Highlights

Reliability Sensitivity with Response Gradient

Siu-Kui Au, Zi-Jun Cao

- Reliability sensitivity as expectation of response gradient
- Kernel smoothing resolves conditioning on zero-probability event
- Sensitivity embedded in Monte Carlo run
- Applicable regardless of the type of sensitivity parameter

Reliability Sensitivity with Response Gradient

Siu-Kui Au[a], Zi-Jun Cao[b]

[a]*School of Civil and Environmental Engineering, Nanyang Technological University, Singapore*
[b]*MOE Key Laboratory of High-Speed Railway Engineering, Institute of Smart City and Intelligent Transportation, Southwest Jiaotong University, Chengdu, China*

Abstract

Engineering risk is concerned with the likelihood of failure and the scenarios when it occurs. The sensitivity of failure probability to change in system parameters is relevant to risk-informed decision making. Computing sensitivity is at least one level more difficult than the probability itself, which is already challenged by a large number of input random variables, rare events and implicit nonlinear 'black-box' response. Finite difference with Monte Carlo probability estimates is spurious, requiring the number of samples to grow with the reciprocal of step size to suppress estimation variance. Many existing works gain efficiency by exploiting a specific class of input variables, sensitivity parameters, or response in its exact or surrogate form. For general systems, this work presents a theory and associated Monte Carlo strategy for computing sensitivity using response values and gradients with respect to sensitivity parameters. It is shown that the sensitivity at a given response threshold can be expressed via the expectation of response gradient conditional on the threshold. Determining the expectation requires conditioning on the threshold that is a zero-probability event, but it can be resolved by the concept of kernel smoothing. The proposed method offers sensitivity estimates for all response thresholds generated in a single Monte Carlo run. It is investigated in a number of examples featuring sensitivity parameters of different nature. As response gradient becomes increasingly available, it is hoped that this work can provide the basis for embedding sensitivity calculations with reliability in the same Monte Carlo run.

Keywords: Kernel smoothing, Markov Chain Monte Carlo, Rare event, Reliability sensitivity, Subset Simulation

1. Introduction

Engineering risk has many facets that cover the likelihood of failure, the scenarios when failure occurs, life-cycle costs, etc. Modern systems are increasingly complex, demanding algorithms that are scalable with the number of input random variables and nonlinearity in input/output relationship, and remain feasible for rare events with low probability. See monographs on engineering reliability [1, 2, 3], a recent a review on structural reliability [4], and a journal special issue on methodology for complex systems [5]. Beyond a given scenario, assessing the change of risk measures or expected performance to perturbations in system parameters/assumptions is relevant to risk-informed decision making, e.g., establishing confidence in conclusions, hypothesis testing of assumptions and reliability-based design; see [6] a recent review. The concept is also practiced in non-engineering fields such as finance, e.g., 'stress testing' of credit risk and the 'Greeks' (sensitivity) of stock option price; see [7] a textbook.

This work is concerned with computing reliability sensitivity for systems when the response and its gradient with respect to (w.r.t.) 'sensitivity parameters' are available, but otherwise they can remain as a 'black-box'. Let

$$Y = f(\mathbf{X}, \alpha) \tag{1}$$

be a scalar response of interest that depends on the random input vector $\mathbf{X} \in \mathbb{R}^n$ and a deterministic (scalar) sensitivity parameter α through the response function f. Failure is defined as the exceedance of Y over a given threshold y, with probability

$$F(y, \alpha) = P(Y \geq y) = \int_{f(\boldsymbol{x},\alpha)\geq y} q(\boldsymbol{x})\, d\boldsymbol{x} \tag{2}$$

where q denotes the probability density function (PDF) of $\mathbf{X}$. Although not explicitly indicated, $P(Y \geq y)$ depends implicitly on α through (1). Viewed as a function, $F(\cdot, \alpha)$ is the 'complementary cumulative distribution function' (CCDF) of Y for a given α. Reliability sensitivity is generally concerned with how F changes with α. In this work we focus on the derivative $\partial F/\partial\alpha$.

1.1. Notes on problem context

As is conventional, despite the name 'reliability', F instead of $1 - F$ is considered because the latter is close to 1 in applications, masking significant digits in quantitative analysis or interpretation. Considering a scalar

Y leads to no loss of generality because the critical response of a system comprising sub-systems connected in series or parallel, or any combination, can be defined as the max or min of component responses cascaded in the same manner. Having said that, there are strategies that allow one to generate reliability estimates for multiple responses with less computations than repeated runs [8, 9], but they are out of the present scope. As indicated by the notation, it is assumed that the PDF q does not depend on α. This is merely to standardize the mathematical context. It does not lead to any loss of generality because parameters that affect the PDF can be absorbed into the definition of f. Having said that, there are works that explore efficiency gain by limiting α to 'distribution parameters', i.e., those affecting q only; see literature review later in Section 1.2. The assumption of a scalar α is just to simplify notation and leads to no loss of generality either, because the gradient of F w.r.t. multiple parameters is simply a vector of the individual derivatives as long as the parameters are 'free', i.e., they do not depend on each other and are not subjected to any equality constraint.

1.2. Fundamental challenge and state of the art

From first principles the sensitivity $\partial F/\partial\alpha$ involves the derivative of F w.r.t. α, while F in turn involves a multivariate integral over the failure domain $\{\boldsymbol{x} \in \mathbb{R}^n : f(\boldsymbol{x}, \alpha) > y\}$ that is only implicitly known when $f(\cdot, \alpha)$ is a black-box. Efficient calculation of F for a given α belongs to a reliabilty estimation problem. It involves resolving the integral, e.g., locating the failure domain and accounting for contributions there. The key challenges are 1) high dimension n that renders deterministic search algorithms unsustainable and intuitions in low dimensions inapplicable or misleading; 2) small probability (F) that makes direct Monte Carlo prohibitive and poses a wide knowlege gap for machine-learning to set off from prior knowledge; 3) implicit nonlinear $f(\cdot, \alpha)$ that makes approximation (e.g., Taylor series) or machine-learning difficult, not to mention the cost of each function evaluation. These challenges are well-awared in the reliability literature, and research is still growing. State-of-the-art strategies for general applications are mostly Monte Carlo in nature, i.e., estimation by averaging over pseudo-randomly generated samples of input random variables. There is a large variety of methods, e.g., direct Monte Carlo that is most robust but prohibitive for rare events; importance sampling when a proposal PDF can be constructed with heuristics or domain-knowledge in various forms [10, 11]; Subset Simulation (SS) [12, 13] (see also the end of Section 4.2) that is less

demanding on prior knoweldge, though at the expense of a machine-learning curve from frequent to rare events that can get tougher along the way; Line Sampling [14, 15] that exploits the probability along different lines of directions in the input variable space; and recent improvements with data science tools such as 'Krigging' [16, 17] that originates from geostatistics, active learning [18, 19] that improves the selection of sample points, and neutral network [20, 21] that offers flexiblity for surrogates and sample generation beyond developers' mathematical facility.

1.2.1. Finite difference

Computing $\partial F/\partial\alpha$ is at least one level more difficult than F, as it needs to resolve both the derivative (w.r.t. α) and integral (over $\boldsymbol{x}$). Requesting the sensitivity for multiple values of y, as is possible for F in a single Monte Carlo run, adds to one's plate. One last resort is finite difference, e.g., $\partial F/\partial\alpha \approx [\widetilde{F}(\alpha+\Delta\alpha) - \widetilde{F}(\alpha-\Delta)]/2\Delta\alpha$, where $\widetilde{F}(\alpha\pm\Delta\alpha)$ denotes the Monte Carlo estimates of F at $\alpha\pm\Delta\alpha$ but based on the same set of N_s (say) samples of $\mathbf{X}$, i.e., 'common random numbers' (CRN) [22]. Steming from central difference, the estimate has a bias of $O(\Delta\alpha^2)$. In addition to the bias, for typical situations where $\widetilde{F}(\alpha+\Delta\alpha)$ and $\widetilde{F}(\alpha-\Delta\alpha)$ have a correlation of $O(\Delta\alpha)$, the variance of the estimate is $O(1/N_s\Delta\alpha)$, requiring at least $N_s = O(\Delta\alpha^{-1})$ samples to suppress. This is a fundamental issue with finite difference of Monte Carlo (even with CRN), and is the original of many suprisingly spurious results when the issue is not awared of.

1.2.2. Distribution parameters

In view of the issue with finite difference, 'stochastic gradient estimation' is an area that explores functions whose expectation is equal or at least tends to the target gradient, in our case the reliability sensitivity. See [23] a monograph in the area of discrete event simulation. Inevitably, efficiency is gained at the expense of limited scope or approximation. Some existing works assume that α affects q but not f so that the differentiation falls on q only, i.e., $\partial F/\partial\alpha = \int_{f(\boldsymbol{x},\alpha)\geq y} \partial q/\partial\alpha\, d\boldsymbol{x}$. In this case strategies have been developed to evaluate $\partial F/\partial\alpha$ as an integral or expectation, e.g., $\partial F/\partial\alpha = E\left[I\left(f(\mathbf{X},\alpha)\geq y\right) S(\mathbf{X},\alpha)\right]$ where $\mathbf{X}\sim q$; $S = q^{-1}\partial q/\partial\alpha$ is a 'score function'; and $I(\cdot)$ is the indicator function, equal to 1 when the argument is true and zero otherwise. Issues with evaluating F carry over to the expectation $E[\cdot]$. Variance reduction methods have been developed using, e.g.,

line sampling [24, 25, 26], Subset Simulation [27, 28], and moving particles method [29].

1.2.3. General parameters

Expressing $\partial F/\partial\alpha$ as an expectation that is conducive to Monte Carlo is more challenging in the general case where α can affect the response function as well. Differentiating F in (2) under integral sign, e.g., using Leibniz rule, is not trivial, because the integral is over the failure domain that depends implicitly on $(\boldsymbol{x}, \alpha)$. Writing $F = \int I\left(f(\boldsymbol{x}, \alpha) \geq y\right) q(\boldsymbol{x}) d\boldsymbol{x}$ removes this dependence, but still does not allow differentiation using ordinary rules, however, because changing α moves the failure (boundary) surface $\{\boldsymbol{x} \in \mathbb{R}^n : f(\boldsymbol{x}, \alpha) = y\}$ that has a non-trivial effect on F. In fact, $\partial I(\cdot)/\partial\alpha$ is zero except at the failure surface where it does not even exist. In this regard, it is possible to write in terms of a surface integral, as shown in Theorem 21 of [30]. Applying the theorem with a limit state function $g(\boldsymbol{x}, \alpha) = y - f(\boldsymbol{x}, \alpha)$ and considering q a function of $(\boldsymbol{x}, \alpha)$, one obtains

$$\frac{\partial F}{\partial \alpha} = \int_{f(\boldsymbol{x},\alpha)\geq y} \frac{\partial q}{\partial \alpha} d\boldsymbol{x} + \int_{f(\boldsymbol{x},\alpha)=y} \frac{\partial f/\partial \alpha}{\|\partial f/\partial \boldsymbol{x}\|} q(\boldsymbol{x}) dS(\boldsymbol{x}) \tag{3}$$

where the second integral is over the failure surface with differential element $dS(\boldsymbol{x})$; $\partial q/\partial\alpha$, $\partial f/\partial\alpha$ and $\partial f/\partial\boldsymbol{x}$ are all functions of $(\boldsymbol{x}, \alpha)$; $\|\cdot\|$ denotes the Euclidean norm. The first and second term in (3) account for the influence of α on q and $f(\boldsymbol{x}, \alpha)$, respectively. The first term is what was treated in methods that limit α to distribution parameters. For its surface integral nature, the second term is more difficult to handle. Monte Carlo is not directly applicable because of the need to sample $\mathbf{X}$ on the failure surface. In view of this, a 'weak approach' was developed, where an augmented CDF was introduced as a smooth approximation to the indicator function to avoid the surface integral and allow Monte Carlo averaging. See [31] that employs sequential importance sampling to adaptively determine the augmented CDF. See [32] and [33] for discussion of score function method, weak approach, and CRN. See also [34] a review on reliability sensitivity.

2. Key contributions and organization of this work

Advancing the state of the art, this work develops a theory and associated Monte Carlo strategy for reliability sensitivity based on response function values and derivatives w.r.t. sensitivity parameters. To facilitate understanding,

we summarize in this section the key contributions and outline the organization of this work.

2.1. Theory

As the key theoretical contribution, we first derive in Section 3 the following formula for sensitivty:

$$\frac{\partial F}{\partial \alpha} = \int t\, p(G = t, Y = y)\, dt \tag{4}$$

where

$$G = \frac{\partial f(\mathbf{X}, \alpha)}{\partial \alpha} \tag{5}$$

is the (random) response gradient, and $p(G = t, Y = y)$ denotes the joint PDF of $\{G, Y\}$ at $\{t, y\}$. By writing $p(G = t, Y = y) = p(G = t|Y = t)\, p(Y = t)$, (4) can be written as

$$\frac{\partial F}{\partial \alpha} = p(Y = y)\, E[G|Y = y] \tag{6}$$

which expresses the sensitivity in an intuitve form in terms of the PDF of Y and the conditional expectation of G given that $Y = y$:

$$E[G|Y = y] = \int t\, p(G = t|Y = y)\, dt \tag{7}$$

Equation (4) is a signficant theoretical result, as it resolves the implicit dependence of $P(Y \geq y)$ on α, now expressing explicitly in terms of an integral that can be interpreted as an expectation. From first glance it may appear that the result is merely an application of chain rule of differetiation, i.e., $\partial F/\partial \alpha = \partial F/\partial y \cdot \partial y/\partial \alpha$, where $\partial F/\partial y$ gives the PDF $p(Y = y)$ (which is correct) and $\partial y/\partial \alpha$ 'somehow' becomes the conditional expectation $E[G|Y = t]$. This looks intuitive but is not correct. In the first place y is a fixed threshold, and hence not a function of α.

Considering the general context of the sensitivity problem in this work, (4) and its intuitive form in (6) are remarkably simple. Effectively, the derivative of probability, which is not trival and is the key challenge of a sensitivity problem, is now passed down to the response function which can be calculated in a deterministic manner. As it turns out, the conditional expectation correctly reflects the probability nature of the subject under differentiation. The proof is not trivial, however, as it involves evaluating the derivative of $P(Y \geq b)$ from first principle of Calculus, i.e., as the limit of the ratio of perturbations. See also Section 2.3 that compares (4) with (3).

2.2. Monte Carlo with kernel smoothing

Beyond the rare event challenge that is typical in the estimation of F, the additional challenge associated with $\partial F/\partial\alpha$ lies in the integral in (4) when the joint PDF of $\{G, Y\}$ is not known, or the conditional expectation $E[G|Y = y]$ in (6). For the latter, Monte Carlo averaging requires samples conditional on the zero-probability event $\{Y = y\}$, which is not feasible. As the second key contribution, we have developed a simple approach that allows $\partial F(y, \alpha)/\partial\alpha$ to be computed together with the CCDF $F(y, \alpha)$ *for all sample values of* y generated in a single Subset Simulation (SS) run. The difficulty with zero probability event is overcome using the concept of 'kernel smoothing', where Monte Carlo samples in the neighborhood of $\{Y = y\}$ are exploited for estimation at the expense of a bias that is nevertheless controllable. Based on (4), we show in Section 4.1 that the sensitivity can be estimated using direct Monte Carlo samples of $Y_k = f(\mathbf{X}_k, \alpha)$ and $G_k = \partial f(\mathbf{X}_k, \alpha)/\partial\alpha$ with $\mathbf{X}_k \sim q$:

$$\frac{\partial F(y,\alpha)}{\partial\alpha} \approx \sum_{k=1}^{N} G_k \, w^{-1} K\left(\frac{Y_k - y}{w}\right) \tag{8}$$

where $K(\cdot)$ is a 'kernel function', e.g., standard Normal PDF, and w is the 'kernel width' (which can depend on y) that is chosen to balance over estimation bias (the narrower the better) and variance (the wider the better).

For the same difficulty with risk analysis for rare events, (8) has large variance for small F where Monte Carlo samples are lacking. Using SS for generating rare event samples, we show in Section 4 that the sensitivity can be estimated by

$$\frac{\partial F(y,\alpha)}{\partial\alpha} \approx \sum_{i=0}^{m-1} P_i N_i^{-1} \sum_{k}^{N_i} G_{ik} \, w_i^{-1} K\left(\frac{Y_{ik} - y}{w_i}\right) \tag{9}$$

In (9), the subscript ik denotes the kth sample in the 'threshold bin' B_i in SS. See Table 1 for the definition of B_i, its probability P_i (on a sample basis), and the number of samples N_i in the bin. They are defined in the same way as in [35] except that B_{m-1} here combines B_{m-1} and B_m there; see (3)-(5) in Section 2 there. The kernel width w_i can depend on the bin i, or even y (not indicated). See (20) one simple choice based on Normal kernel.

Table 1: Definition of threshold bins $\{B_i\}_{i=0}^{m-1}$ in Subset Simulation

Bin	Probability	No. of samples
$B_i = \{y_i \le Y < y_{i+1}\}$ $0 \le i \le m-2$	$P_i = p_0^i(1-p_0)$	$N_i = (1-p_0)N$
$B_{m-1} = \{Y \ge y_{m-1}\}$	$P_{m-1} = p_0^{m-1}$	$N_{m-1} = N$
$y_0 = -\infty$		

2.3. Remarks on related ideas

Although the context is completely different, the key formula (4) and the technique behind the derivation (Section 3) were discovered after a recent work on analyzing the 'failure mixing rate' in optimizing MCMC for rare events [36]. On the other hand, compared with (3) that is a surface integral over the (potentially high dimensional) input variable space ($\boldsymbol{x}$), (4) is conceptually simpler as it is only a one-dimensional integral over the response gradient space. The conditional expectation form in (6) allows one to develop probabilistic intuitions. It opens up possibilities for general data-analytics such as kernel regression; see the literature note at the end of Section 4 later. While this work was developed independently, on hindsight the resulting approximation via kernel smoothing with direct Monte Carlo samples is similar to [31] that was based on (3); see (8) in this work and (11) in [31]. Nevertheless, the developed algorithms are different, i.e., based on SS in this work, and based on sequential importance sampling in [31]. As seen in (9) and demonstrated in the numerical examples (Section 5), the use of SS allows sensitivity calculations to be naturally embeded in existing codes of SS and to be obtained for all response threshold values (y) generated in an SS run.

2.4. Organization of this work

The remaining parts of this work are organized as follows. Section 3 presents the theory and proves the key formula (4). Section 4 discusses Monte Carlo estimation via (4) with kernel smoothing. Section 5 investigates the proposed methodology using a number of examples of different nature; see Table 2 for a summary. Sensitivity parameters of special nature are illustrated, e.g., those with non-zero response gradient but zero sensitivity in F, such as α_3 in Example 1 (Normal response); those with a significant chance

of zero response gradient, such as α_2 (second story stiffness) in Example 2 (shear building buckling); those displaying weak correlation between response and gradient, but still carrying significant sensitivity in F, such as α_2 (natural frequency) in Example 3 (first passage problem); and those in applications (Example 4, pile design) that may prefer finite difference over analytical means for response gradient to save coding effort.

3. Theory of sensitivity and response gradient

In this section we present the main theory, which shows that the sensitivity $\partial F/\partial\alpha$ is given by (4). The derivation assumes that the response $Y = f(\mathbf{X}, \alpha)$ and its gradient $G = \partial f(\mathbf{X}, \alpha)/\partial\alpha$ exist almost surely, and they have continuous joint, conditional and marginal PDFs. This is to justify their appearance in various places, and the Taylor approximation in (13). Although the problem contexts are completely different, the technique used here is similar to that in Section 7.2 of a recent work [36] investigating the derivative of failure mixing rate w.r.t. hyperparameters in optimizing MCMC for rare events.

Technically, α affects $F(y, \alpha) = P(Y \geq y)$ through $Y = f(\mathbf{X}, \alpha)$, and hence any probability about Y. The effect is implicit, and so direct differentiation of F is not possible. We will start from first principles, i.e., derivative as the limit of the ratio of perturbations. Consider a perturbation $\Delta\alpha$ in α. To the first order of $\Delta\alpha$, Y will be perturbed to

$$Y' = Y + G\Delta\alpha \tag{10}$$

Replacing Y in $F = P(Y \geq y)$ by Y' and rearranging gives the perturbed F:

$$F' = P(Y \geq y - G\Delta\alpha) \tag{11}$$

Unlike the usual reasoning in Calculus, it is not legtimate to make the Taylor approximation $F' \approx P(Y \geq y) - G\Delta\alpha\, \partial P(Y \geq y)/\partial y$ because G is random and it is generally correlated with Y. In view of this, we condition on $G = t$ to remove the randomness:

$$F' = \int P(Y \geq y - t\Delta\alpha | G = t)\, p(G = t)\, dt \tag{12}$$

For every t, we can now make the Taylor approximation for small $\Delta\alpha$:

$$\begin{aligned} &P(Y \geq y - t\Delta\alpha | G = t) \\ &= P(Y \geq y | G = t) - t\,\Delta\alpha \frac{\partial}{\partial y} P(Y \geq y | G = t) + o(\Delta\alpha) \\ &= P(Y \geq y | G = t) + t\,\Delta\alpha\, p(Y = y | G = t) + o(\Delta\alpha) \end{aligned} \tag{13}$$

where $p(Y = y|G = t)$ denotes the conditional PDF of Y at y given that $\{G = t\}$. Substituting (13) into (12) gives three integrals. The first is equal to F. Substracting it from LHS gives the perturbation $\Delta F = F' - F$. Dividing by $\Delta\alpha$ and noting $p(Y = y|G = t)\, p(G = t) = p(G = t, Y = y)$ gives

$$\frac{\Delta F}{\Delta\alpha} = \int t\, p(G = t, Y = y) dt + \frac{o(\Delta\alpha)}{\Delta\alpha} \tag{14}$$

Taking $\Delta\alpha \to 0$ gives the required result in (4).

4. Monte Carlo with kernel smoothing

In this section we develop a Monte Carlo method for efficient calculation of reliability sensitivity, leveraging the key formula in (4). A Monte Carlo approach is adopted to allow general applications, where only the point-wise value of the response function and its gradient are required, but otherwise no other information is needed. Subset Simulation (SS) is adopted for generating samples of frequent to rare events. As will be seen shortly, the method allows one to calculate $\partial F(y, \alpha)/\partial\alpha$ in the same SS run for $F(y, \alpha)/\partial\alpha$, covering all values of y generated in the SS run from high to low probability.

Equation (4), or its conditional expectation form in (6), suggests averaging over Monte Carlo samples of G. A fundamental difficulty arises from the conditioning on $\{Y = y\}$, which is a zero-probability event. In typical situations, Monte Carlo samples of G are associated with random values of Y, and so it is impossible to obtain samples with $Y = y$ exactly. In this work we overcome this difficulty using the concept of 'kernel smoothing'. It allows one to use samples in the neighborhood of $\{Y = y\}$ for Monte Carlo averaging at the expense of introducing a bias, which is nevertheless controllable. In what follows, we first introduce the idea and discuss how the integral in (4) can be estimated by direct Monte Carlo, i.e., using samples of $\{Y, G\}$ where $\mathbf{X} \sim q$. After that we extend the idea to SS, where the samples are

conditional on different threshold bins of Y. Kernel smoothing is not new and is a topic in itself in data-science. See a literature note near the end of Section 4.3 later.

4.1. Direct Monte Carlo

The basic idea of kernel smoothing in our context is to approximate the integral in (4) by a double integral over t and y so that it can be estimated by Monte Carlo averaging over samples of $\{G, Y\}$. To conform with the condition $\{Y = y\}$, a weighting is introduced so that only contributions of the samples in the neighborhood of $\{Y = y\}$ are accounted for. Specifically, let $K(\tau)$ be a user-defined univariate PDF. For its role in the problem that will become clear shortly, it is called a 'kernel function', and is often chosen to have a centralized shape around zero and significant probability over a length scale of $O(1)$, e.g., standard Normal PDF. It can be easily verified that for given $w > 0$ and $y \in \mathbb{R}$, a function of τ in the form $w^{-1}K\left((\tau - y)/w\right)$ is still a univariate PDF, but now it is centered at y and varies over a scale of w. Consider now an integral of the product of sensitivity and kernel:

$$\begin{aligned} J(y,\alpha) &= \int \frac{\partial F(\tau,\alpha)}{\partial \alpha} w^{-1} K\left(\frac{\tau - y}{w}\right) d\tau \\ &= \iint t\, w^{-1} K\left(\frac{\tau - y}{w}\right) p(G = t, Y = \tau)\, dt\, d\tau \\ &= E\left[G\, w^{-1} K\left(\frac{Y - y}{w}\right)\right] \end{aligned} \tag{15}$$

where the expectation is taken with $\{G, Y\}$ distributed as $p(G = t, Y = y)$; the second equality is obtained by substituting (4). Intuitively, from the first expression in (15), $J(y,\alpha)$ collects the contributions of $\partial F(\tau,\alpha)/\partial\alpha$ over a neighorhood of width w around $\tau = y$. As $w \to 0$, the function $w^{-1}K\left((\tau - y)/w\right)$ tends to a Dirac Delta function centered at $\tau = y$. One can then expect that $J(y,\alpha) \to \partial F(y,\alpha)/\partial\alpha$. This is the idea of kernel smoothing in our context, where J is taken as a proxy for $\partial F/\partial\alpha$. The merit is that it can be estimated by averaging the term inside $E[\cdot]$ over Monte Carlo samples of $\{G, Y\}$, which is feasible now because the zero-probability condition $\{Y = y\}$ has been removed. See the resulting estimator in (8). Of course, this feasibility is made at the expense of a bias, but which can be suppressed with sufficiently small w. In particular, if $K(\cdot)$ is symmetric

about 0 and $\partial F(\tau,\alpha)/\partial\alpha$ is twice differentiable w.r.t. τ at $\tau = y$, then the bias is second order, i.e.,

$$J(y,\alpha) = \frac{\partial F}{\partial \alpha} + O(w^2) \tag{16}$$

This can be reasoned by substituting the Taylor series

$$\frac{\partial F(\tau,\alpha)}{\partial\alpha} = \frac{\partial F(y,\alpha)}{\partial\alpha} + \frac{\partial^2 F(y,\alpha)}{\partial y \partial\alpha}(\tau - y) + O\left((\tau-y)^2\right) \tag{17}$$

into the first expresion in (15) and noting that upon integration the zeroth order term gives the target $\partial F(y,\alpha)/\partial\alpha$, the first order term gives zero, and the reminder (second order) term gives $O(w^2)$. The choice of w should balance between estimation bias and variance. See Section 4.3 later.

4.2. Subset Simulation

The idea of kernel smoothing in the last section has resolved the difficulty of zero-probability conditioning, but issues with rare events still remain. Unless the number of samples is very large ($\propto 1/F$), direct Monte Carlo does not offer enough samples (if any) of $\{G, Y\}$ for estimating small F or its sensitivity $\partial F/\partial\alpha$. In this work we use Subset Simulation (SS) to address rare events, generating samples covering frequent to small regimes of the CCDF of Y. As the samples in SS are conditional on a series of thresholds, their occurence needs to be properly weighted in the Monte Carlo average. Below we show how that can be done to give a proper estimator.

Consider a conventional SS run where N samples of $\mathbf{X}$ and $Y = f(\mathbf{X},\alpha)$ are generated at each simulation level i $(= 0, 1, \ldots, m-1)$ conditional on $\{Y \geq b_i\}$, with $b_0 = -\infty$, and $b_1 < b_2 < \cdots < b_m$ an increasing sequence of generated thresholds such that on a sample basis $P(Y \geq b_{i+1} | Y \geq b_i) = p_0$, i.e., the 'level probability'. For sensitivity calculation, the theory in this work requires that the response gradient $G = \partial f(\mathbf{X},\alpha)/\partial\alpha$ be also calculated for all samples of $\mathbf{X}$. Let the samples be grouped into a series of 'threshold bins' $\{B_i\}_{i=0}^{m-1}$ that partition the real line. Table 1 summarizes the definition of the bins B_i, their probability P_i and the number of samples N_i they contain. Using the theorem of total probability, (15) can be written as

$$J(y,\alpha) = \sum_{i=0}^{m-1} E\left[G\, w^{-1} K\left(\frac{Y-y}{w}\right) \middle| B_i \right] P(B_i) \tag{18}$$

The conditional expectation $E[\cdot|B_i]$ can be estimated by averaging over the samples in B_i:

$$E\left[G\, w^{-1} K\left(\frac{Y-y}{w}\right)\middle| B_i\right] \approx N_i^{-1} \sum_{k=1}^{N_i} G_{ik} w^{-1} K\left(\frac{Y_{ik}-y}{w}\right) \tag{19}$$

where the subscript ik denotes the kth sample of B_i. Although it has not be explicitly indicated, the kernel width w can depend on y, i.e., 'adaptive', to account for the density of $\{Y_{ik}\}$ around y or other effects such as relationship with G. Substituting (19) into (18) and writing $P(B_i) = P_i$ gives the key estimation formula in (9). In theory the threshold y can be any value, although for proper estimation it should be limited to the range of values generated in an SS run. Taking computer coding into consideration, a convenient way is to calculate the sensitivity at the generated thresholds of Y. In this manner, an SS run gives an estimate of the CCDF curve $F(y, \alpha)$ (as conventional), as well as the 'sensitivity curve' $\partial F(y, \alpha)/\partial\alpha$ (new in this work) for all generated values of y.

As long as the conditional samples in different bins are available, the methodology here is applicable regardless of how they are generated in SS, e.g., independent component Metropolis random walk in the debut paper [37], conditional sampling [38] that is an elegant finite-infinite dimensional equivalent [39], Hamiltonian Monte Carlo [40, 41] or its Riemannian variant [42, 43], and affine transformation [44, 45]. Aside, see recent applications of SS to autonomous driving [46], climate simulation [47] and fusion energy systems [48].

4.3. Kernel width

As mentioned before, the key idea in kernel smoothing that resolves the issue of zero-probabilty conditioning $\{Y = y\}$ is the use of kernel $K(\cdot)$ that allows averaging over Monte Carlo samples of $\{Y \neq y\}$ but at the same time weighting them so that only those sufficiently close to $\{Y = y\}$ matters. The weights are controlled by the kernel width w, which should be chosen to balance over estimation bias and variance. A small w can reduce bias because only samples near $\{Y = y\}$ receive a significant weight, but it will increase the variance of Monte Carlo average because the effective number of samples is reduced; and vice-versa. In between the extremes of small-bias-high-variance and high-bias-low-variance lies the optimal w. Intuitively, the optimal w depends on the total number of samples. For example, more samples means lower variance and hence can accomodate a smaller w.

Similar to many algorithm parameters, the choice of kernel width is an open issue that depends on the distribution of data but which is not known a priori. In this work we adopt the Scott's rule [49, 50] for its simplicity and theoretical elegance (optimal for Normal data). For $y \in B_i$ (Table 1), the kernel width is taken as

$$w_i = \sigma_Y \left(\frac{4}{3N_i} \right)^{1/5} \tag{20}$$

where N_i is the number of samples in B_i, and

$$\sigma_Y^2 = E[Y^2] - E[Y]^2 \tag{21}$$

is the variance of Y when $\mathbf{X} \sim q$. The statistical moments can be estimated in a single SS run using the samples in all bins, by virtue of the theorem of total probability ($r = 1, 2$):

$$E[Y^r] = \sum_{i=0}^{m-1} E[Y^r | B_i] P_i = \sum_{i=0}^{m-1} P_i N_i^{-1} \sum_{k=1}^{N_i} Y_{ik}^r \tag{22}$$

While the specific formula and relevant statistic of the optimal kernel width depends on the criterion and distribution of samples, the scaling $w_i \propto N_i^{-1/5}$ has an origin that can be reasoned intuitively as follows. A Monte Carlo average with kernel smoothing of width w is effectively an average of $w\,N_s$ samples, where N_s is the effective number of independent samples. Its variance is therefore $O(w^{-1}N_s^{-1})$. If the bias is $O(w^p)$ for some p (say), the mean squared error (MSE) of the average from the target is of the form $a\,w^{-1}N_s^{-1} + b\,w^{2p}$, where a and b are constants that depend on the distribution and location of estimate (y). Minimizing the MSE gives the optimal width of $w = (a/2bpN)^{1/(2p+1)}$. Taking $p = 2$ for second order bias (see (16)) then gives $w \propto N_s^{-1/5}$ as implied by (20). Intuitively, a larger N_s allows one to reduce the bias by narrowing w, but the latter should proceed at a slower rate than $O(N_s^{-1})$ to suppress the growth of variance due to the lack of samples in the $O(w)$ neighborhood of $\{Y = y\}$.

As a literature note, kernel smoothing is a nonparametric technique for estimation involving conditional expectation given one or more features by constructing a weighting, i.e., kernel function. See [51] an early work by Nadaraya, and [52] a monograph in econometrics. Besides the kernel, performance also depends on the locally weighted regression forms, e.g., constant,

linear or polynomial. The form in this work is a simple conventional choice often referred as the Nadaraya-Watson model, which corresponds to a constant locally weighted regression form. See Chapter 5 of [53] that discusses general 'local polynomial estimators', [54] a review, and [55] a general review on nonparametric regression.

5. Illustrative examples

In this section we present a series of examples to illustrate the theory and computational methodology. As is conventional in SS, the input random variables $\{X_i\}_{i=1}^n$ are assumed to be i.i.d. N(0,1). In each example we first present the setup, how to calculate the response $Y = f(\mathbf{X}, \alpha)$ and its gradient $\partial Y/\partial\alpha$ w.r.t. sensitivity parameter α. If available, analytical solution for the failure probability $F = P(Y \geq y)$ and its sensitivity $\partial F/\partial\alpha$ will also be presented. In the numerical investigation, F and $\partial F/\partial\alpha$ will be calculated using SS in the same Monte Carlo run. The results of sensitivity will be compared with a benchmark. The latter is taken as the analytical solution when available (Examples 1 and 2), or otherwise as the central difference of direct Monte Carlo estimates with common random numbers (Examples 3 and 4). Results of a single SS run will be presented to give an idea of typical results. Statistics from 1000 independent runs will also be presented to assess the bias and variance of sensitivity estimates.

SS will be implemented for $m = 3$ levels with a level probability of $p_0 = 0.1$ and $N = 1000$ samples per level. This is intended to provide samples populating the CCDF of Y from 1 down to at least $p_0^3 = 10^{-3}$. There will be $m = 3$ bins, i.e., B_0 and B_1 with $(1 - p_0)N = 900$ samples, and B_3 with $N = 1000$ samples. Conditional sampling [38] is used for generating MCMC samples. In level $i \geq 1$ the candidate (before checking $Y \geq b_i$) is generated as $\mathbf{X}' = a_i\mathbf{X} + s_i\mathbf{Z}$ where $\mathbf{X}$ is the current sample, $\mathbf{Z}$ is an i.i.d. N(0,1)^n vector independent of $\mathbf{X}$, $a_i \in (0, 1)$ is a correlation (hyper-) parameter and $s_i = \sqrt{1 - a_i^2}$. We take $a_i = 0.5(1 + u_i/v_i)$ where $u_i = \overline{\Phi}^{-1}(p_0^i)$ and $v_i = \overline{\Phi}^{-1}(p_0^{i+1})$; $\overline{\Phi}(\cdot)$ denotes the standard Normal CCDF. This choice minimizes the correlation between successive MCMC samples of Normal response, and is found to be a good for other applications [56].

Four examples serving complementary purposes are considered. See Table 2 for a summary. The first example (Section 5.1) considers Normal response with a simple relationship betwen input variables and output response, providing a clear exposition of theory with analytical solution for

benchmarking. All sensitivity parameters have well-behaved response gradients. Parameter α_3 is designed to illustrate the possibility of non-zero response gradient but zero reliability sensitivity, prompting for care in applying intuitions and interpreting sensitivity results in proper scales.

The second example (Section 5.2) considers the global buckling of a shear building. It serves to outline the basic ingredients for general structures while admiting analytical solution for the special case of shear building for benchmarking. The example is scalable with the number of stories and input random variables, although for illustration a five-storied building with five i.i.d. Lognormal random variables for the floor loads is considered. The mean value of random floor loads (α_1) and the deterministic stiffness of the second story (α_2) are considered to illustrate different types of sensitivity parameters. The latter is designed to illustrate the case when the response gradient has a significant chance to be zero. This is typical in systems with components connected in series, where system failure is insensitive to components that have a relatively low failure rate than others. As will be seen, this renders the influence of the parameter a 'rare event' itself, inflating the variance of sensitivity estimate and calling for advanced strategies for variance reduction.

The third example (Section 5.3) considers the first passage failure of a single-degree-of-freedom (SDOF) structure subjected to white noise excitation. Despite the SDOF nature, the problem is complicated for its large number of input random variables ($n = 400$) in modeling the excitation, the same large number of failure modes induced by the possible peak response time, and dynamics that induces non-trivial correlation among them. The problem is representative of a high dimensional problem, and has been considered in previous works investigating MCMC for rare events [56, 36]. See also a review on first passage problems [57]. As analytical solution for failure probability or sensitivity is not available, central difference with direct Monte Carlo is used for benchmarking. Sensitivities to the damping ratio (α_1) and natural frequenty (α_2) are considered, which are of very different nature. The latter is designed to illustrate the possibility of a parameter displaying weak correlation between the response and its gradient, but still carrying signficant reliability sensitivty. As will be seen, the resulting sensitivity estimate can have large variance, calling for advanced strategies for variance reduction.

The fourth example (Section 5.4) considers the risk of excessive settlement in the serviceability limit state design of pile foundation. The problem is considered for its relevance in application, and complication in its spatially varying soil friction angle along the depth. The latter is modeled by $n =$

120 input random variables transformed through Cholesky decomposition of correlation matrix to produce a Lognormal random field with specified mean, standard deviation and correlation length. Sensitivities to the pile diameter (α_1) and the mean of friction angle (α_2) are considered, whose results are intuitive but still non-trivial especially in the correlation between response and its gradient. Different from other examples, the response gradient is calculated using finite difference, which is a scenario in applications where only response function values can be accessed, or it is preferred over analytical means to simplify coding effort. Similar to Example 3, analytical solution for sensitivity is not available, and so central difference with direct Monte Carlo is used for benchmarking.

Table 2: Summary of examples. Input random variables $\{X_i\}_{i=1}^n$ i.i.d. N(0,1). (Output) response $Y = f(\mathbf{X}, \alpha_i)$. α_i = sensitivity parameter. Failure probability $F = P(Y \geq y)$, sensitivity $\partial F/\partial \alpha_i$

No.	n	Y	$\partial Y/\partial \alpha$	Remarks
1	2	Normal response $Y =$ $\alpha_1 + (\alpha_2^2 - \alpha_3)^{1/2} X_1 + \alpha_3 X_2$	(26) and (27)	α_3 has $\partial Y/\partial \alpha_3 \neq 0$ but $\partial F/\partial \alpha_3 = 0$ Analytical solution see (24) and (25)
2	5	Shear building buckling $Y = \lambda_0/\lambda$ λ = from eigenvalue problem (30) λ_0 = normalizing factor Lognormal vertical loads $W_i = w\, e^{a+bX_i}$ a and b from (35) $\alpha_1 \equiv w$ $\alpha_2 \equiv$ 2nd story stiffness	(32), solve matrix equation (33)	α_2 is difficult, with 80% chance of $\partial Y/\partial \alpha_2 = 0$ Analytical solution see (41), (44) and (45)

3	400	Linear SDOF first passage $Y = \max_{j=0}^{n-1} \lvert u(j\Delta t)\rvert$ $\ddot{u} + 2\zeta\omega\dot{u} + \omega^2 u = W$ $\omega = 2\pi$ (rad/s), $\zeta = 1\%$ $u(0) = \dot{u}(0) = 0$ $W_i = (\frac{2\pi S}{\Delta t})^{1/2} X_{i+1}$ $0 \le i \le n-1$ $\Delta t = 0.05$ s, $S = 0.86$ N^2/(rad/s) $\alpha_1 \equiv \zeta, \alpha_2 \equiv \omega$	(48), solve ODE (49) and (50)	α_2 is difficult, apparently weak correlation between Y and $\partial Y/\partial\alpha_2$, but still signficant sensitivity $\partial F/\partial\alpha_2$ No analytical solution Benchmark with central difference of F from direct Monte Carlo. Relative step size 1%
4	120	Pile serviceability design $Y = F_{50}/Q_{sls}$ F_{50} = design load (fixed) Q_{sls} = effective soil resistance (random) See (54) and Table 3 Soil friction angles $\{\phi_i'\}_{i=1}^n \sim$ Lognormal with mean $\mu = 0.5585$ rad (32°), c.o.v. = 17%, and correlation length 4 m $\alpha_1 \equiv B$ (pile diameter) $\alpha_2 \equiv \mu$	Central difference with relative step size 0.1%	No analytical solution Benchmark with central difference of F from direct Monte Carlo. Relative step size 1%

5.1. Normal response

Consider the response defined by

$$Y = \alpha_1 + (\alpha_2^2 - \alpha_3^2)^{1/2} X_1 + \alpha_3 X_2 \tag{23}$$

where $\{\alpha_1, \alpha_2, \alpha_3\}$ are sensitivity parameters and $\{X_1, X_2\}$ are i.i.d. N(0,1). It can be reasoned that $Y \sim N(\alpha_1, \alpha_2^2)$ and so

$$F(y, \boldsymbol{\alpha}) = P(Y \ge y) = \Phi(-z) \qquad z = \frac{y - \alpha_1}{\alpha_2} \tag{24}$$

where $\Phi(\cdot)$ denotes the standard Normal CDF. Taking derivatives gives

$$\frac{\partial F}{\partial \alpha_1} = \alpha_2^{-1}\phi(z) \qquad \frac{\partial F}{\partial \alpha_2} = \alpha_2^{-1} z\,\phi(z) \qquad \frac{\partial F}{\partial \alpha_3} = 0 \tag{25}$$

5.1.1. Analytical verification with theory

We now determine the sensitivities using the theory in this work, i.e., (6), and verify that they are the same as those in (25). The basic steps are 1) derive $\partial Y/\partial \alpha_i$, 2) determine the conditional distribution and hence expectation of $\partial Y/\partial \alpha_i$ given $\{Y = y\}$, and 3) substitute all results into (6). Direct differentiation of (23) gives

$$\frac{\partial Y}{\partial \alpha_1} = 1 \qquad \frac{\partial Y}{\partial \alpha_2} = \alpha_2(\alpha_2^2 - \alpha_3^2)^{-1/2} X_1 \tag{26}$$

$$\frac{\partial Y}{\partial \alpha_3} = -\alpha_3(\alpha_2^2 - \alpha_3^2)^{-1/2} X_1 + X_2 \tag{27}$$

Since $\partial Y/\partial \alpha_1$ is a constant, it is clear that $E[\partial Y/\partial \alpha_1 | Y = y] = 1$. For $E[\partial Y/\partial \alpha_2 | Y = y]$, we need to determine $E[X_1|Y = y]$. Note from (23) that $Y \sim N(\alpha_1, \alpha_2^2)$ is jointly Normal with $X_1 \sim N(0, 1)$ with covariance $(\alpha_2^2 - \alpha_3^2)^{1/2}$. Recall from standard results that if $Z_1 \sim N(\mu_1, \sigma_1^2)$ and $Z_2 \sim N(\mu_2, \sigma_2^2)$ are jointly Normal with covariance c, then $E[Z_2|Z_1] = (Z_1 - \mu_1)\, c/\sigma_1^2$. Applying this result gives

$$E[X_1|Y = y] = \alpha_2^{-2}(\alpha_2^2 - \alpha_3^2)^{1/2}(y - \alpha_1) \tag{28}$$

Multiplying this with the PDF of Y, i.e., $\alpha_2^{-1}\phi(z)$ where $z = (y - \alpha_2)/\alpha_3$, gives the same expression for $\partial F/\partial \alpha_2$ in (25). Finally, the benchmark result $\partial F/\partial \alpha_3 = 0$ in (25) appears counter-intuitive, as $\partial Y/\partial \alpha_3$ in (27) is generally non-zero. Equation (6) does give the same result, however. To see this, evaluating the (unconditional) covariance of $\{Y, \partial Y/\partial \alpha_3\}$ using Y in (23) and $\partial Y/\partial \alpha_3$ in (27) shows that it is zero. Since Y and $\partial Y/\partial \alpha_3$ are jointly Normal, this implies that are independent and so $E[\partial Y/\partial \alpha_3 | Y = y] = E[\partial Y/\partial \alpha_3] = 0$.

5.1.2. Single SS run

For illustration, consider $\alpha_1 = \alpha_2 = 1$ and $\alpha_3 = 0.5$. Figure 1 shows the results of a typical SS run. In all plots, the blue dots and red lines indicate the results of SS and benchmark, respectively. The vertical green lines mark the threshold levels $\{y_1, y_2, y_3\}$ at CCDF values $\{p_0^i\}_{i=1}^m = \{10^{-1}, 10^{-2}, 10^{-3}\}$.

Counting from the left to right, the $(1-p_0)N = 900$ samples to the left of the first vertical line go to B_0; those (same number) between the first and second line go to B_1; the remaining $N = 1000$ samples go to B_2. The quality of estimates for F and $\partial F/\partial\alpha$ to the right of the third line is deteriorating fast as it runs out of samples there.

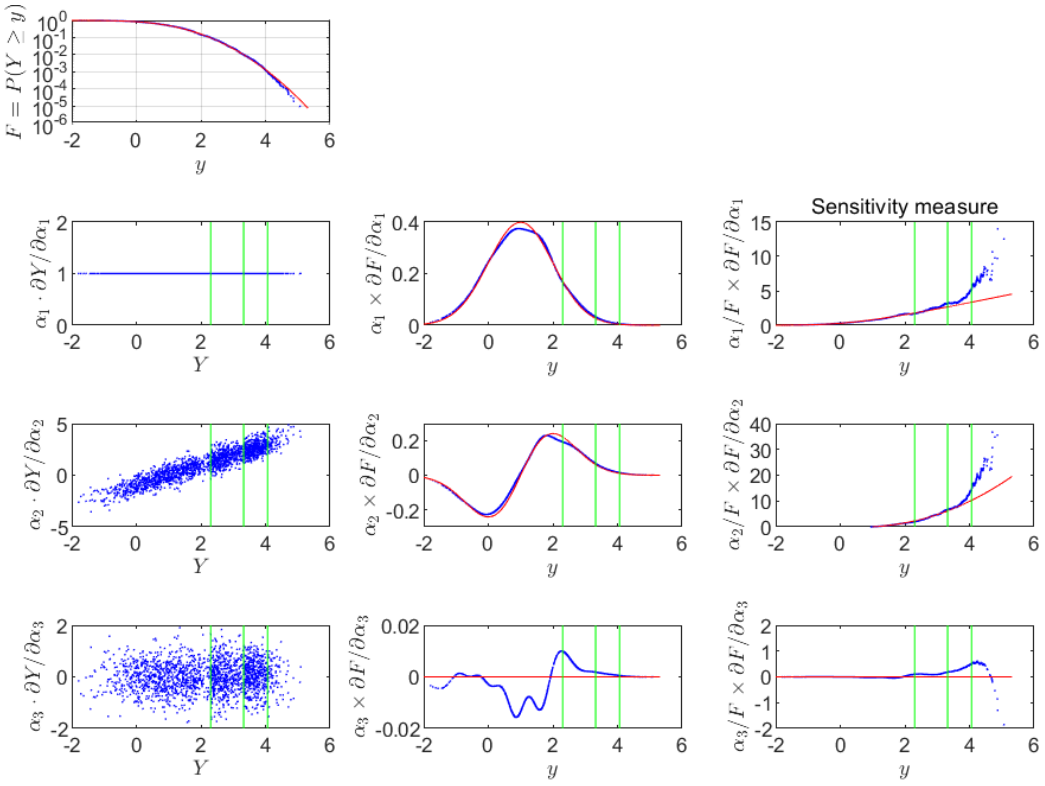

Figure 1: Example 1 (Normal response). Results of a typical SS run with $m = 3$ levels, $p_0 = 0.1$ and $N = 1000$ samples per level. Blue = SS, red = Benchmark, vertical green lines = thresholds at $\{10^{-1}, 10^{-2}, 10^{-3}\}$. The right column shows the fractional change of F per fractional change of sensitivity parameter, the recommended sensitivity measure.

The top-left plot shows the CCDF estimate of Y, the basic output of SS that is commonly reported. The rows below the top show the results that illustrate the theory and computation of this work, one row for each sensitivity parameter, and logically from left to right. Consider for example the second row, i.e., α_1. The left plot shows the samples of response gradient (normalized) versus the response values in all bins. The former has been multiplied with the sensitivity parameter (α_1) so that it has the same unit as

the response. The response gradient normalized this way can be interpreted as the rate of change of Y per fractional change of sensitivity parameter. Consistent with the first expression in (26), the sample values of $\partial Y/\partial\alpha_1$ are all equal to 1.

The middle column in Figure 1 shows the sensitivity estimate (normalized) using SS and kernel smoothing, i.e., (9). It is dimensionless, and can be interpreted as the rate of change of F per fractional change of the sensitivity parameter. For α_1, the sensitivity takes the shape of a Normal PDF. This is consistent with the exact solution in the first expression of (25). The SS (blue) estimate generally agrees with the benchmark (red). It populates over the central as well as the upper tail (low probability) region. The sensitivity tends to zero for large y because the PDF of Y does. This is typical and can be explained by (6), where $\partial F/\partial\alpha = p(Y = y)\, E[G|Y = y] \to 0$ if $E[G|Y = y] < \infty$. Of course, this trivial diminution does not imply that the parameter is no longer sensitivity for large y. It is just a matter of scaling and perspective, similar to the trivial convergence to zero of the variance $p(1-p)/N$ of direct Monte Carlo estimate for failure probability $p \to 0$.

In view of the trivial deminution, a more useful measure is to further divide the sensitivity by the CCDF value F. This leads to the sensitivity measure in the right column of Figure 1, which can be interpreted as the fractional change of F due to fractional change in the sensitivity parameter. The PDF and hence diminution effect has been removed. On a fractional basis, the sensitivity w.r.t. α_1 indeed increases as y increases. For example, at $F \approx 10^{-3}$ ($y \approx 4$), a 1% change in α_1 leads to about 3% change in F. The same 1% change in α_2 leads to 10% change in F. In this sense F is 3 times more sensitive to α_2 than α_1.

The third row in Figure 1 shows the results for α_2. They are qualitatively similar to those of α_2, except that the response gradient (left plot) is now correlated with the response, and the sensitivity (middle plot) can be negative; see the second expression in (25) and (26). The results for α_3 in the bottom row are qualitatively different from those for α_1 or α_2. Recall from Section 5.1.1 that α_3 does not affect F (hence zero sensitivity) but $\partial Y/\partial\alpha_3$ is not zero. The scatter plot on the left displays a lack of correlation, not surprising because $\partial Y/\partial\alpha_3$ and Y are indeed uncorrelated. Despite this, the sensitivity estimate in the middle or right plot is non-zero, simply because of estimation error. This reflects the importance of viewing sensitivity estimates with the right scaling and order of magnitude in perspective. For example, the oscillatory behavior with y in the middle plot is merely an estimation

error, which can be misleading when the order of magnitude is not paid attention to. The sensitivity estimate on the right plot is not zero because of estimation error, but it has a small magnitude ($< 1\%$) compared to other parameters, delivering the correct message of low sensitivity. The divergence beyond the third vertical line is expected, as it runs out of samples there.

5.1.3. Bias and variance from 1000 SS runs

Figure 2 shows the sample mean and $\pm 1\sigma$ bound of the sensitivity estimate from 1000 i.i.d. SS runs. It has the same layout as Figure 1 except that the scatter plots in the left column are omitted, as they are no longer relevant. The top-left plot shows that the sample mean of CCDF agrees with the benhcmark. This is consistent with common findings and not surprising because SS is asymptotically unbiased for sufficiently large N. The new investigation that matters lies in the rows below. For the plots in the middle column, all lines are visually close, i.e., the sample mean of SS and its $\pm 1\sigma$ bounds all gather around the benchmark. This is largely attributed to the fact that the sensitivity in the middle column is dominated by the PDF of Y. The PDF estimate derives it quality from the CCDF estimate, which is of good quality as evidenced from the top-left plot. A more critical assessment lies in the plots of the right column, the sensitivity measure recommended in this work. For y beyond y_2, i.e., the middle vertical green line, a visual departure of the blue dots (SS mean) from the red line (benchmark) can be seen to increase with y, implying an increasing bias. The gap between the $\pm 1\sigma$ bounds also widens, not suprising because it runs out of samples for $y > y_3$ (the rightmost vertical line). For $y < y_3$, the growth of variance is somewhat suppressed, by virtue of SS that populates samples up to around y_3. The bias can be further reduced by using a smaller kernel width, though at the expense of widening the $\pm 1\sigma$ gap. The sensitivity estimate at the left end (small y) of the plot diverges because it runs out of samples there. It is immaterial because that region of low thresholds is not of concern in applications.

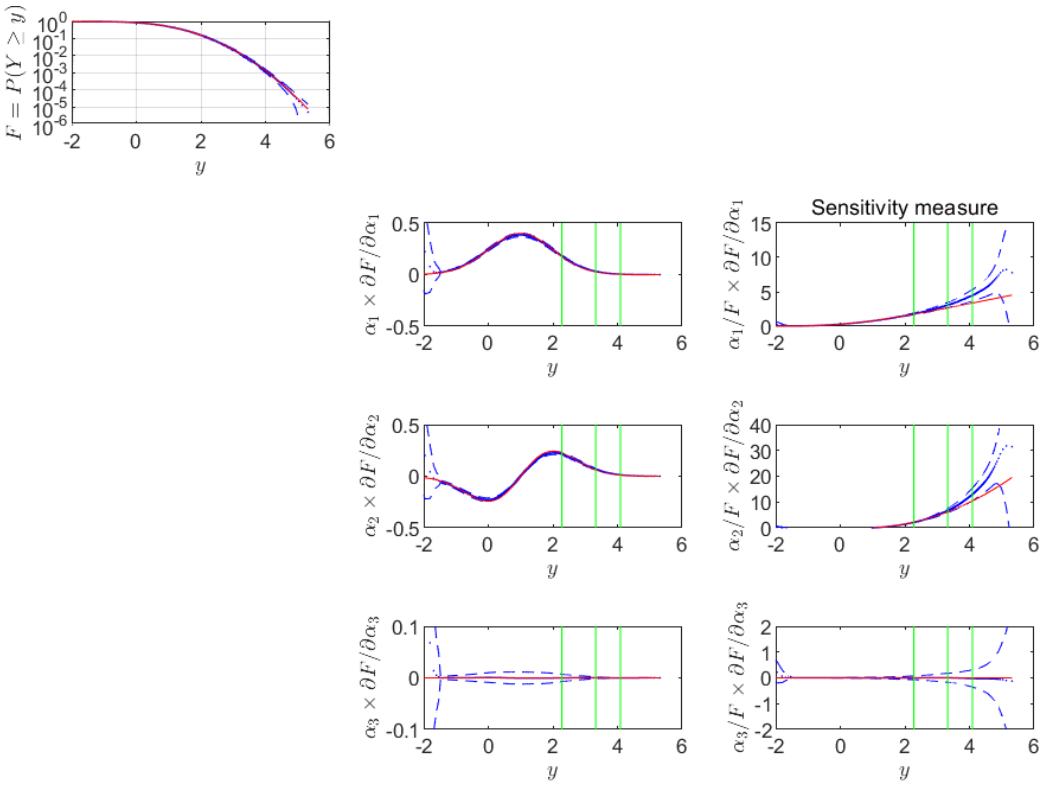

Figure 2: Example 1 (Normal response). Sample mean (blue solid line) and $\pm 1\sigma$ bounds (blue dashed line) from 1000 SS runs. Red = Benchmark. Vertical green lines = thresholds at $\{10^{-1}, 10^{-2}, 10^{-3}\}$ of SS sample mean. Same layout as Figure 1 except that scatter plots in the left column are omitted.

5.2. Buckling load of shear building

Consider the buckling of a structure with linear-elastic stiffness matrix $\mathbf{K}$, and geomatric stiffness matrix $\mathbf{K}_g$ that accounts for first order P-Delta effect. To assess buckling risk, we define the 'response' as

$$Y = \frac{\lambda_0}{\lambda} \tag{29}$$

where λ_0 is a normalizing constant, see Section 5.2.4 later for specific value; λ is the critical buckling load factor (the higher the better), equal to the smallest eigenvalue of the generalized eigenvalue problem:

$$\mathbf{K}\mathbf{u} = \lambda \mathbf{K}_g \mathbf{u} \tag{30}$$

where $\mathbf{u}$ is the eigenvector, i.e., buckling shape. For an n-storied shear building with lateral story stiffness k_i $(1 \le i \le n_s)$,

$$\mathbf{K} = \begin{bmatrix} k_1 + k_2 & -k_2 & & & \\ -k_2 & k_2 + k_3 & -k_3 & & \\ & -k_3 & \ddots & \ddots & \\ & & \ddots & k_{n-1} + k_n & -k_n \\ & & & -k_n & k_n \end{bmatrix} \tag{31}$$

The geomatrix stiffness matrix $\mathbf{K}_g$ has the same form as $\mathbf{K}$ except that k_i is replaced by W_i/H_i, where W_i is the floor load and H_i is the story height.

5.2.1. Response derivatives

Differentiating (29) gives, for a generic sensitivity parameter α,

$$\frac{\partial Y}{\partial \alpha} = -\frac{\lambda_0}{\lambda^2} \frac{\partial \lambda}{\partial \alpha} \tag{32}$$

For general $\mathbf{K}$ and $\mathbf{K}_g$, the derivative of eigenvalue λ and its eigenvector $\mathbf{u}$ (though not required here) can be obtained by solving a matrix equation of augmented dimension [58]:

$$\begin{bmatrix} \mathbf{K} - \lambda \mathbf{K}_g & -\mathbf{K}_g \mathbf{u} \\ -\mathbf{u}^T \mathbf{K}_g & 0 \end{bmatrix} \begin{bmatrix} \partial \mathbf{u}/\partial \alpha \\ \partial \lambda / \partial \alpha \end{bmatrix} = \begin{bmatrix} -\left(\partial \mathbf{K}/\partial \alpha - \lambda\, \partial \mathbf{K}_g/\partial \alpha\right) \mathbf{u} \\ \frac{1}{2} \mathbf{u}^T (\partial \mathbf{K}_g / \partial \alpha) \mathbf{u} \end{bmatrix} \tag{33}$$

See a review on eigenvalue problem sensitivity [59]. In (33), the upper partition results from taking derivative of the eigenvalue equation in (30). The lower partition results from taking derivative of the scaling constraint $\mathbf{u}^T \mathbf{K}_g \mathbf{u} =$ constant. For non-singular $\mathbf{K}$ and $\mathbf{K}_g$ it can be shown that the coefficient matrix on the LHS is always invertible, and hence the stability of numerical solution is guaranteed; see Section 3.2 of [58]. Also, since the coefficient matrix is symmetric, efficient algorithms such as LU factorization can be used for solution. For a given eigen pair $(\lambda, \mathbf{u})$, the coefficient matrix remains the same regardless of the sensitivity parameter α under question, and so the factorization, which can be time-consuming for large structures, only needs to be performed once.

5.2.2. Uncertainty modeling

Let the floor loads W_i be independent Lognormal with mean w_i and coefficient of variation (c.o.v. = standard deviation/mean) δ_i. To fit into

the context of this work where the PDF of input random variables does not depend on sensitivity parameters, we represent the W_i's in terms of i.i.d. $N(0,1)$ variates X_i's as

$$W_i = w_i \, e^{a_i + b_i X_i} \tag{34}$$

where

$$a_i = -\ln \sqrt{1+\delta_i^2} \qquad b_i = \sqrt{\ln(1+\delta_i^2)} \tag{35}$$

are the mean and standard deviation of the $\ln(\cdot)$ of a Lognormal variate with mean 1 and standard deviation δ_i.

5.2.3. Analytical solution for benchmarking

For general structures, λ requires solving the eigenvalue problem in (30). For a shear building with $\mathbf{K}$ and $\mathbf{K}_g$ of the form in (31), however, all eigenvalues λ_i and eigenvectors $\mathbf{u}_i$ can be obtained analytically as

$$\lambda_i = \frac{k_i H_i}{W_i} \tag{36}$$

and $\mathbf{u}_i$ is a vector equal to 1 at all entries from i to n, and zero elsewhere. Essentially, the ith mode corresponds to buckling in the ith story only. This analytical result can be verified by noting that $\mathbf{K}\mathbf{u}_i$ gives a vector equal to k_i at the ith entry and zero elsewhere; the same for $\mathbf{K}_g\mathbf{u}_i$ except that the ith entry is W_i/H_i. The two vectors are therefore colinear, implying that $\mathbf{u}_i$ is an eigenvector. The ith row of $\mathbf{K}\mathbf{u}_i = \lambda_i \mathbf{K}_g \mathbf{u}_i$ reads $k_i = \lambda_i W_i/H_i$, which gives (36).

By noting $\lambda = \min_{1\le i\le n} \lambda_i$ and using (34), Y in (29) is now given by

$$Y = \lambda_0 \max_{1\le i\le n} \left\{ \frac{w_i}{k_i H_i} e^{a_i + b_i X_i} \right\} \tag{37}$$

Since $\{Y < y\}$ is equivalent to all terms in the $\max\{\cdot\}$ less than y, and the terms are independent, we have

$$P(Y<y) = \prod_{i=1}^{n} \Phi(\bar{x}_i) \tag{38}$$

where

$$\bar{x}_i = \frac{1}{b_i}\left[\ln\left(\frac{k_i H_i y}{w_i \lambda_0} \right) - a_i \right] \tag{39}$$

Using (38), $P(Y \ge y) = 1 - P(Y<y)$ and its sensitivities can be obtained analytically. See (41), (44) and (45) later in the numerical example.

5.2.4. Illustrative scenario

For numerical investigation, consider a shear building with $n = 5$ stories and the following nominal properties: uniform story height $H = 3.5$ m and stiffness $k = 250$ kN/mm, and i.i.d. Lognormal floor loads W_i with uniform mean $w = 100$ kN and c.o.v. 10%. The normalizing constant λ_0 in (29) is taken as the value of λ when W_i's are at their mean value. Equation (36) gives $\lambda_0 = 250 \cdot 3500/1000 = 875$.

We will investigate the sensitivity of $P(Y \geq y)$ w.r.t. the common mean of floor load $w \equiv \alpha_1$ and the second story stiffness $k_2 \equiv \alpha_2$, both about their nominal values (100 kN, 250 kN/mm). All other quantities are fixed at their nominal values. The different nature of α_1 and α_2 serves to demonstrate applicability to parameters that affect the distribution of input random variables (α_1) and those that affect the response directly (α_2).

When applying SS for sensitivity, to be consistent wtih typical applications, λ will be obtained by solving the eigenvalue problem in (30), and its derivative $\partial\lambda/\partial\alpha$ by solving (33). In the latter equation, it can be reasoned from (31) that $\partial\mathbf{K}/\partial\alpha_2$ has the same form as $\mathbf{K}$ with k_2 replaced by 1 and all other k_i's by 0. On the other hand, recall that $\mathbf{K}_g$ has the same form as $\mathbf{K}$ except that k_i is replaced by W_i/H_i, and $W_i = \alpha_1 \exp(a + bX_i)$. This implies that $\partial\mathbf{K}_g/\partial\alpha_1$ has the same form as $\mathbf{K}$ except that k_i is replaced by $H^{-1}\exp(a + bX_i)$.

For benchmarking, the analytical solution in (37) is given by ($n = 5$)

$$Y = \lambda_0 \max\left\{\frac{\alpha_1}{kH}e^{a+bX_i}, \frac{\alpha_1}{\alpha_2 H}e^{a+bX_2}, \cdots, \frac{\alpha_1}{kH}e^{a+bX_n}\right\} \tag{40}$$

where $a = -4.975 \times 10^{-3}$ and $b = 0.9975 \times 10^{-3}$ (4 digits) are obtained from (35) with mean 1 and c.o.v. 10%. Equation (38) then gives

$$P(Y < y) = \Phi(\bar{x}_1)^{n-1}\Phi(\bar{x}_2) \tag{41}$$

where

$$\bar{x}_1 = \frac{1}{b}\left[\ln\left(\frac{kHy}{\alpha_1\lambda_0}\right) - a\right] \tag{42}$$

$$\bar{x}_2 = \frac{1}{b}\left[\ln\left(\frac{\alpha_2 Hy}{\alpha_1\lambda_0}\right) - a\right] \tag{43}$$

Taking derivatives gives the sensitivities

$$\frac{\partial P(Y<y)}{\partial \alpha_1} = -\frac{1}{b\alpha_1}\left[(n-1)\frac{\phi(\bar{x}_1)}{\Phi(\bar{x}_1)} + \frac{\phi(\bar{x}_2)}{\Phi(\bar{x}_2)}\right] P(Y<y) \tag{44}$$

$$\frac{\partial P(Y<y)}{\partial \alpha_2} = \frac{1}{b\alpha_2}\frac{\phi(\bar{x}_2)}{\Phi(\bar{x}_2)} P(Y<y) \tag{45}$$

Clearly, $P(Y \geq y) = 1 - P(Y < y)$ and $\partial P(Y \geq y)/\partial \alpha_i = -\partial P(Y < y)/\partial \alpha_i$.

5.2.5. Numerical results

Figure 3 shows the results of a typical SS run, in the same manner as Figure 1. The CCDF in the top-left plot agrees with benchmark. In the second row, the left plot shows that the response gradient of α_1 is fully correlated with the response. This can be seen by taking α_1 outside the $\max\{\cdot\}$ in (40) so that it reads $Y = \alpha_1 \times (\cdot)$ and hence $\partial Y/\partial \alpha_1 = (\cdot)$ differs from Y only by a deterministic factor (α_1). The quality of sensitivity estimate for α_1 in the middle and right plot is similar to that for α_1 or α_2 in Figure 1 for Example 1. The sensitivity estimate for α_2 has a lower quality, as evidence from the middle and right plot at the bottom. This is attributed to the fact that the response gradient has a more complicated structure (hence estimation variance), as seen in the bottom left plot. Although the points at the lower part of the plot show a full correlation similar to what was seen for α_1, there are many points at zero response gradient. This can be explained by (40). When the second term is greater than all other terms, $Y = \alpha_2^{-1} \times (\cdot)$ and so $\partial Y/\partial \alpha_2 = -\alpha_2^{-2} \times (\cdot)$ is fully negatively correlated with Y. Otherwise Y does not depend on α_2 and so $\partial Y/\partial \alpha_2 = 0$. Intuitively, the second story stiffness (α_2) matters to the critical buckling load only when the second story is most critical. In the present case where $\alpha_1 = \alpha_2 = 1$, the terms in the $\max\{\cdot\}$ are i.i.d. and so the chance of the second story being critical is $1/n_s = 1/5$. In other words there is a 80% chance that $\partial Y/\partial \alpha_2 = 0$. This creates a disparity in the samples of response gradient, inflating the estimation variance in a similar manner as averaging indicator function values of rare events. Figure 4 shows the sample mean and $\pm 1\sigma$ bounds of sensitivity estimates, in the same manner as Figure 2. The statistics are consistent with the single run results in Figure 3.

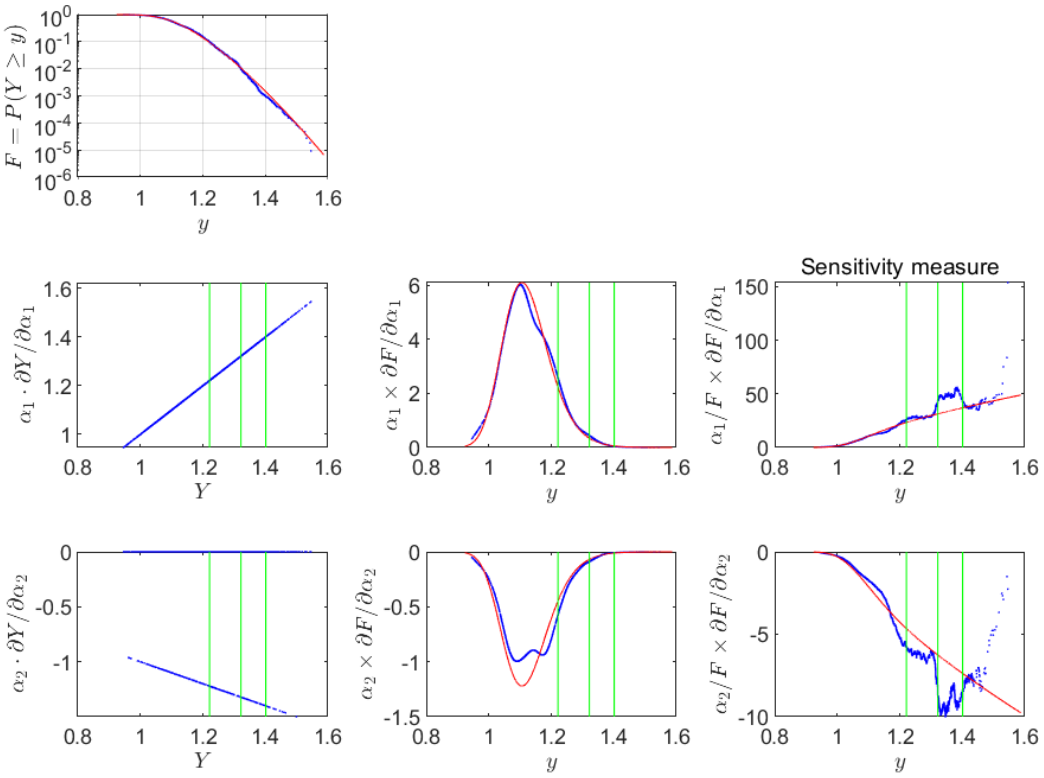

Figure 3: Example 2 (Shear building buckling). Results of a typical run. Same legend as Figure 1

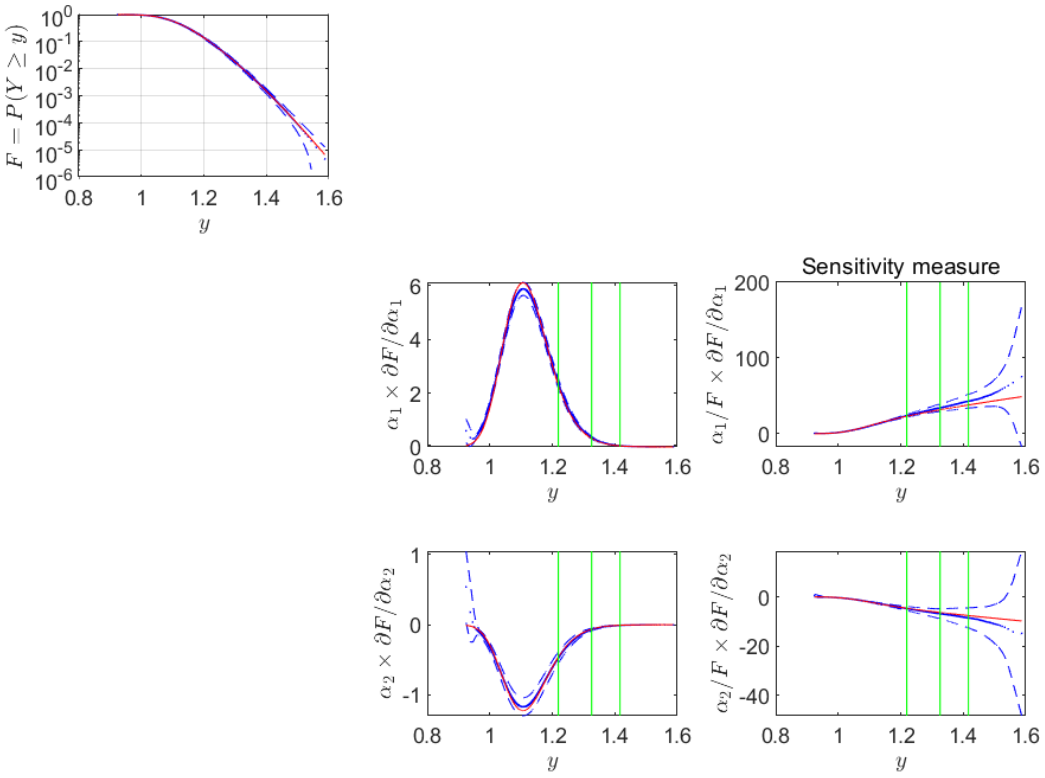

Figure 4: Example 2 (Shear building buckling). Sample mean (blue solid line) and $\pm 1\sigma$ bounds (blue dashed line) from 1000 SS runs. Same legend as Figure 2

5.3. Linear SDOF first passage problem

Consider a linear SDOF structure subjected to white noise excitation. The response Y is defined as

$$Y = \max_{0 \le t \le T} |u(t)| \tag{46}$$

where $T = 20$ s is the duration; $u(t)$ is the displacement starting from rest $u(0) = \dot{u}(0) = 0$ and thereafter follows the equation of motion

$$\ddot{u}(t) + 2\zeta\omega\dot{u}(t) + \omega^2 u(t) = W(t) \tag{47}$$

with $\omega = 2\pi$ (rad/s, natural frequency), $\zeta = 1\%$ (damping ratio), and $W(t)$ is white noise with two-sided power spectral density (PSD) $S = 0.86N^2/(rad/s)$. This example was considered recently in [56] when investigating the optimal choice of correlation parameter in conditional sampling MCMC, and in [36] on a theory for optimal MCMC for rare events. Here, the sensitivty of $P(Y \ge y)$ w.r.t. ζ and ω will be investigated. From (46), for a generic α,

$$\frac{\partial Y}{\partial \alpha} = \chi \, \frac{\partial u(t_*)}{\partial \alpha} \tag{48}$$

where t_* is the peak response time, i.e., when $|u|$ is maximum, and χ is the sign of $u(t_*)$ that accounts for the effect of $|\cdot|$. For general t, $\partial u(t)/\partial\alpha$ can be obtained by solving an equation resulting from differentiating (47) w.r.t. α. Let $u_\zeta(t) = \partial u(t)/\partial\zeta$ and $u_\omega(t) = \partial u(t)/\partial\omega$. Then (omitting t)

$$\ddot{u}_\zeta + 2\zeta\omega\dot{u}_\zeta + \omega^2 u_\zeta = -2\omega\dot{u} \tag{49}$$

$$\ddot{u}_\omega + 2\zeta\omega\dot{u}_\omega + \omega^2 u_\omega = -2\zeta\dot{u} - 2\omega u \tag{50}$$

These equations differ by the RHS only, and their LHS have the same functional form as the LHS of (47). When $u(t)$ is given, their RHS are known and they can be solved in the same way as (47). Computation is performed in discrete time at a time interval of $\Delta t = 0.05$ s and using `lsim` in Matlab. There are $n = T/\Delta t = 400$ input random variables in the discrete white noise $W_i = \sqrt{2\pi S/\Delta t}\, X_{i+1}$ $(0 \le i \le n-1)$, where $\{X_i\}_{i=1}^n$ are the i.i.d. N(0,1) input random variables.

5.3.1. Numerical investigation

In this example, analytical solution is not available. The benchmark for the CCDF of Y is obtained by direct Monte Carlo with 10^7 i.i.d. samples. Using the same set of input random variables $\mathbf{X}$, i.e., common random numbers, the sensitivity benchmark is obtained by central difference with a relative step of 1%. This requires an additoinal 10^7 response function evaluations for the forward step, and another 10^7 for the backward step.

Figure 5 shows the results for a typical run, in the same way as Figure 1. In the top-left plot the CCDF estimate by SS agrees with the benchmark, as expected. In the middle row, the left plot shows that the response gradient of α_1 (damping ratio) is negative, and is negatively correlated with the response. The former is intuitive; the latter is not trivial. The quality of sensitivity estimate in the middle and right plot is similar to those in Example 1, suggesting similar difficulty. This is not the case for α_2 (natural frequency) in the bottom row. The bottom middle plot shows a spurious fluctuation with y; the bottom right plot shows a significant departure from the benchmark as y increases. The difficulty with α_2 this time is more subtle than α_2 in Example 2, and the mechanism is not well-understood at the time of writing. Visual inspection of the scatter plot in the bottom left would suggest that the response gradient may be uncorrelated from the response. The bottom right plot defies this, as the sensitivity there has a similar order of magntiude as α_1 (damping ratio). In this sense, the natural frequency (α_2) is a more difficult parameter than the damping ratio (α_1).

Figure 6 shows the sample mean and $\pm 1\sigma$ bounds of sensitivity estimates, in the same way as Figure 2. The results are consistent with those in Figure 5 for a single run. The $\pm 1\sigma$ gap for the sensitivity of α_2 (natural frequency) is one order of magnitude wider than that for α_1 (damping ratio).

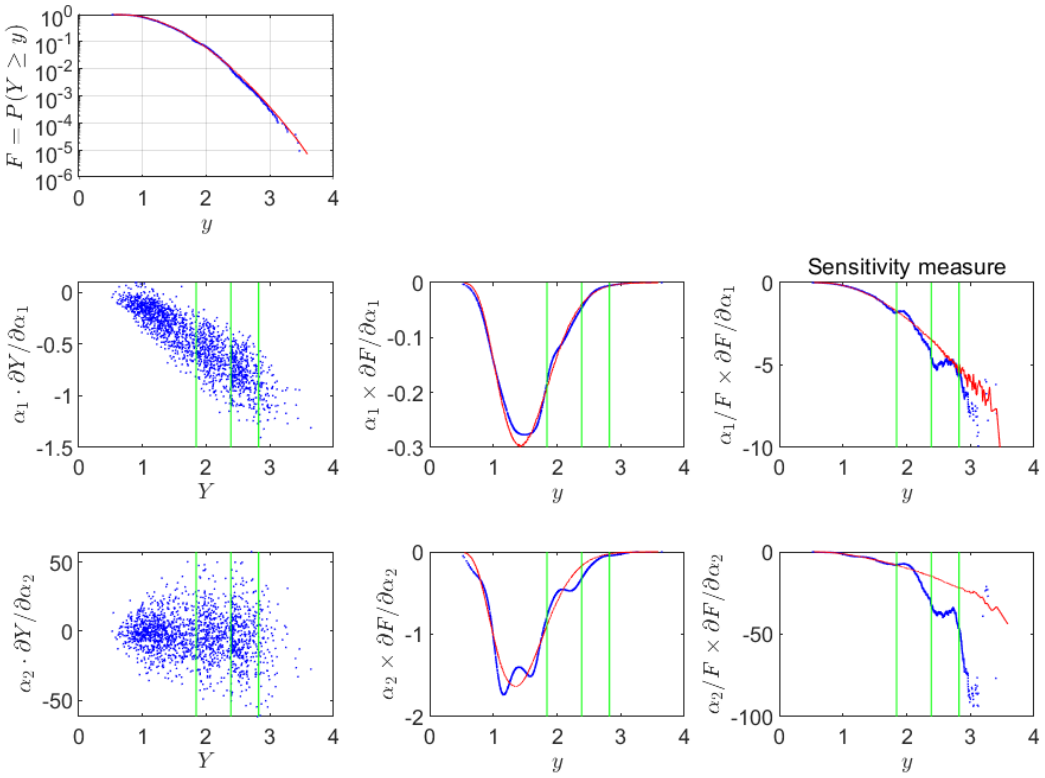

Figure 5: Example 3 (SDOF first passage). Results of a typical run. Same legend as Figure 1

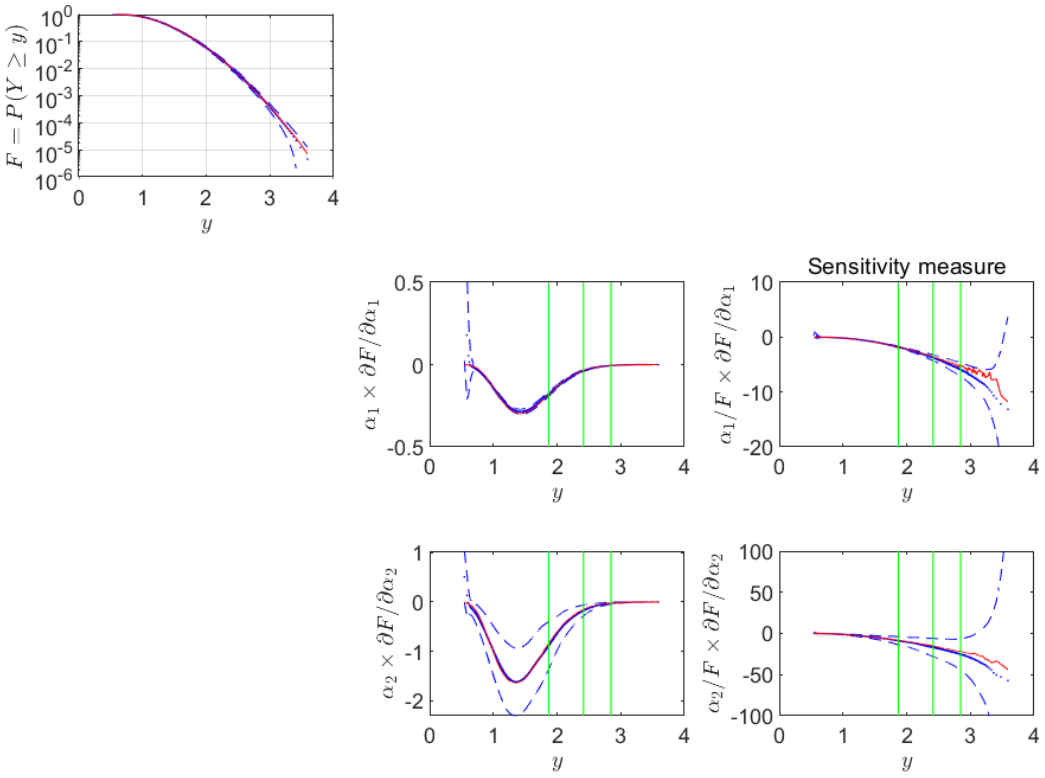

Figure 6: Example 3 (SDOF first passage). Sample mean (blue solid line) and $\pm 1\sigma$ bounds (blue dashed line) from 1000 SS runs. Same legend as Figure 2

5.4. Pile serviceability design

This example is adapted from Section 6 of [60] that investigated reliability-based design with SS implemented in spreadsheets. Consider assessing the 'failure' as inadequacy of a concrete pile in loose sand against serviceability limit state (SLS) of excessive settlement. The pile has a circular section with diameter $B = 0.9$ m and depth $D = 8$ m into the soil. The latter has uncertainty in the friction angle ϕ' that depends on depth z. The dash here indicates that the quantity is associated with effective stress, a convention in geotechnical engineering. The friction angle is modeled as a spatially stationary Lognormal random field with mean μ and c.o.v. δ. The correlation between $\ln \phi_i$ at depth z_i and $\ln \phi_j$ at z_j is assumed to be

$$R_{ij} = \exp\left(-\frac{2}{\lambda}|z_i - z_j|\right) \tag{51}$$

where $\lambda = 4$ m is a correlation length. For computation, a soil depth of 12 m is considered, which is sufficient to cover the relevant influence zone of resisance (see later). Dividing it into differential segments of equal thickness $d = 0.1$ m, the uncertainty in frictional angle is characterized by the values $\{\phi'_i\}_{i=1}^n$ at depths $z_i = (i - 1/2)d$, where $n = 120$ is the number of random variables. The random vector $\boldsymbol{\phi}' = [\phi'_1, \ldots, \phi'_n]^T$ can be represented by a i.i.d. N(0,1) vector $\mathbf{X}$ by

$$\boldsymbol{\phi}' = \mu \exp\left(u\mathbf{1} + s\mathbf{L}\mathbf{X}\right) \tag{52}$$

where $\exp(\cdot)$ operates entry by entry; $\mathbf{1}$ denotes an $n \times 1$ vector of ones; $u = -\ln\sqrt{1+\delta^2}$ and $s = \sqrt{\ln(1+\delta^2)}$ are respectively the mean and standard deviation of the $\ln(\cdot)$ of a Lognormal variate with mean 1 and standard deviation δ; and $\mathbf{L}$ is the lower triangular matrix from Cholesky decomposition $\mathbf{R} = \mathbf{L}\mathbf{L}^T$, where $\mathbf{R}$ is the $n \times n$ correlation matrix of $\{\ln \phi'_i\}_{i=1}^n$ with the (i, j)-entry given by (51).

The response Y is defined as the ratio of load to effective resistance, the higher the worse. We will investigate the sensitivity of $P(Y \geq y)$ w.r.t. the pile diameter and the mean of the friction angle, i.e., $\alpha_1 \equiv B$ and $\alpha_2 \equiv \mu$. Following the notation in Section 6 of [60], the response is given by

$$Y = \frac{F_{50}}{Q_{sls}} \tag{53}$$

where F_{50} is the design load,

$$Q_{sls} = 0.625\, a \left(\frac{y_a}{B}\right)^b (Q_{side} + Q_{tip} - W) \tag{54}$$

$$Q_{side} = \pi B\, d\, (K/K_0)_n K_0 \sum_{i=1}^{N_s} \sigma_i' \tan \phi_i' \tag{55}$$

$$Q_{tip} = 0.25\pi B^2 \left[0.5B(\gamma - \gamma_w) N_\gamma \zeta_{\gamma s} \zeta_{\gamma d} \zeta_{\gamma r} + D(\gamma - \gamma_w) N_q \zeta_{qs} \zeta_{qd} \zeta_{qr}\right] \tag{56}$$

$$W = 0.25\pi B^2 D\, (\gamma_c - \gamma_w) \tag{57}$$

are respectively the resistance capacity at SLS, side resistance, tip resistance, and effective pile weight. The parameters appearing in these equations are summarized in Table 3, categorized by their nature. The top category shows the source of uncertainty, i.e., $\{\phi_i'\}_{i=1}^n$. The next category contains the ones affected by $\{\phi_i'\}_{i=1}^n$, essentially through the mean friction angle $\bar{\phi}'$. It also contains the pile diameter B, which is a sensitivity parameter (α_1). The parameters at the bottom category are constants.

Table 3: Parameters in (54) to (57) of the pile example in Section 5.4

Symbol	Value/formula	Description
ϕ'_i, $1 \le i \le n$	Lognormal random, mean $\mu = 0.5585$ rad ($32°$), c.o.v. $\delta = 17\%$, correlation length $\lambda = 4$ m. See (52)	Friction angle at depth $z_i = (i - 1/2)\,d$; $n = 120$ (no. of discretized layers), $d =$ 0.1 m (layer thickness)
$\bar{\phi}'$		Average ϕ in tip resistance influence zone, from $\min\{8B, D\}$ above to $3.5B$ below pile tip
B	0.9 m	Pile diameter
N_q	$\tan^2(\pi/4 + \bar{\phi}'/2)\exp(\pi \tan \bar{\phi}')$	Bearing capacity factor (load)
N_γ	$2(N_q + 1)\tan \bar{\phi}'$	Bearing capacity factor (weight)
ζ_{qs}	$1 + \tan \bar{\phi}'$	Correction factor
ζ_{qd}	$1 + 2(\tan \bar{\phi}')(1 - \sin \bar{\phi}')^2 \tan^{-1}(D/B)$	Correction factor
N_s	Integer part of D/d	No. of differential layers in side resistance influence zone of length D
σ'_i	$(\gamma - \gamma_w)z_i$	Effective stress at depth z_i
F_{50}	800 kN	Design load
D	8 m	Pile depth
y_a	25 mm	Allowable displacement
γ	20 kN/m^3	Unit weight of soil
γ_w	9.81 kN/m^3	Unit weight of water
γ_c	24 kN/m^3	Unit weight of concrete
$(K/K_0)_n$	1	Nominal operative in-situ horizontal stress coefficient ratio
K_0	1	At-rest coefficient of horizontal soil stress
a, b	4, 0.4	Constants
$\zeta_{\gamma s}, \zeta_{\gamma d}, \zeta_{\gamma r}, \zeta_{qr}$	0.6, 1, 1, 1	Correction factors

In this example, the response derivatives $\partial Y/\partial \alpha_i$ are computed by central difference with a relative step size of 0.1%. This is the case in applications where only response function values are available. Strictly speaking, this will induce an additional bias in the resulting sensitivity estimate, although it is negligible compared to the bias due to kernel smoothing. Note that B ($\equiv \alpha_1$) affects Q_{sls}, Q_{side} and Q_{tip} directly through its appearance in (54) to (56), and indirectly through $\bar{\phi}'$, because the latter is averaged over the tip resistance influence zone from $\min\{8B, D\}$ above to $3.5B$ below the pile tip; see Table 3. On the other hand, μ ($\equiv \alpha_2$) is the mean value of $\{\phi'_i\}_{i=1}^n$, which in turn affects Q_{side} through $\tan(\cdot)$ in the sum of (55), and Q_{tip} in (56) through $\bar{\phi}'$ that affects N_q, N_γ, ζ_{qs} and ζ_{qd}, also in a trigonometric manner; see the second category in Table 3.

5.4.1. Numerical results

Figures 7 and 8 show the results of a single run and statistics from 1000 indepenent runs, analogous to Figures 1 and 2 of Example 1. Similar to

Example 3 (SDOF first passage), analytical solution for failure probability or sensitivity is not available. The benchmark is taken as the central difference (relative step size 1%) of Monte Carlo estimates for $P(Y \geq y)$, each based on 10^7 samples. Generally, the quality of sensitivity estimates is similar to that of Example 1, reflecting sensitivity parameters of similar difficulty. For both α_1 ($\equiv B$) and α_2 ($\equiv \mu$), the response is negatively correlated with its gradient. The former has a larger scatter. These are not trivial to explain. Both parameters have a reliability sensitivity in the order of a few tens. At a failure probability of 0.1%, i.e., $y \approx 1$, the sensitivities (right column) of α_1 and α_2 are about 25 and 50, respectively. This implies that, e.g., a 1% change in α_1 will lead to about 25% change in $P(Y \geq y)$. Also, α_2 is twice as sensitive than α_1.

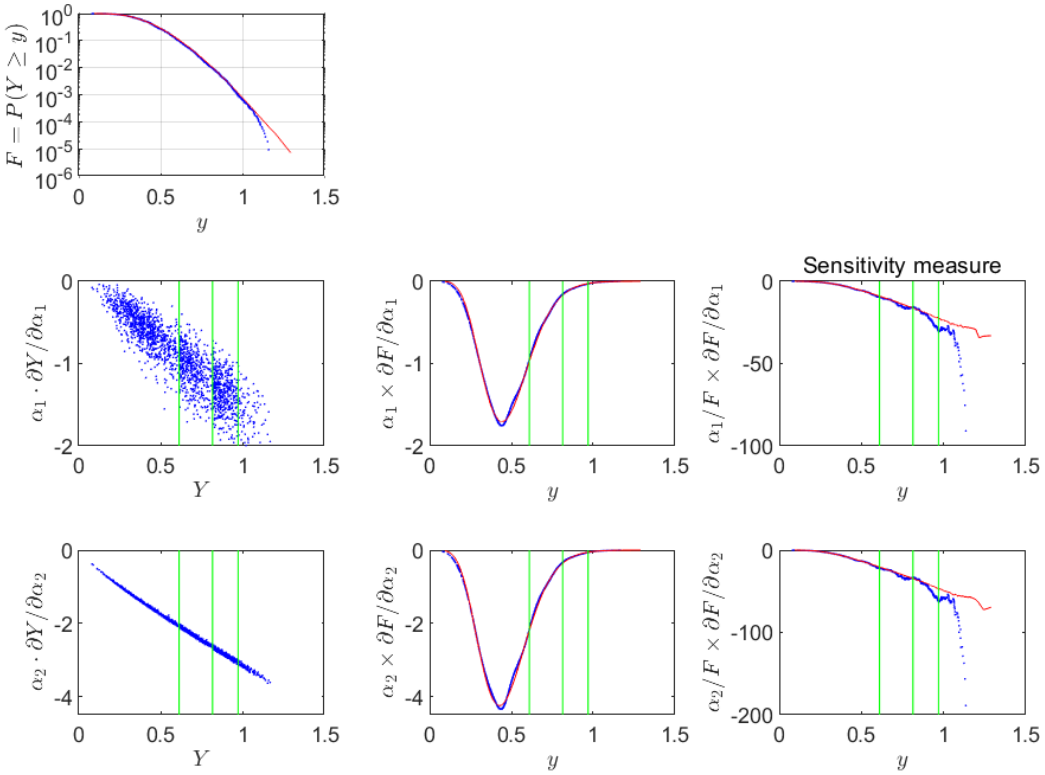

Figure 7: Example 4 (pile design). Results of a typical run. Same legend as Figure 1

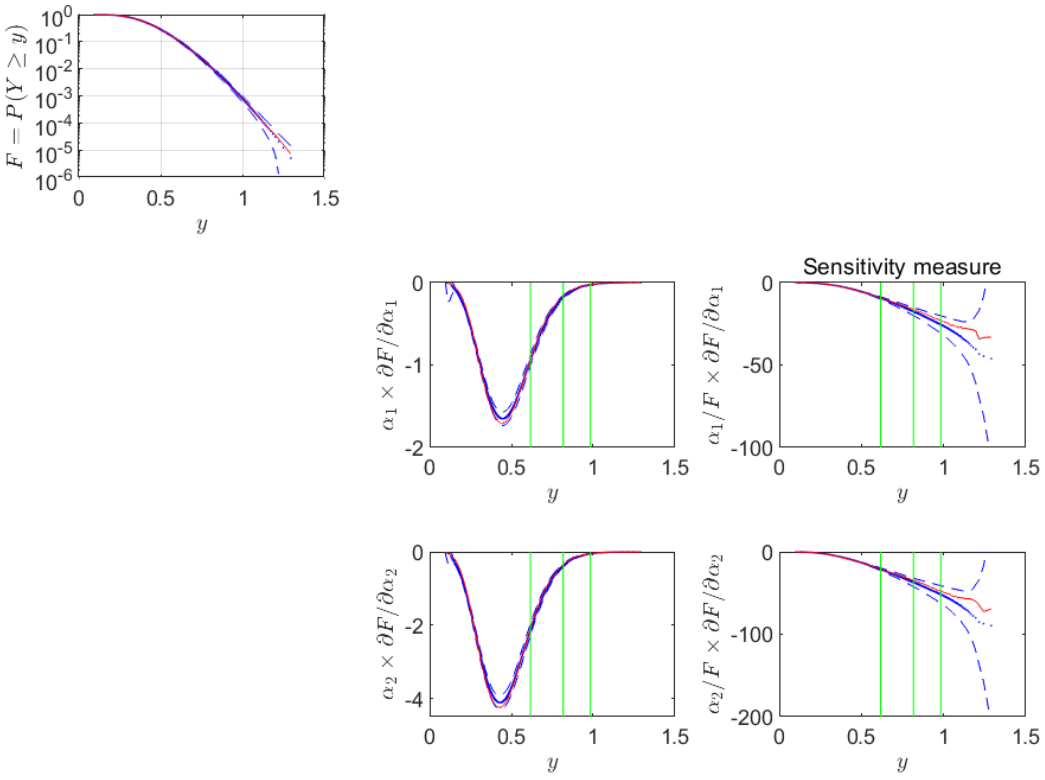

Figure 8: Example 4 (pile design). Sample mean (blue solid line) and $\pm 1\sigma$ bounds (blue dashed line) from 1000 SS runs. Same legend as Figure 2

6. Conclusions

As the key theoretical contribution, this work has derived a formula that expresses reliability sensitivity $\partial P(Y \geq y)/\partial \alpha$ in terms of response gradient $\partial Y/\partial \alpha$. See (4) as an integral and (6) as a conditional expectation. Based on the formula a Monte Carlo method has been developed to calculate $\partial P(Y \geq y)/\partial \alpha$ together with $P(Y \geq y)$ in the same run and for all generated values of y. See (9), which leverages Subset Simulation for generating rare event samples, and kernel smoothing for resolving the issue of zero-probability conditioning. The latter is at the expense of a bias, but which can be controlled by the kernel width that is chosen to balance over estimation bias and variance. See (20) a simple choice used in this work. The theory and Monte Carlo method have been investigated in a number of examples, illustrating sensitivity parameters of different nature that prompt attention, e.g., α_3 in

Example 1 (Normal response) has non-zero response gradient but zero sensitivity; α_2 in Example 2 (buckling) has a significant chance of zero response gradient; α_2 in Example 3 (first passage) displays weak correlation between response and its gradient but still carries significant reliability sensitivity; and α_1 (pile diameter) whose sensitivity is intuitive but not trivial.

6.1. Limitations and future work

Limitations and potential future work are in order. The theory assumes the response function to be sufficiently smooth to justify the existence of gradient and continuity of their joint/marginal PDFs. The Monte Carlo method has not exploited the relationship between the response and its gradient for variance reduction, which can be a future topic. The use of kernel smoothing was motivated by the integral form in (4), but it is not the only means. For example, the expectation form in (6) motivates the general perspective of regressing $\partial Y/\partial\alpha$ against Y, which opens up the possibility of data analytics such as Gaussian process regression. In this regard, the input random variables $\mathbf{X}$ may be seen as hidden variables, and it may be possible to incorporate their information for better estimating the conditional expectation in (6). On extension of theory, it should be possible to obtain the second derivative $\partial^2 P(Y \geq y)/\partial\alpha_i\partial\alpha_j$. The derivation should be similar to Section 7.3 of [36] and the result should be similar to (if not the same as) (13) there. It has not been done in this work because there are questions that deserve a separate study, e.g., how/whether second derivatives are used in applications, and the required precision; whether kernel smoothing can still be used to resolve zero-probability conditioning, especially for the derivative term in (13) there.

6.2. Future vision with gradient

This work is part of an effort to explore the use of gradient information for reliability calculations and related objectives, hoping to break bottlenecks when only response values are used. Such information may be limited in the past, but can become a norm in the near future, thanks to the advent of computer hardware (e.g., GPU), computer codes (e.g., Pytorch) and increasingly frequent use of neural network models. A future vision is that system response will be computed together with gradients in an embedeed manner for various purposes, e.g., reliability sensitivity, model updating, training surrogate models, and design optimization. This calls for a vitalization in reliability methodologies to be in par with future response calculation codes.

7. Acknowledgments

The research presented in this paper is supported by Academic Research Fund Tier 1 (RG68/22) from the Ministry of Education, Singapore. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not reflect the views of the funders.

References

[1] O. D. Ditlevsen, H. O. Madsen, Structural Reliability Methods, John Wiley & Sons, 1996.

[2] R. E. Melchers, A. T. Beck, Structural Reliability Analysis and Prediction, John Wiley & Sons, Ltd, 2017.

[3] A. D. Kiureghian, Structural and System Reliability, Cambridge Press, 2022.

[4] B. Ellingwood, M. Maes, F. M. Bartlett, A. T. Beck, C. Caprani, A. D. Kiureghian, L. Dueñas-Osorio, N. Galvão, R. Gilbert, J. Li, J. Matos, Y. Mori, I. Papaioannou, R. Parades, D. Straub, B. Sudret, Development of methods of structural reliability, Structural Safety 113 (3 2025). doi:10.1016/j.strusafe.2024.102474.

[5] A. Teixeira, E. Zio, H. Z. Huang, Special issue on safety, reliability and availability of complex systems and structures, Reliability Engineering and System Safety (2023).

[6] W. Hu, S. Cheng, J. Yan, J. Cheng, X. Peng, H. Cho, I. Lee, Reliability-based design optimization: a state-of-the-art review of its methodologies, applications, and challenges, Structural and Multidisciplinary Optimization 67 (2024) 168. doi:10.1007/s00158-024-03884-x.
URL `https://link.springer.com/10.1007/s00158-024-03884-x`

[7] R. Kosowski, S. Neftci, Principles of Financial Engineering, 3rd Edition, Academic Press, 2015.

[8] H. S. Li, Y. Z. Ma, Z. Cao, A generalized Subset Simulation approach for estimating small failure probabilities of multiple stochastic responses, Computers and Structures 153 (2015) 239–251. doi:10.1016/j.compstruc.2014.10.014.

[9] W. Xia, Z. Liao, Enhanced generalized Subset Simulation with multiple importance sampling for reliability estimation, Computers and Structures 313 (6 2025). doi:10.1016/j.compstruc.2025.107741.

[10] I. Papaioannou, C. Papadimitriou, D. Straub, Sequential importance sampling for structural reliability analysis, Structural Safety 62 (2016) 66–75.

[11] A. Deo, K. Murthy, Achieving efficiency in black-box simulation of distribution tails with self-structuring importance samplers, Operations Research (1 2023). doi:10.1287/opre.2021.0331.

[12] S. K. Au, Y. Wang, Engineering Risk Assessment with Subset Simulation, John Wiley & Sons, 2014.

[13] K. M. Zuev, Subset Simulation Method for Rare Event Estimation: An Introduction, Springer, 2021, pp. 1–25.

[14] M. de Angelis, E. Patelli, M. Beer, Advanced line sampling for efficient robust reliability analysis, Structural Safety 52 (2015) 170–182. doi:10.1016/j.strusafe.2014.10.002.

[15] C. Dang, M. A. Valdebenito, P. Wei, J. Song, M. Beer, Bayesian active learning line sampling with log-normal process for rare-event probability estimation, Reliability Engineering & System Safety 246 (2024) 110053.

[16] F. Xin, P. Wang, Y. Chen, R. Yang, F. Hong, A new uncertainty reduction-guided single-loop kriging coupled with Subset Simulation for time-dependent reliability analysis, Reliability Engineering and System Safety 261 (9 2025). doi:10.1016/j.ress.2025.111065.

[17] H. Dai, D. Li, M. Beer, Adaptive kriging-assisted multi-fidelity Subset Simulation for reliability analysis, Computer Methods in Applied Mechanics and Engineering 436 (3 2025). doi:10.1016/j.cma.2024.117705.

[18] C. Dang, M. Beer, Semi-Bayesian active learning quadrature for estimating extremely low failure probabilities, Reliability Engineering & System Safety 246 (2024) 110052.

[19] P. Rajak, P. Roy, Efficient computing technique for reliability analysis of high-dimensional and low-failure probability problems using active learning method, Probabilistic Engineering Mechanics 77 (2024) 103662.

[20] Y. Xu, S. Kohtz, J. Boakye, P. Gardoni, P. Wang, Physics-informed machine learning for reliability and systems safety applications: State of the art and challenges, Reliability Engineering and System Safety 230 (2 2023). doi:10.1016/j.ress.2022.108900.

[21] Y. Shi, P. Wei, K. Feng, D.-C. Feng, M. Beer, A survey on machine learning approaches for uncertainty quantification of engineering systems, Machine Learning for Computational Science and Engineering 1 (2025) 11. doi:10.1007/s44379-024-00011-x.
URL https://link.springer.com/10.1007/s44379-024-00011-x

[22] P. Glasserman, D. D. Yao, Some guidelines and guarantees for common random numbers, Management Science 38 (1992) 884–908.

[23] M., J. F. Hu, Conditional Monte Carlo: Gradient Estimation and Optimization Applications, Springer, 1997. doi:10.1007/978-1-4614-6293-1.

[24] Z. Lu, S. Song, Z. Yue, J. Wang, Reliability sensitivity method by line sampling, Structural Safety 30 (2008) 517–532. doi:10.1016/j.strusafe.2007.10.001.

[25] M. A. Valdebenito, H. A. Jensen, H. B. Hernández, L. Mehrez, Sensitivity estimation of failure probability applying line sampling, Reliability Engineering and System Safety 171 (2018) 99–111. doi:10.1016/j.ress.2017.11.010.

[26] M. A. Valdebenito, H. B. Hernández, H. A. Jensen, Probability sensitivity estimation of linear stochastic finite element models applying line sampling, Structural Safety 81 (11 2019). doi:10.1016/j.strusafe.2019.06.002.

[27] S. Song, Z. Lu, H. Qiao, Subset Simulation for structural reliability sensitivity analysis, Reliability Engineering and System Safety 94 (2009) 658–665. doi:10.1016/j.ress.2008.07.006.

[28] H. A. Jensen, F. Mayorga, C. Papadimitriou, Reliability sensitivity analysis of stochastic finite element models, Computer Methods in Applied Mechanics and Engineering 296 (2015) 327–351. doi:10.1016/j.cma.2015.08.007.

[29] C. Proppe, Local reliability based sensitivity analysis with the moving particles method, Reliability Engineering and System Safety 207 (3 2021). doi:10.1016/j.ress.2020.107269.

[30] K. Breitung, Asymptotic Approximations for Probability Integrals, Springer-Verlag, 1994.

[31] I. Papaioannou, K. Breitung, D. Straub, Reliability sensitivity estimation with sequential importance sampling, Structural Safety 75 (2018) 24–34. doi:10.1016/j.strusafe.2018.05.003.

[32] A. J. Torii, On sampling-based schemes for probability of failure sensitivity analysis, Probabilistic Engineering Mechanics 62 (10 2020). doi:10.1016/j.probengmech.2020.103099.

[33] A. J. Torii, A. A. Novotny, A priori error estimates for local reliability-based sensitivity analysis with monte carlo simulation, Reliability Engineering and System Safety 213 (9 2021). doi:10.1016/j.ress.2021.107749.

[34] M. Chiron, J. Morio, S. Dubreuil, A review on local failure probability sensitivity analysis, Applied Sciences (Switzerland) 13 (11 2023). doi:10.3390/app132112021.

[35] S. K. Au, Augmenting approximate solutions for consistent reliability analysis, Probabilistic Engineering Mechanics 22 (2007) 77–87. doi:10.1016/j.probengmech.2006.08.004.

[36] S. K. Au, Optimality conditions for MCMC in rare event risk analysis, Reliability Engineering & System Safety 265 (2026) 111539. doi:10.1016/j.ress.2025.111539.

[37] S. K. Au, J. L. Beck, Estimation of small failure probabilities in high dimensions by Subset Simulation, Probabilistic Engineering Mechanics 16 (2001) 263–277.

[38] I. Papaioannou, W. Betz, K. Zwirglmaier, D. Straub, Mcmc algorithms for subset simulation, Probabilistic Engineering Mechanics 41 (2015) 89–103.

[39] S. K. Au, E. Patelli, Rare event simulation in finite-infinite dimensional space, Reliability Engineering and System Safety 148 (2016) 67– 77.

[40] S. Duane, A. D. Kennedy, B. J. Pendleton, D. Roweth, Hybrid monte carlo, Physics Letters B 195 (1987) 216–222.

[41] Z. Wang, M. Broccardo, J. Song, Hamiltonian Monte Carlo methods for Subset Simulation in reliability analysis, Structural Safety 76 (2019) 51–67.

[42] M. Girolami, B. Calderhead, Riemann manifold Langevin and Hamiltonian Monte Carlo methods, Journal of Royal Statistical Society B 73 (2011) 123–214.

[43] W. Chen, Z. Wang, M. Broccardo, J. Song, Riemannian Manifold Hamiltonian Monte Carlo based Subset Simulation for reliability analysis in non-gaussian space, Structural Safety 94 (2022) 102134.

[44] J. Goodman, J. Weare, Ensemble samplers with affine invariance, Communications in Applied Mathematics and Computational Science 5 (2010) 65–80.

[45] M. D. Shields, D. G. Giovanis, V. S. Sundar, Subset Simulation for problems with strongly non-Gaussian, highly anisotropic, and degenerate distributions, Computers and Structures 245 (2021) 106431.

[46] Y. Tian, A. Fu, H. Zhang, L. Tang, J. Sun, Accelerated verification of autonomous driving systems based on adaptive Subset Simulation, IEEE Transactions on Intelligent Vehicles (2024) 1–11doi:10.1109/TIV.2024.3449947.

[47] J. Finkel, P. A. O'Gorman, Bringing statistics to storylines: Rare event sampling for sudden, transient extreme events, Journal of Advances in Modeling Earth Systems 16 (6 2024). doi:10.1029/2024MS004264.

[48] P. C. A. Simon, C. T. Icenhour, G. Singh, A. D. Lindsay, C. Bhave, L. Yang, A. Riet, Y. Che, P. Humrickhouse, P. Calderoni, M. Shimada, Moose-based Tritium migration analysis program, version 8 (TMAP8) for advanced open-source Tritium transport and fuel cycle modeling, Fusion Engineering and Design 214 (5 2025). doi:10.1016/j.fusengdes.2025.114874.

[49] D. Scott, On optimal and data-based histograms, Biometrika 66 (1979) 605–610.

[50] D. Scott, Multivariate Density Estimation: Theory, Practice, and Visualization, John Wiley & Sons, 1992.

[51] E. Nadaraya, On estimating regression, Theory of Probability and its Applications 9 (1964) 141–142.

[52] H. Bierens, Topics in advanced econometrics: estimation, testing, and specification of cross-section and time series models, Cambridge University Press, 1994.

[53] M. Wand, M. Jones, Kernel Smoothing, Chapman and Hall/CRC, 1994.

[54] C. G. Atkeson, A. W. Moore, S. Schaal, Locally weighted learning, Artificial Intelligence Review 11 (1997) 11–73.

[55] P. Čížek, S. Sadıkoğlu, Robust nonparametric regression: A review, Wiley Interdisciplinary Reviews: Computational Statistics 12 (5 2020). doi:10.1002/wics.1492.

[56] S. K. Au, X. Zhou, Adaptive proposal length scale in Subset Simulation, Reliability Engineering and System Safety 261 (9 2025). doi:10.1016/j.ress.2025.111069.

[57] Y. Jiang, X. Zhang, M. Beer, M. G. Faes, C. Papadimitriou, H. Zhou, First excursion probability of dynamical systems: A review on computational methods, Mechanical Systems and Signal Processing 232 (6 2025). doi:10.1016/j.ymssp.2025.112751.

[58] I. Lee, G. Jung, An efficient algebraic method for the computation of natural frequency and mode shape sensitivities-part i. distinct natural frequencies, Computers & Structures 62 (1997) 429–435.

[59] R. M. Lin, J. E. Mottershead, T. Y. Ng, A state-of-the-art review on theory and engineering applications of eigenvalue and eigenvector derivatives, Mechanical Systems and Signal Processing 138 (4 2020). doi:10.1016/j.ymssp.2019.106536.

[60] Y. Wang, Z. Cao, Expanded reliability-based design of piles in spatially variable soil using efficient Monte Carlo simulations, Soils and Foundations 53 (2013) 820–834. doi:10.1016/j.sandf.2013.10.002.