_{\mathcal{F}_j}(\mathbf{u}) + \text{const} \end{aligned} \quad (12)$$

The constant term in Eq. (12) can be dropped since it leaves Hamilton’s equations intact. The term $-\log I_{\mathcal{F}_j}(\mathbf{u})$ introduces a *potential barrier* to the system, so that proposals outside the failure domain \mathcal{F}_j have infinite potential energy, i.e., areas outside \mathcal{F}_j cannot be reached by the Hamiltonian system.

As long as the trajectories of the Hamiltonian system lie in the failure domain \mathcal{F}_j , the term $-\log I_{\mathcal{F}_j}(\mathbf{u})$ in Eq. (12) is zero, and the Hamiltonian system has an analytical solution expressed by

$$\begin{aligned} \mathbf{u}(t) &= \mathbf{p}_{init} \sin t + \mathbf{u}_{init} \cos t \\ \mathbf{p}(t) &= \mathbf{p}_{init} \cos t - \mathbf{u}_{init} \sin t \end{aligned} \quad (13)$$

where \mathbf{p}_{init} and \mathbf{u}_{init} denote initial momentum and initial position, respectively. Using the analytical solution Eq. (13), efficient algorithms of HMC-SS for Gaussian space can be developed, implementations details of the algorithms can be seen in Wang et.al 2019.

For non-Gaussian space simulations, the Hamiltonian is written as

$$\begin{aligned} H(\mathbf{q}, \mathbf{p}) &= V(\mathbf{q}) + K(\mathbf{p}) \\ &= -\log \pi(\mathbf{q}|\mathcal{F}_j) \\ &\quad + \frac{\mathbf{p}^T \mathcal{M}^{-1} \mathbf{p}}{2} \\ &= -\log \pi(\mathbf{q}) + \frac{\mathbf{p}^T \mathcal{M}^{-1} \mathbf{p}}{2} \\ &\quad - \log I_{\mathcal{F}_j}(\mathbf{q}) + \text{const} \end{aligned} \quad (14)$$

It is seen that Eq. (14) has the same form as Eq. (12), thus similar algorithms can be used to sample from $\pi(\mathbf{q}|\mathcal{F}_j)$. The only difference is that instead of analytical solutions, one may need a numerical integration technique to solve the Hamilton's equations. One could use well-known methods (e.g. *leapfrog method*) to numerically solve Hamilton's equations. Due to the numerical errors introduced by numerical methods, a Metropolis accept-reject rule needs to be applied after the proposal:

$$\begin{aligned} \min[1, \exp(-H(\mathbf{q}^*, \mathbf{p}^*) + H(\mathbf{q}, \mathbf{p}))] = \\ \min[1, \exp(-V(\mathbf{q}^*, \mathbf{p}^*) + V(\mathbf{q}, \mathbf{p}) \\ - K(\mathbf{q}^*, \mathbf{p}^*) + K(\mathbf{q}, \mathbf{p}))]. \end{aligned} \quad (15)$$

Note that if the Hamilton's equations can be solved analytically (e.g. Eq. (13) for Gaussian space simulation), Eq. (15) is not needed.

2. HMC-SS EQUIPPED WITH GAUSSIAN PROCESS METAMODELING

Although HMC-SS has high efficiency in performing rare event simulations compared with the crude MCS approach, for problems with complex model functions, the evaluation of thousand times (which is typical in HMC-SS) of model functions can be impractical. To further enhance the efficiency of HMC-SS for problems with complex model functions, a HMC-SS

combined with active learning based Gaussian process metamodeling is developed in Broccardo et.al 2018. The main algorithm of the approach is summarized as follows.

Algorithm 2. Active learning HMC-SS

Step 1. A test set $S = [\mathbf{x}_1, \dots, \mathbf{x}_s]$ is generated by a variation of the HMC-SS method (iso-HMC-SS), which provides an iso-density sampling scheme.

From S , an initial experimental design, $\mathcal{X} = [\mathbf{x}_1, \dots, \mathbf{x}_d]$ with $d \ll s$ is subsampled. Exact outcomes of the computational model are computed for each sample of the experimental design, i.e. $\mathcal{Y} = [\mathcal{M}(\mathbf{x}_1), \dots, \mathcal{M}(\mathbf{x}_d)]$

Step 2. The Gaussian process metamodel is defined based on the input output pairs $\mathcal{X} \sim \mathcal{Y}$.

Step 3. A three-fold estimate of the probability of failure, \hat{P}_f^+ , \hat{P}_f^0 , \hat{P}_f^- is computed via HMC-SS based on the metamodel.

Step 4. The sample $\mathbf{x}^* \in S$ which maximize a given learning criterion is added to the experimental design, i.e. $\mathcal{X} \leftarrow [\mathcal{X}; \mathbf{x}^*]$; and the corresponding exact solution $\mathcal{M}(\mathbf{x}^*)$ is added to \mathcal{Y} such that $\mathcal{Y} \leftarrow [\mathcal{Y}; \mathcal{M}(\mathbf{x}^*)]$. Next a new test set, S , is generated via iso-HMC-SS. Finally, the algorithm returns to Step 2 for an update of the metamodel.

Step 5. The iterations are terminated when a stopping criterion is met.

3. NUMERICAL EXAMPLES

3.1. Banana distribution sampling

Consider a "banana-shaped" PDF given as follows.

$$\begin{aligned} f_{XY}(x, y) = \frac{1}{Z} \exp \left[-\frac{1}{2(1-\rho^2)} \left(\frac{x^2}{a^2} \right. \right. \\ \left. \left. + a^2 \left(y - b \frac{x^2}{a^2} - ba^2 \right)^2 \right. \right. \\ \left. \left. - 2\rho x \left(y - b \frac{x^2}{a^2} \right. \right. \right. \\ \left. \left. \left. - ba^2 \right) \right) \right] \end{aligned} \quad (16)$$

where Z is a normalizing constant. In this example, we set $a = 1.15$, $b = 0.5$, $\rho = 0.9$,

Figure 1 a) shows the contour plot of the “banana-shaped” distribution.

We use leapfrog based HMC to sample from the banana-shaped distribution, starting with a seed $\mathbf{q}_0 = (4,5)$ at the far tail region of the distribution. The trajectory of 100 HMC iterations are shown in the left plot of Figure 1 b). For comparison, the trajectory of 100 iterations of the Metropolis Hastings algorithm using a uniform transition distribution within a square of width 1 are shown in the right plot of Figure 1 c). One can observe the efficiency of HMC in reaching the bulk of the probability density in only one iteration. This example shows that HMC is particularly suitable for probability densities that are “narrow” and confined in specific region of the space. This is the typical case of high dimensional spaces, where the bulk of probability lies in specific confined regions named typical set.

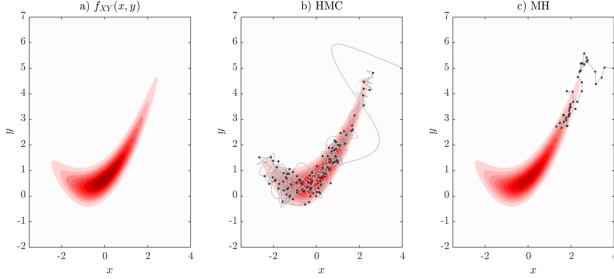


Figure 1. Trajectories obtained from HMC and MH algorithms

3.2. First-passage probability estimation

Consider a single degree of freedom (SDOF) linear oscillator under seismic loading defined by the differential equation

$$m\ddot{X}(t) + c\dot{X}(t) + kX(t) = -m\ddot{U}_g(t) \quad (17)$$

where $X(t)$, $\dot{X}(t)$ and $\ddot{X}(t)$ denote the displacement, velocity and acceleration of the oscillator, respectively. We set the mass $m = 6 \times 10^4$ kg, stiffness $k = 2.0 \times 10^7$ N/m, damping $c = 2m\zeta\sqrt{k/m}$ with the viscous damping ratio $\zeta = 10\%$. The initial natural period of this SDOF oscillator is $T = 0.34$ s. The ground acceleration $\ddot{U}_g(t)$ is modeled by white noise process.

Now we consider the first-passage probability $\Pr \left[\max_{t \in (0,10)} X(\mathbf{u}, t) > x \right]$. The first

passage probabilities for threshold $x = 0.020$ m, $x = 0.025$ m and $x = 0.030$ m are computed using SS, and the results are compared with the solution obtained from crude MCS with 1.0×10^6 runs. Table 1 illustrates the results.

Table 1. Performance of HMC-SS for first-passage problem

Threshold	HMC			CWMH			MCS
	\hat{P}_f	c.o.v	NG	\hat{P}_f	c.o.v	NG	P_f
0.020	6.73×10^{-3}	0.21	2773	6.67×10^{-3}	0.21	2782	6.80×10^{-3}
0.025	7.65×10^{-5}	0.32	4492	7.55×10^{-5}	0.35	4483	8.20×10^{-5}
0.030	2.90×10^{-7}	0.42	6400	3.14×10^{-7}	0.58	6400	-

3.3. Active learning HMC-SS example

In this study, a variation of the four-branch benchmark function (Schobi 2016) is presented in details. The four-branch function is often used in structural reliability analysis to describe a series system with four distinct limit state surfaces, and it can be expressed as follows.

$$\begin{aligned} \mathcal{M}(x) \\ = \min \left\{ \begin{aligned} &c_1 + \kappa(x_1 - x_2)^2 \pm \frac{x_1 + x_2}{\sqrt{2}} \\ &\frac{c_3}{\sqrt{2}} \pm (c_2 x_1 - x_2) \end{aligned} \right\} \quad (18) \end{aligned}$$

where $c_1 = 6.5$, $c_2 = 1.5$, $c_3 = 8.5$, and $\kappa = 0.06$. Moreover, it is assumed $G(x) = \mathcal{M}(x)$. The joint PDF is based on a mixture of two banana shaped distributions (Eq. (16)) with equal weight, i.e.

$$\begin{aligned} f_{XY}(x, y) \\ = \frac{1}{Z} \sum_{i=1}^2 \exp \left[-\frac{1}{2(1-\rho^2)} \left(\frac{x^2}{a^2} \right. \right. \\ \left. \left. + a^2 \left(y - b_i \frac{x^2}{a^2} - b_i a^2 \right)^2 \right. \right. \\ \left. \left. - 2\rho x \left(y - b_i \frac{x^2}{a^2} - b_i a^2 \right) \right) \right] \quad (19) \end{aligned}$$

where $a = 0.9$, $b_1 = 0.4$, $b_2 = -0.4$, $\rho = 0.95$, Figure 2 b) shows the contour plot of the distribution.

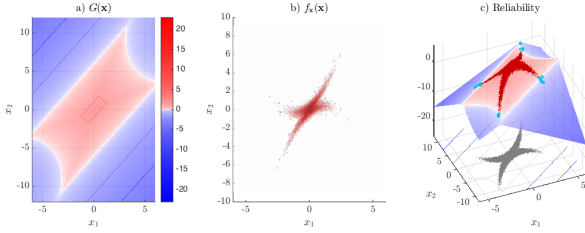


Figure 2: a) The four-branch limit state function. Red safe domain, blue failure domain. b) The mixture of “banana-shaped” distribution, with HMC samples. c) The structural reliability problem.

Table 2 reports the results of a single run of the adaptive algorithm versus the results of a classical MCS with 10^7 samples. Only 30 evaluations (16 of the initial experimental design + 14 iterations) of the original model function are necessary to evaluate the reliability of the system. The c.o.v. of the HMC-SS estimated is computed with 50 repeated runs, after the stopping criterion is met. Therefore, only metamodel evaluations are used and not new evaluation of the model function.

To test the robustness of the metamodel convergence, the full analysis is replicated for 50 times with different initial experimental designs. The results of the analysis are reported in Table 3, compared with the same MCS simulation of Table 2. Despite the results depends on 50 different metamodels, the c.o.v. of the probabilities is comparable to the one of Table 2. This suggests that most of the variability on the estimate is due to the HMC-SS statistical errors.

Table 2: Four branch problem: Estimate of P_f based on a single run of the adaptive design algorithm

Method	\hat{P}_f	c.o.v.
MCS	5.86×10^{-4}	0.013
AL-HMC-SS	5.87×10^{-4}	0.248

Table 3: Four branch problem: Estimate of \hat{P}_f based on 50 runs of the adaptive design algorithm

Method	\hat{P}_f	c.o.v.	Model Evaluations
MCS	5.86×10^{-4}	-	10^7
AL-HMC-SS	5.87×10^{-4}	0.281	16+22.08

4. CONCLUSIONS

In this paper, the HMC-SS method for reliability analysis is reviewed. The HMC-SS method aims to solve reliability problems with complex, high dimensional PDF, complex model (limit-state) function, and low failure probability. The current limitation of the method lies in the sequential nature of the HMC based scheme, which constitutes a drawback for parallel computing. Moreover, only first-order differentiable PDFs are allowed to perform HMC. Futures studies will focus on computational efficiency and parallelization of the proposed scheme.

5. REFERENCES

- Au SK, Beck JL. Estimation of small failure probabilities in high dimensions by subset simulation. Probabilistic Engineering Mechanics 2001;16(4):263–277.
- Duane S, Kennedy A.D, et al. Hybrid Monte Carlo. Physics Letters B 1987; 195(2):216–222.
- Neal R. M. MCMC using Hamiltonian dynamics. Handbook of Markov Chain Monte Carlo 2011; CRC Press, New York.
- Wang Z, Broccardo M, Song J. Wang, Z., M. Broccardo, and J. Song (2018). Hamiltonian monte carlo methods for subset simulation in reliability analysis. Structural Safety 2019; 76: 51-67.
- Broccardo M, Wang Z, Marelli S, Song J, Sudret B. Hamiltonian Monte Carlo-based subset simulation using Gaussian process metamodeling. IFIP8 Reliability and Optimization of Structural Systems, June 26-29, 2018 Zurich.
- Schöbi, R., B. Sudret, and S. Marelli (2016). Rare event estimation using polynomial-chaos kriging. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering 3(2), D4016002.