Journal of Computational Physics 229 (2010) 8966–8980

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Evaluation of failure probability via surrogate models

Jing Li, Dongbin Xiu *

Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA

ARTICLE INFO

Article history:
Received 22 March 2010
Received in revised form 24 June 2010
Accepted 16 August 2010
Available online 26 August 2010

Keywords:
Failure probability
Sampling
Polynomial chaos
Stochastic computation

ABSTRACT

Evaluation of failure probability of a given system requires sampling of the system response and can be computationally expensive. Therefore it is desirable to construct an accurate surrogate model for the system response and subsequently to sample the surrogate model. In this paper we discuss the properties of this approach. We demonstrate that the straightforward sampling of a surrogate model can lead to erroneous results, no matter how accurate the surrogate model is. We then propose a hybrid approach by sampling both the surrogate model in a "large" portion of the probability space and the original system in a "small" portion. The resulting algorithm is significantly more efficient than the traditional sampling method, and is more accurate and robust than the straightforward surrogate model approach. Rigorous convergence proof is established for the hybrid approach, and practical implementation is discussed. Numerical examples are provided to verify the theoretical findings and demonstrate the efficiency gain of the approach.

1. Introduction

Computing failure probability of a given system is a fundamental problem in many fields such as risk management, structural design, reliability based optimization, etc. The task can be easily formulated mathematically, for it is essentially a problem of evaluating multivariate integrals in domains defined by failure modes. However, in practice it is challenging to accomplish. The reasons are mainly two folds: one is that the domain can be (highly) irregular, and the other is that the integrals often reside in high-dimensional spaces.

The most straightforward approach is the direct Monte Carlo sampling (MCS). An estimate of the failure probability can be readily obtained by counting the number of samples that lie within the failure domain. The concern is the computational cost, for the total number of samples required is proportional to the inverse of the failure probability. A rule of thumb is that, in order to obtain an estimate with reasonable accuracy one needs to generate at least 10 failure samples. Therefore for reliable systems (those with small failure probability), a very large number of samples are required for MCS. Since each sample point requires a simulation of the underlying system, MCS can be prohibitively time consuming. A number of improvements have been proposed, e.g., directional sampling [3,8,20,21], (adaptive) importance sampling [1], line sampling [18], subset simulation [2], etc.

For complex systems one usually can only afford a very limited number of simulations. Hence none of the sampling based methods seem affordable. Non-sampling approaches were then investigated. The most widely used methods are the First Order Reliability Method (FORM) and the Second Order Reliability Method (SORM). In both methods, the closest point (geometrical interpretation) on the limit state function, the function that separates failure and safe regions, to the origin in the standard Gaussian space is first located. This point is often called the most probable point (MPP), the beta point, or the design point. The limit state function is then approximated close to the MPP by a linear function (in FORM) or a quadratic function

* Corresponding author. Tel.: +1 765 496 2846.
E-mail addresses: li176@math.purdue.edu (J. Li), dxiu@purdue.edu (D. Xiu).

0021-9991/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2010.08.022

(in SORM), and the resulting approximate integral is further approximated by asymptotic method. See, for example, [16,11,7,15]. Many improvements have been made since their early developments [5,6,17,19,36]. And it is generally regarded that FORM and SORM are very effective methods for weakly nonlinear problems, especially in their asymptotic regimes [24,27,26].

A different approach, which has the flavor of both sampling methods and non-sampling methods, is response surface method (RSM). In RSM, one first conducts (usually) deterministic sampling of the system and then constructs a response surface as an approximation of the limit state function. An estimate of the failure probability is then obtained by sampling the response surface. Since the response surface is of analytic form, usually a multivariate polynomial, sampling is achieved with no simulation effort and can be done with a large number of samples. The idea has been pursued long ago [29,10,4], but has not shown significant advantages in a large class of problems over the FORM/SORM. It is widely acknowledged that the construction of the response surface is crucial. This is a challenging problem and it is where most later developments have focused on [25,12,14,22,23].

The method proposed in this paper also relies on response surface, which will be termed by a synonymous phrase *surrogate*. This is done intentionally because traditionally response surfaces are constructed by methods based on regressions, design of experiments, or some statistical algorithms, most of which are heuristic. Here we will employ the notion of surrogate to include some of the more recent constructions that are mathematically rigorous and with established convergence properties and error analysis. In particular, we introduce generalized polynomial chaos (gPC) methodology as a surrogate model. The gPC method, an extension of the seminal work on polynomial chaos by R. Ghanem (cf. [13]), has become one of the most widely used methods for stochastic computation since its introduction in [34]. It is based on multidimensional approximation theory and in many problems more efficient than other methods. Unlike the traditional response surfaces, which usually take the form of linear or quadratic polynomials, the gPC expansions can be readily constructed for high orders and result in superior accuracy. Many work has been devoted to gPC methods. For a detailed overview, see [31,32].

The real contribution of this paper, however, lies beyond the use of gPC expansion. (In fact, using gPC as a surrogate model has been attempted in [22,23].) Here we first demonstrate that *the existing straightforward surrogate (or response surface) approach can result in erroneous estimates for the failure probability, no matter how accurate the surrogate model is.* Consequently the lack of robustness of RSM is a fundamental issue rather than an accuracy issue. Having identified the core issue, we then present a hybrid approach that combines sampling of the surrogate and the original system. The idea is that in the most part of the probability space that are reasonably "away" from the limit state, samples of the surrogate (which are easy to obtain) are used, and only in the region "close" to the limit state are the samples the original system used. By doing so, only a (very) small portion (proportional to the failure probability) of the samples are from the original system and therefore the overall cost is much smaller than sampling the system in the entire probability space. On the other hand, the use of these small number of samples of the original system can prevent the loss of accuracy of the direct surrogate approach. Consequently the hybrid approach is much more efficient than the direct MCS and offers higher accuracy compared to the direct surrogate approach. Rigorous analysis is conducted and we present convergence proof for the hybrid approach. The analysis reveals a desirable feature of the hybrid approach. That is, the convergence of the failure probability estimate does not require the convergence of the surrogate model. This is a useful fact in practice because it allows one to use any surrogate model with finite accuracy. We then proceed to propose an alternative implementation of the hybrid method. This is an iterative algorithm that is more flexible and efficient, although the accuracy of the underlying surrogate model will affect the accuracy of the failure probability estimate. The properties of the two algorithms are discussed.

The paper is organized as follows. After setting up the problem in Section 2, we briefly review the MCS and RSM methods in Section 3, two of the existing methods that are directly related to our algorithms, and present arguments of the invalidity of the traditional RSM methods. The new hybrid method is proposed in Section 4, where both the numerical algorithms and the convergence proof are presented. A set of numerical examples are shown in Section 5 to examine the properties of the hybrid method, before the paper is concluded in Section 6.

2. Problem setup

Let $Z = (Z_1, Z_2, \ldots, Z_{n_Z}) : \Omega \to \mathbb{R}^{n_Z}$ be a n_Z-dimensional random vector with distribution function $F_Z(z) = \text{Prob}\,(Z \leqslant z)$, where Ω is the probability space and $z \in \mathbb{R}^{n_Z}$. Reliability problem is traditionally defined by a (scalar) limit state function, also called performance function, $g(Z)$, where $g(Z) < 0$ defines the failure domain Ω_f and $g(Z) \geqslant 0$ the safe domain. That is,

$$\Omega_f \triangleq \{Z : g(Z) < 0\}. \tag{2.1}$$

And the probability of failure P_f is defined as

$$P_f = \text{Prob}(Z \in \Omega_f) = \int_{\Omega_f} dF_Z(z) = \int \chi_{\Omega_f}(z) dF_Z(z), \tag{2.2}$$

where χ is the characteristic function satisfying

$$\chi_A(z) = \begin{cases} 1 & \text{if} \quad z \in A, \\ 0 & \text{if} \quad z \notin A. \end{cases} \tag{2.3}$$

It is obvious that the hyper-surface defined by $g(Z) = 0$, called limit state surface, plays an important role. A proper approach to determine the failure probability will depend on both the geometry of the limit state surface and the dimensionality n_Z.

It should be noted that in most practical cases the form of the limit state function $g(Z)$ is not known explicitly. In fact, very often the random vector Z represents the random inputs into a complex system and $g(Z)$ represents a complicated relation between the inputs and the failure modes via the solution of the system. That is, $g(Z)$ depends on the solution of a potential highly complex stochastic system. Hence, no explicit form of $g(Z)$ is available. And $g(Z)$ can only be evaluated by simulations of the system and can be computational intensive to obtain.

3. Existing approaches

Here we briefly review the existing approaches. In particular, we focus on the Monte Carlo sampling (MCS) and response surface method (RSM), because they are directly related to the method proposed in this paper. The methods of FORM and SORM will not be reviewed here. Instead we refer the interested readers to reviews such as [24,27,26].

3.1. Monte Carlo simulation

The most straightforward way to compute the failure probability is Monte Carlo sampling (MCS), where random samples are generated according to the distribution of Z and those land in the failure region are counted to estimate the failure probability. Let $z^{(i)} \in \mathbb{R}^{n_Z},\ i = 1, \ldots, M$, be a set of samples for the random vector Z, then a MCS estimate of the failure probability is

$$P_f^{mc} = \frac{1}{M} \sum_{i=1}^{M} \chi_{\{g(z)<0\}}(z^{(i)}). \tag{3.1}$$

Note, hereafter we will, with a slight abuse of notation, use the shorthanded notation $\{g(z) < 0\}$ to stand for the set $\{z : g(z) < 0\}$. This is done merely for notational convenience.

Although easy to implement, the MCS approach can be costly in practice, for each sample point requires a full-scale simulation of the underlying system. And usually a large number of samples is required to obtain an accurate estimate of the failure probability. The approach can be much more difficult when the failure probability is small, as in this case most of the samples will not land in the failure region. Variations of MCS exists, see, for example, [3,8,20,21,1,18,2].

3.2. Response surface method

One of the alternatives to MCS is response surface method (RSM). Here one conducts an accurate (stochastic) simulation of the underlying system and constructs a response surface (RS) for the limit state function. That is, we seek

$$\tilde{g}(Z) \approx g(Z),$$

where the approximation is assessed in a proper norm/metric and is satisfactory. We then estimate the failure probability based on the RS, that is,

$$P_f^{RS} = \frac{1}{M} \sum_{i=1}^{M} \chi_{\{\tilde{g}(z)<0\}}(z^{(i)}). \tag{3.2}$$

A distinct advantage of this approach is that sampling $\tilde{g}$ does not require any simulation effort of the original system. Hence, we can use an arbitrarily large number of samples to obtain the estimate. The only simulation effort is spent on constructing the RS and is a one-time cost. Therefore the accuracy of the surrogate is important, and there have been some research efforts devoted to this line of approach [25,12,14,22,23].

We now present counter-examples to demonstrate that such a straightforward use of RS will result in erroneous results for the failure probability estimate. More importantly, the accuracy of the RS can not help improve the estimate. In fact, no matter how accurate the RS is, the resulting failure probability estimate can still be incorrect.

Example 3.1. Consider a trivial state function $g(Z) = 0$. Thus the true failure probability Prob $(g(Z) < 0) = 0$. Now let a surrogate model be

$$g_N(Z) = -\frac{1}{N}, \quad N \geqslant 1,$$

which implies

$$\lim_{N\to\infty} g_N(Z) = g(Z), \quad \forall Z.$$

However, $\widetilde{P}_f = \text{Prob}(g_N < 0) = 1$ for any finite N. Therefore, the estimate $\widetilde{P}_f$ is completely off even though g_N approximates g not only point-wise but also uniformly. Similarly, one can use a response surface

$$g_N(Z) = \frac{1}{N}\sin(kZ),$$

where k is an arbitrary wave number. The same conclusion holds. Though $g_N(Z) \to g(Z)$ as $N \to \infty$, for any finite N, $\widetilde{P}_f = \text{Prob}(g_N < 0) \approx 1/2$ and is an incorrect result.

Example 3.2. Let $Z \sim U[-1,1]$ be a random variable uniformly distributed in $[1,1]$ and consider a state function

$$g(Z) = \begin{cases} -1, & Z \in [-1,0), \\ -0.5, & Z = 0, \\ 0, & Z \in (0,1]. \end{cases}$$

Naturally, $g(Z) \in L^2_{[-1,1]}$, and Prob $(g(Z) < 0) = 0.5$. We now construct an orthogonal polynomial approximation to g. Let $g_N(Z)$ be the series expansion for g using Legendre polynomials. Its explicit form exists

$$g_N(Z) = -\frac{1}{2} + \sum_{n=0}^{N} \frac{(-1)^n(4n+3)(2n)!}{2^{2n+2}(n+1)!n!} P_{2n+1}(Z),$$

where P_n are the classical Legendre polynomials. This is a global polynomial approximation to a discontinuous function and we have the L^2 convergence $\lim_{N\to\infty}\|g - g_N\|_{L^2_{[-1,1]}} = 0$. However, for any finite N, the expansion suffers from Gibbs' oscillation. Consequently, Prob $(g_N(Z) < 0) = C$, where the constant $C \approx 3/4$.

The key message behind these counter-examples is that a response surface can be as accurate as one wishes, measured by one of the classical norms or even point-wise, but the resulting estimate of failure probability can be erroneous and produce catastrophic results in practice.

4. Hybrid approach

We now present a hybrid approach for failure probability computation. The approach seeks to combine the advantage of the robustness of MCS and the ease of sampling of response surface method. By doing so, the new method significantly improves the efficiency of the MCS and addresses the accuracy issue of the direct RSM. Hereafter we will replace the notion of response surface with *surrogate model* because the traditional RSM is mostly based on heuristic constructions such as regression, design of experiments, etc. We use surrogate model to incorporate the more recent rigorous constructions, which include, but are not limited to, those based on generalized polynomial chaos (gPC).

4.1. Algorithm

Let $g_N(Z)$ be a surrogate model for the limit state function $g(Z)$. That is, $g_N(Z) \approx g(Z)$ in a proper norm. The hybrid method consists of the following steps:

1. Constructing an accurate surrogate model g_N for the limit state function g. We require g_N to approximate g in a proper strong norm, typically a L^p-norm with $p \geqslant 1$.
2. Identify/prescribe a threshold value $\gamma > 0$. Usually this is a small positive number.
3. Conduct random sampling of Z by generating $z^{(i)} \in \mathbb{R}^{n_Z}$, $i = 1,\ldots,M$, according to the distribution of Z.
4. For each $i = 1,\ldots,M$, evaluate $g_N(z^{(i)})$. If $|g_N(z^{(i)})| \leqslant \gamma$, evaluate $g(z^{(i)})$.
5. Finally, the failure probability is estimated as

$$P_f^h = \frac{1}{M}\sum_{i=1}^{M} \chi_{\widetilde{\Omega}_f}(z^{(i)}), \tag{4.1}$$

where the approximate failure region is defined as

$$\widetilde{\Omega}_f \triangleq \{g_N(Z) < -\gamma\} \cup \{\{|g_N(Z)| \leqslant \gamma\} \cap \{g(Z) < 0\}\}. \tag{4.2}$$

The key idea behind the algorithm is that we use the samples from the surrogate g_N whenever their values are reasonably large ($|g_N(z)| > \gamma$), and only use the samples of the true limit state g when $|g_N|$ is smaller than the threshold. By doing so, only a (small) portion of the samples is from sampling g (which is costly) and a (large) portion is from sampling the surrogate g_N (which is inexpensive). Therefore, this hybrid approach combines the accuracy of MCS and efficiency of surrogate approaches.

The parameter γ determines the efficiency of the algorithm. The smaller the value of γ, the smaller the region where samples of the original system g are required. Hence, small γ is preferred. However, to ensure the accuracy of the algorithm, γ can not be arbitrarily small. Its value is directly related to the accuracy of the surrogate model g_N. This will be discussed in details in the convergence study of the algorithm in the following sections.

4.2. Construction of surrogate

Construction of the surrogate model is a critical step. And a surrogate of high accuracy is always preferred. Though there exist many means of constructing surrogate models, here we discuss an approach based on the generalized polynomial chaos (gPC) methodology, which has become widely popular in the field of stochastic computing. We emphasize that the hybrid approach presented here requires *only the expression of a surrogate model*. How one constructs the surrogate model is in principle independent of the implementation of the hybrid approach. Therefore, in practice, one is free to use any kind of surrogate model and not restricted to the gPC models discussed here. For example, one can employ an empirical surrogate model. The conditions to ensure the accuracy of the hybrid method is rather mild, and will be presented in the following section.

Let $\mathbf{i} = (i_1, i_2, \ldots, i_{n_Z}) \in \mathbb{N}_0^{n_Z}$ be a multi-index with $|\mathbf{i}| = i_1 + i_2 + \cdots + i_{n_Z}$, and let $N \geqslant 0$ be an integer. The Nth-degree gPC expansion of $g(Z)$ takes the following form

$$g_N(Z) = \sum_{|\mathbf{i}|=0}^{N} a_{\mathbf{i}} \Phi_{\mathbf{i}}(Z). \tag{4.3}$$

where $\{a_{\mathbf{i}}\}$ are expansion coefficients that are to be determined, and $\{\Phi_{\mathbf{i}}(Z)\}$ are n_Z-dimensional orthogonal polynomials of degree up to N, satisfying

$$\mathbb{E}[\Phi_{\mathbf{i}}(Z)\Phi_{\mathbf{j}}(Z)] = \int_\Omega \Phi_{\mathbf{i}}(z)\Phi_{\mathbf{j}}(z) dF_Z(z) = \delta_{\mathbf{ij}}, \quad 0 \leqslant |\mathbf{i}|, \ |\mathbf{j}| \leqslant N. \tag{4.4}$$

Here $\delta_{\mathbf{ij}} = \prod_{d=1}^{n_Z} \delta_{i_d j_d} = 1$ if $\mathbf{i} = \mathbf{j}$ and 0 otherwise, and is the multivariate Kronecker delta function. The orthogonality relation indicates a correspondence between the distribution of Z and the type of the orthogonal polynomial basis $\{\Phi_{\mathbf{i}}(Z)\}$. For example, when dF_Z is a Gaussian measure corresponding to normal distribution, the orthogonality relation (4.4) defines Hermite orthogonal polynomials. Therefore, we have Hermite polynomial chaos corresponds to Gaussian distribution, Legendre polynomial chaos corresponds to uniform distribution, etc. For detailed discussions of these relations, see [34].

For a given stochastic system, there are different ways to obtain the expansion coefficients. This is where different numerical approaches deviate from each other. Roughly speaking, two of the most popular approaches are the Galerkin method and the collocation method. We will not engage in a detailed discussion on this and refer the interested readers to references such as [31,32]. For the purpose of this paper, it suffices to assume that a particular numerical method is employed and the expansion coefficients are obtained, and the resulting gPC solution is a satisfactory approximation to the true solution. The most common norm to measure the accuracy is the L^2 norm, which implies

$$\|g(Z) - g_N(Z)\|_{L^2_\Omega}^2 := \int_\Omega |g(z) - g_N(z)|^2 dF_Z(z) \to 0, \quad N \to \infty. \tag{4.5}$$

And the convergence rate depends on the regularity of g. The smoother g is, the faster the convergence rate. When g is relatively smooth, a low-degree expansion may achieve sufficient accuracy and results in significant efficiency gain.

Note that when the true solution possesses sufficient smoothness in the random space, it is possible for the gPC approximation to converge in a stronger norm such as the L^∞ norm. However, here we will adopt the rather mild and more general L^p, $p \geqslant 1$, norm, which includes the most commonly used L^2 norm. More importantly, hereafter we will not require the convergence property of g_N, as in the form of (4.5). We will only assume that g_N is an approximation of g and its error can be quantified in L^p norm. Though gPC approach is discussed in details here, in practice one can adopt any kind of surrogate with finite accuracy in L^p norm.

4.3. Error analysis

Here we present an error estimate of the hybrid approach for failure probability estimate. Again we remark that only L^p error of the surrogate model g_N is required. For notational convenience, we define, for a real $\gamma \geqslant 0$,

$$\begin{aligned} P_{N,\gamma} &\triangleq \int_\Omega \chi_{\{g_N(Z)<-\gamma\}} dF_Z(z), \\ Q_{N,\gamma} &\triangleq \int_\Omega \chi_{\{-\gamma\leqslant g_N(Z)\leqslant\gamma\}\cap\{g(Z)<0\}} dF_Z(z). \end{aligned} \tag{4.6}$$

From the algorithm in Section 4.1, it is clear that $P_{N,\gamma}$ stands for the failure probability computed from sampling the surrogate g_N "away" from the limit state, and $Q_{N,\gamma}$ represents the failure probability computed from sampling g when g_N is "close" to the limit state. The overall estimate of the failure probability by the hybrid algorithm is the summation of the two, i.e.,

$$P_f^h = P_{N,\gamma} + Q_{N,\gamma}. \tag{4.7}$$

Naturally, P_f^h becomes the surrogate/response surface estimate in (3.2) when γ approaches zero. And when γ approaches infinity, P_f^h becomes the MCS estimate P_f^{mc} in (3.1). The following result holds.

Theorem 4.1. *Assume, in the definition (2.2) of the failure probability P_f, the limit state function $g(Z) \in L^p_\Omega$, $p \geqslant 1$, and $g_N(Z)$ is its approximation in L^p-norm, i.e., for $p \geqslant 1$,*

$$\|g(Z) - g_N(Z)\|_{L^p_\Omega} = \left(\int_\Omega |g(z) - g_N(z)|^p dF_Z(z) \right)^{1/p}.$$

Let $P_{N,\gamma}$ and $Q_{N,\gamma}$ be defined as in (4.6) and P_f^h as in (4.7). Then for all $\epsilon > 0$, there exits $\gamma_N > 0$ such that for all $\gamma > \gamma_N$

$$\left| P_f - P_f^h \right| < \epsilon. \tag{4.8}$$

More precisely,

$$\gamma_N = \frac{1}{\epsilon^{1/p}} \|g(Z) - g_N(Z)\|_{L^p_\Omega}. \tag{4.9}$$

Proof. Since for any $\gamma > 0$,

$$\begin{aligned} \|g(Z) - g_N(Z)\|^p_{L^p_\Omega} &= \int_\Omega |g(z) - g_N(z)|^p dF_Z(z) \\ &= \int_\Omega \chi_{\{|g(z)-g_N(z)|>\gamma\}} |g(z) - g_N(z)|^p dF_Z(z) + \int_\Omega \chi_{\{|g(z)-g_N(z)|\leqslant\gamma\}} |g(z) - g_N(z)|^p dF_Z(z) \\ &\geqslant \gamma^p \int_\Omega \chi_{\{|g(z)-g_N(z)|>\gamma\}} dF_Z(z), \end{aligned} \tag{4.10}$$

it follows that

$$\int_\Omega \chi_{\{|g(z)-g_N(z)|>\gamma\}} dF_Z(z) \leqslant \frac{1}{\gamma^p} \|g(z) - g_N(z)\|^p_{L^p_\Omega}. \tag{4.11}$$

As a result, for any given $\epsilon > 0$, let $\gamma > \gamma_N$, which is defined in (4.9), we have

$$\int_\Omega \chi_{\{|g(z)-g_N(z)|>\gamma\}} dF_Z(z) < \epsilon. \tag{4.12}$$

From the definition (2.2) of P_f, we have

$$\begin{aligned} P_f = \int_\Omega \chi_{\{g(z)<0\}} dF_Z(z) &= \int_\Omega \chi_{\{g_N(z)<-\gamma\}\cap\{g(z)<0\}} dF_Z(z) + \int_\Omega \chi_{\{-\gamma\leqslant g_N(z)\leqslant\gamma\}\cap\{g(z)<0\}} dF_Z(z) \\ &+ \int_\Omega \chi_{\{g_N(z)>\gamma\}\cap\{g(z)<0\}} dF_Z(z) := P_1 + P_2 + P_3, \end{aligned} \tag{4.13}$$

where P_1, P_2, P_3 are the three integrals in the above expressions, respectively.

Since $\{g_N(z) < -\gamma\} = (\{g_N(z) < -\gamma\} \cap \{g(z) < 0\}) \cup (\{g_N(z) < -\gamma\} \cap \{g(z) \geqslant 0\})$, and $(\{g_N(z) < -\gamma\} \cap \{g(z) \geqslant 0\}) \subset \{|g_N(z) - g(z)| > \gamma\}$, we have

$$P_{N,\gamma} < P_1 + \int_\Omega \chi_{\{|g(z)-g_N(z)|>\gamma\}} dF_Z(z). \tag{4.14}$$

On the other hand, $(\{g_N(z) < -\gamma\} \cap \{g(z) < 0\}) \subset \{g_N(z) < -\gamma\}$, then we have

$$P_1 \leqslant P_{N,\gamma}. \tag{4.15}$$

It is easy to see, from the definitions in (4.6), that $P_2 = Q_{N,\gamma}$. For P_3, since $\{g_N(z) > \gamma\} \cap \{g(z) < 0\} \subset \{|g(z) - g_N(z)| > \gamma\}$,

$$0 \leqslant P_3 \leqslant \int_\Omega \chi_{\{|g(z)-g_N(z)|>\gamma\}} dF_Z(z). \tag{4.16}$$

Upon combining Eqs. (4.14)–(4.16) we obtain

$$P_{N,\gamma} - \int_\Omega \chi_{\{|g(z)-g_N(z)|>\gamma\}} dF_Z(z) + Q_{N,\gamma} \leqslant P_1 + P_2 + P_3 \leqslant P_{N,\gamma} + Q_{N,\gamma} + \int_\Omega \chi_{\{|g(z)-g_N(z)|>\gamma\}} dF_Z(z). \tag{4.17}$$

Together with (4.13), (4.12) and (4.7), the main conclusion holds. □

The theorem states that the estimate by the hybrid method (4.7) will converge to the true failure probability P_f for any threshold value γ that is larger than γ_N. Therefore γ_N is a lower bound of the threshold to guarantee convergence of the failure probability estimate. It is also clear from the proof that the γ_N given in (4.9) is an upper bound estimate. Hereafter we will refer γ_N of (4.9) as the *critical threshold*, which is determined by the L^p error of the surrogate model g_N and the prescribed error control ϵ. For a given error control ϵ, the more accurate g_N is, the smaller γ_N is. Hence the smaller the region where samples of the original model g is required and the larger the efficiency gain. It is also worth noting that the convergence

of the failure probability P_f^h does not require the convergence of the surrogate g_N. That is, γ_N from (4.9) is merely a number providing the critical threshold value to separate the two sampling regions. For any surrogate g_N with finite accuracy, any choice of γ larger than γ_N would be sufficient. This fact is useful for practical computation because it guarantees the convergence of the failure probability estimate for any given surrogate model with fixed accuracy.

In practice, the integrals $P_{N,\gamma}$ and $Q_{N,\gamma}$ need to be calculated numerically. The straightforward approach is to use MCS. Let $z^{(i)}$, $i = 1,\ldots,M$, be samples generated by the distribution of Z. Then,

$$\widetilde{P}_{N,\gamma} = \frac{1}{M}\sum_{i=1}^{M} \chi_{\{g_N<-\gamma\}}(z^{(i)}), \tag{4.18}$$

$$\widetilde{Q}_{N,\gamma} = \frac{1}{M}\sum_{i=1}^{M} \chi_{\{|g_N|\leqslant\gamma\}}(z^{(i)}) \cdot \chi_{\{g<0\}}(z^{(i)}), \tag{4.19}$$

are MCS approximations of $P_{N,\gamma}$ and $Q_{N,\gamma}$, respectively. Subsequently, the numerical estimate

$$\widetilde{P}_f^h = \widetilde{P}_{N,\gamma} + \widetilde{Q}_{N,\gamma} \tag{4.20}$$

will converge to P_f^h as the number of samples increases. Note again, for an ensemble of large size $M \gg 1$, most of its samples lie in $\widetilde{P}_{N,\gamma}$, which can be evaluated easily by evaluating the surrogate g_N. Only a small portion of the ensemble lies in the estimate $\widetilde{Q}_{N,\gamma}$, which requires evaluations of the original system g.

4.4. A practical iterative algorithm

The convergence of the hybrid method, as stated in Theorem 4.1, requires the knowledge of γ_N, the critical threshold value, which in turns depends on the L^p error of the surrogate model g_N. In practice, the error of the surrogate is often hard to obtain and the choice of the threshold value γ becomes "free"–not a desirable situation. Here we propose an iterative scheme that avoids the choice of γ.

Let $M \gg 1$ be the total number of samples, and $S_z = \{z^{(i)}\}_{i=1}^M$ be the sample set generated from the probability distribution F_Z. Let δM be an integer (much) smaller than M for the "step size" and $\eta \geqslant 0$ a small number for stopping criterion of the following iteration. Let integer k be the iteration count.

- **Initialization**:
  - Set $k = 0$, $M^{(k)} = 0$, and $S_z^{(k)} = \emptyset$.
  - Estimate the failure probability using the surrogate model g_N. That is, for $k = 0$, let

$$P_f^{(k)} = \frac{1}{M}\sum_{i=1}^{M} \chi_{\{g_N<0\}}(z^{(i)}). \tag{4.21}$$

  - Sort $\{|g_N(z^{(i)})|\}_{i=1}^M$ in ascending order.
- **Iteration**: At k-th iteration ($k > 0$), do the following.
  - Identify the $(M^{(k)} + 1)$ to $(M^{(k)} + \delta M)$ elements in the sorted sequence of $|g_N|$ and their corresponding samples in the set S_z. Denote $\delta S_z^{(k)}$ the set for these samples, and let $S_z^{(k)} = S_z^{(k-1)} \cup \delta S_z^{(k)}$. Obviously the size of $\delta S_z^{(k)}$ is δM.
  - Evaluate the original model g at the sample points in the set $\delta S_z^{(k)}$.
  - Update the failure probability estimate using the values of g on $\delta S_z^{(k)}$.

$$P_f^{(k)} = P_f^{(k-1)} + \frac{1}{M}\sum_{z^{(i)}\in\delta S_z^{(k)}} \left[-\chi_{\{g_N<0\}}(z^{(i)}) + \chi_{\{g<0\}}(z^{(i)})\right]. \tag{4.22}$$

  - If $|P_f^{(k)} - P_f^{(k-1)}| \leqslant \eta$, exit; if not, let $k \leftarrow k + 1$, $M^{(k)} = M^{(k-1)} + \delta M$, and repeat the iteration.

Note the first estimate (4.21) in the initialization step is simply the estimate by the surrogate model. In each iteration step, δM samples whose values of g_N are the "closest" to zero are simulated by the original model g and used in the probability estimation. The iteration will terminate when the stopping criterion is reached, or when the iteration step reaches $\lceil M/\delta M\rceil$. The latter case corresponds to the situation when all of g_N samples are replaced by those of g. And the hybrid estimate P_f^h will become the MCS estimate P_f^{mc}. Therefore, the convergence of the iterative scheme is easily established as

$$P_f^{(k)} \to P_f^{mc}, \quad k \to \lceil M/\delta M\rceil, \tag{4.23}$$

which in turn will converge to the exact P_f with increasing number (M) of samples. However, in practice the iteration would be required to exit early by satisfying the stopping criterion η. This indicates that a numerical convergence of the iteration is reached. In this manner, the iterative algorithm can be considered as a "predictor–corrector" scheme, where the estimate by the surrogate g_N serves as a predictor and the samples close to the limit state are iteratively replaced by their true model g values to correct the probability estimate. The converged failure probability now depends on the accuracy of the underlying surrogate g_N.

We also note that if one has a good error estimate of g_N, then the critical threshold value γ_N can be estimated by (4.9). In this case one can check, at each iteration, if all the samples in $S_z \setminus S_z^{(k)}$ have values of $|g_N|$ larger than γ_N then the iteration can be terminated.

From a practical point of view, one can also prescribe a maximum iteration number $k_{\max}$, which can be specified based on ones simulation capacity. If $M_{\max}$ is the maximum number of Monte Carlo simulations one can afford for the original system g, then $k_{\max} = \lceil M_{\max}/\delta M \rceil$.

4.5. A discussion on the two algorithms

The fundamental idea behind the hybrid method is to separate the integration domain into two regions: one that is "close" to the limit state $g = 0$ where sampling of the original system g is used, and elsewhere in the domain sampling of the surrogate g_N is used. In the above sections, two different algorithms are proposed. The algorithm in Section 4.1 is non-iterative and hereafter will be referred to as the *direct algorithm*, whereas the algorithm in Section 4.4 will be called the *iterative algorithm*. Here we make some brief comments about the algorithms.

- *Direct Algorithm.* The obvious advantage of the direct algorithm from Section 4.1 is its convergence property. The algorithm is guaranteed to be accurate (within the prescribed error control), as stated in Theorem 4.1, for any surrogate model g_N with finite accuracy, so long as one can identify a proper threshold value γ to separate the integration domain. The value of γ should be larger than the critical threshold value γ_N, which depends on the accuracy of g_N as in (4.9). And this is the limitation of the direct algorithm, because in practice one rarely knows the accuracy of the surrogate and can not easily specify the threshold γ.
- *Iterative Algorithm.* The iterative algorithm proposed in Section 4.4 does not require the knowledge of the threshold. Instead the algorithm starts to replace the samples of g_N that are the closest to zero by samples of g and then update the probability estimate. The procedure is conducted iteratively till a convergence is established. Though in principle the result will converge to the MCS result once all the samples of g_N are replaced by those of g. In practice it is desirable to stop the procedure as early as possible by using a convergence criterion parameter, and thus inducing less sampling of g. This implies that the iteration reaches *numerical convergence* for the given surrogate, and the resulting probability estimate now depends on the underlying g_N. Therefore, the accuracy of g_N will directly affect the accuracy of the probability estimate.

In summary, the direct algorithm can always produce an accurate failure probability estimate, as long as one can specify a proper threshold value γ, which may not be easy to do in practice. Any surrogate model g_N can be used, though the less accurate ones will induce larger threshold and render the direct algorithm less efficient. Therefore the accuracy of the surrogate g_N will have a direct impact of the efficiency of the direct algorithm.

The iterative algorithm does not require the specification of the threshold parameter. But its accuracy will depend on the accuracy of the surrogate g_N, and g_N with better accuracy will be preferred. Thus the direct impact of the accuracy of g_N is on the accuracy of the failure probability estimate.

5. Numerical examples

In this section we present some numerical examples to demonstrate the properties of the hybrid approach. The examples are chosen in such a way that they include cases involving random ODE and PDE, scalar equation and system of equations, linear problems and nonlinear problems, as well as uni-variate problems and multivariate problems. The computational complexity of all examples is not particularly high, and this allows us to focus on the properties of the failure probability estimation. Both the direct algorithm and the iterative algorithm are considered. In all numerical examples, we conduct the traditional MCS with $M = 10^6$ samples and use the results as the reference solutions to examine the accuracy of the hybrid methods. Same sequences of $M = 10^6$ random numbers are utilized so that the comparison between MCS and hybrid algorithms are meaningful. In all simulations by the iterative algorithm we employ $\delta M = 100$ and $\eta = 0$.

5.1. Ordinary differential equation

Consider a random ODE

$$\frac{du}{dt} = -Zu, \quad u(0) = u_0, \tag{5.1}$$

where the decay rate coefficient Z is a random variable with probability distribution $F_Z(z) = \text{Prob}\,(Z \leqslant z)$. The exact solution of (5.1) is:

$$u(t, Z) = u_0 e^{-Zt}. \tag{5.2}$$

Suppose the failure probability is $P_f = \text{Prob}[g(u(t,Z)) < 0]$, where the limit state function g is, for a given value u_d,

$$g(u(t, Z)) = u(t, Z) - u_d, \tag{5.3}$$

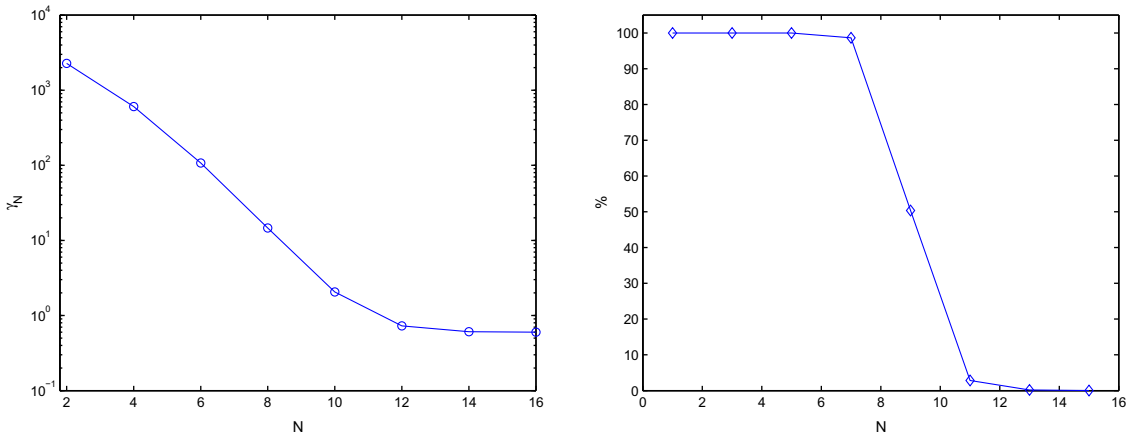

Fig. 5.1. Example 1, direct algorithm. Decay of the critical threshold γ_N (left) and the percentage of the required number of samples of g (right) for different orders of gPC expansion N. (Accuracy requirement is $\epsilon = 10^{-6}$.)

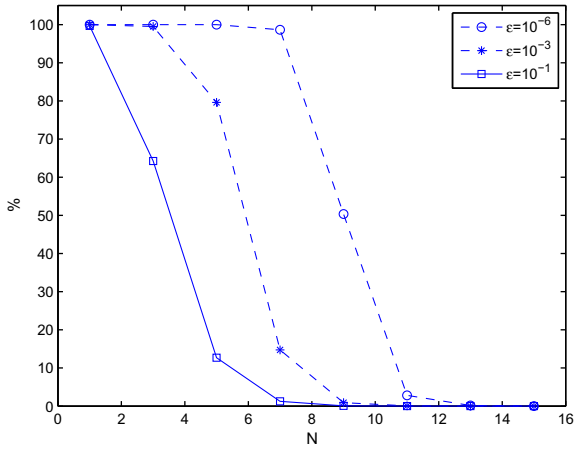

Fig. 5.2. Example 1, direct algorithm. Percentage of the required number of samples of g to reach error control of $\epsilon = 10^{-6}, 10^{-3}, 10^{-1}$, for different orders of gPC expansion N.

Here we employ the gPC Galerkin method to solve (5.1) ([34]). Let u_N be the gPC solution with order N. Then the error can be evaluated as

$$e_N^2(t) = \|u_N(t,Z) - u(t,Z)\|_{L^2_\Omega}^2 = \int_\Omega |u_N(t,z) - u(t,z)|^2 dF_Z(z). \tag{5.4}$$

According to Theorem 4.1, the critical threshold γ_N can be determined by (4.9).

We fix $u_0 = 1$, $t = 1$, and assume $Z \sim \mathcal{N}(\mu, \sigma^2)$ is a Gaussian random variable with mean $\mu = -2$ and standard deviation $\sigma = 1$. Let $u_d = 0.5$ and the exact failure probability can be computed as $P_f = 0.003539050776086$. The reference solution by MCS is $P_f^{mc} = 0.003561$.

We first consider the direct algorithm with an accuracy requirement of $\epsilon = 10^{-6}$. On the left of Fig. 5.1, the critical threshold γ_N from (4.9) is shown for increasing order of gPC expansions. We observe fast (exponential) decay of γ_N with respect to N. This is expected because the decay of γ_N is the same as the L^2 error of g_N, as indicated by (4.9). The fast decay of γ_N indicates that at higher orders, the region of sampling of g becomes smaller very fast. This can be seen on the right of Fig. 5.1, where the percentage of the required number of samples of g to reach the accuracy requirement is shown. The accuracy requirement is set to be very high ($\epsilon = 10^{-6}$). We observe that while at lower orders the required percentage is high (close to 100%), once the surrogate order passes a critical order (around 8), the percentage drops drastically. This indicates that the direct algorithm can produce accurate failure probability estimate with very few samples of the original model g. If we relax the accuracy control, then the number of samples of g will decrease. This is evident from Fig. 5.2, where the percentages of required samples of g for accuracy control of $\epsilon = 10^{-6}, 10^{-3}, 10^{-1}$ are reported. It can be seen that with more relaxed accuracy control, lower order gPC can produce satisfactory failure probability estimate with smaller number of samples of g. We

Table 5.1
Example 1, iterative algorithm. Number of samples (#) of the original model g for convergence and errors of the converged failure probability estimates, for different order (N) of gPC expansions.

N	1	3	5	7	9	11	13	15
#	428,900	28,700	1000	6000	600	300	300	200
Error	5.7(−3)	7.0(−4)	2.5(−3)	4.9(−4)	5.8(−4)	0	0	0

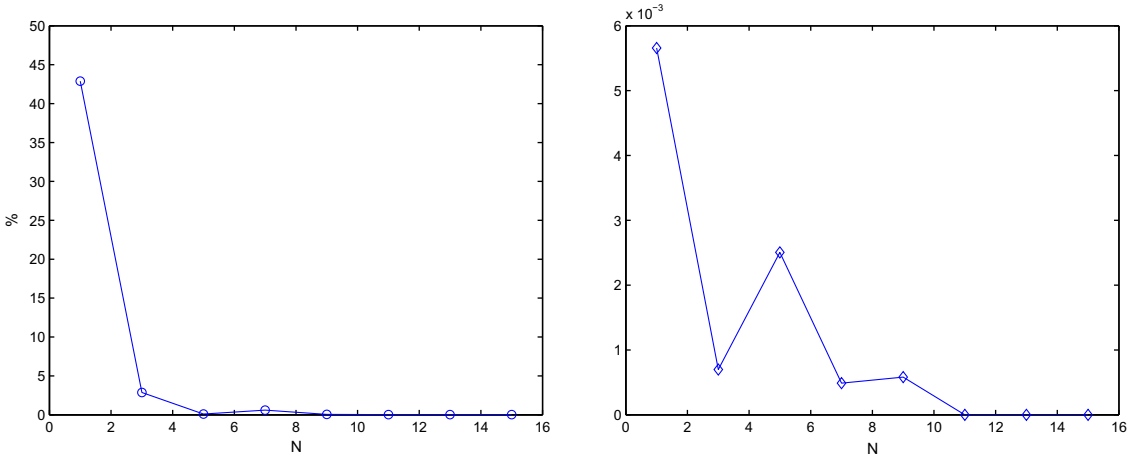

Fig. 5.3. Example 1, iterative algorithm. Percentage of the number of samples of g (left) and errors of the failure probability estimates (right), for different order of gPC expansions.

also remark that these percentages are calculated based on (4.9), which is an upper bound. This implies that the required percentage in actual computation may be (much) lower.

Next we examine the iterative algorithm presented in Section 4.4. In Table 5.1, the numerical results are tabulated for various orders of gPC surrogate g_N. The same results are also graphed in Fig. 5.3. It can be seen that at lower order expansions, more samples of the true model g are required to produce converged solutions, which become more accurate as the order of g_N increases. The percentage of the samples of g drops quickly at higher orders. And at sufficiently high orders, i.e. $N = 11, 13, 15$, a couple of hundreds of samples of g can help produce the exactly the reference MCS solutions. The percentages in these computations are lower than those estimated by the direct algorithm (see Fig. 5.2). This is consistent with the fact that the results of Fig. 5.2 are upper bound estimates.

5.2. Multivariate benchmark

Here we consider a multivariate benchmark problem, the Example 6 from [9]. The failure function was defined as

$$g(X) = X_1 + 2X_2 + 2X_3 + X_4 - 5X_5 - 5X_6 + 0.001 \sum_{i=1}^{6} \sin(100X_i) \tag{5.5}$$

where $X_i \sim LN(\mu_i, \sigma_i)$ for $i = 1, 2, \ldots, 6$, are independent log-normal random variables. Let $\{Z_i\}_{i=1}^6$ be i.i.d. $\mathcal{N}(0,1)$ Gaussian random variables. We employ Hermite polynomials to approximate $\{X_i\}_{i=1}^6$.

$$X_i \approx X_{N,i}(Z_i) = \sum_{k=0}^{N} c_{i,k} H_k(Z_i), \quad c_{i,k} = e^{\mu_i + \frac{\sigma_i^2}{2}} \frac{\sigma_i^k}{k!}. \tag{5.6}$$

We then construct our gPC surrogate model as

$$g_N(Z) = X_{N,1} + 2X_{N,2} + 2X_{N,3} + X_{N,4} - 5X_{N,5} - 5X_{N,6}.$$

A direct approximation of the nonlinear terms involving sine functions is avoided because of their small coefficients. It is easy to see that

$$\|g - g_N\|_{L^2} \leqslant \|X_1 - X_{N,1}\|_{L^2} + 2\|X_2 - X_{N,2}\|_{L^2} + 2\|X_3 - X_{N,3}\|_{L^2} + \|X_4 - X_{N,4}\|_{L^2} + 5\|X_5 - X_{N,5}\|_{L^2} + 5\|X_6 - X_{N,6}\|_{L^2} + 0.006$$

And this relation can be used to estimate the critical threshold γ_N. Note that in this particular example the gPC surrogate g_N *will not converge* to g because of the neglect of the sine terms.

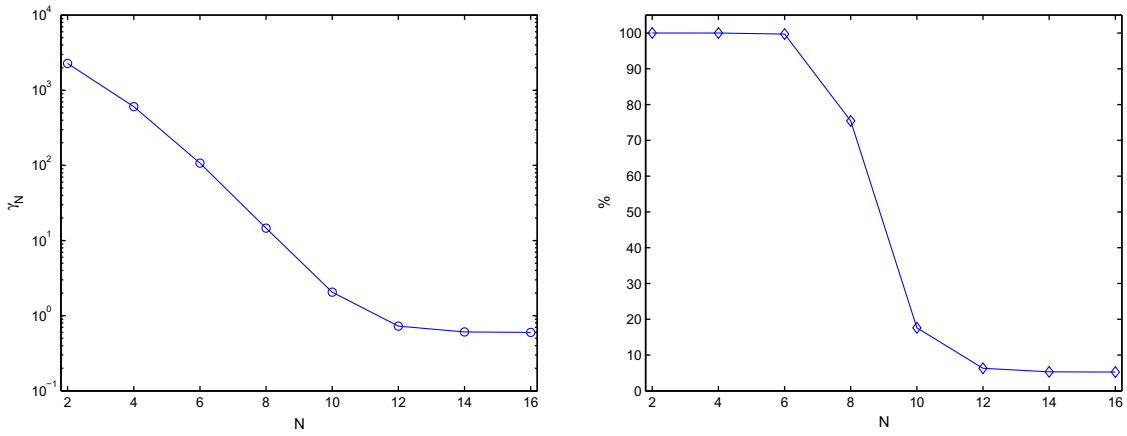

Fig. 5.4. Example 2, direct algorithm. Decay of the critical threshold γ_N (left) and the percentage of the required number of samples of g (right) for different orders of gPC expansion N. (Accuracy requirement is $\epsilon = 10^{-4}$.)

Table 5.2
Example 2, iterative algorithm. Number of samples (#) of the original model g for convergence and error of the converged failure probability estimates, for different order (N) of gPC expansions.

N	2	4	6	8	10	12	14	16
#	178,900	18,600	200	400	600	600	600	600
Error	2.2(−3)	4.2(−3)	4.9(−4)	1.1(−4)	1.0(−6)	0	0	0

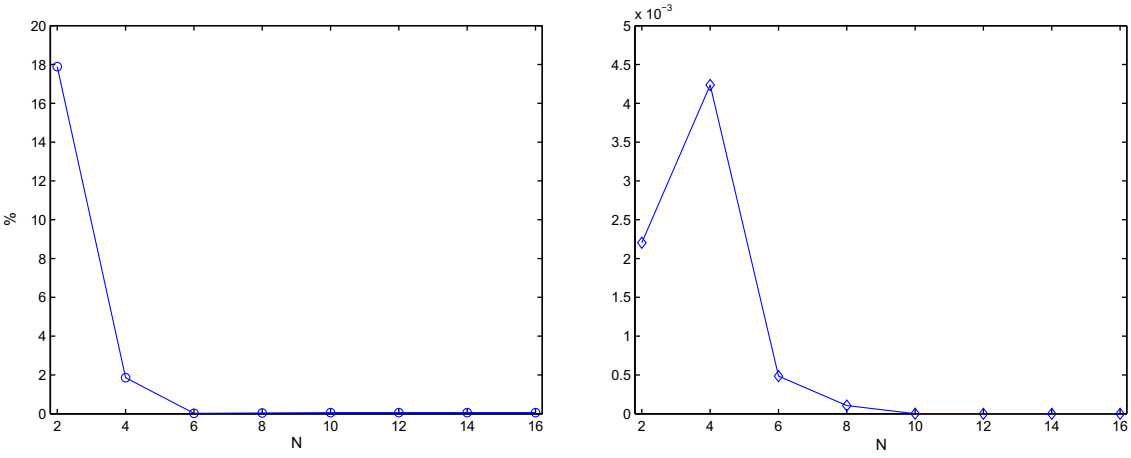

Fig. 5.5. Example 2, iterative algorithm. Percentage of the number of samples of g (left) and error of the failure probability estimates (right), for different order of gPC expansions.

We consider the failure probability $P_f = \text{Prob}\,(g(Z) < 0)$ with $\mu_1 = 0.12$, $\sigma_1 = 1$; $\mu_2 = 0.12$, $\sigma_2 = 1$; $\mu_3 = 0.12$, $\sigma_3 = 1$; $\mu_4 = 0.12$, $\sigma_4 = 0.5$; $\mu_5 = 0.05$, $\sigma_5 = 1$; $\mu_6 = 0.04$, $\sigma_6 = 1$. The reference MCS estimate is $P_f^{mc} = 0.654615$.

For the direct algorithm, we prescribe an error control of $\epsilon = 10^{-4}$. In Fig. 5.4, the critical threshold γ_N is plotted on the left, and the percentage of the number samples of g required to reach the error control is plotted on the right. Again we observe the fast decrease of the critical threshold and notice it will not converge to zero because the surrogate g_N does not converge. However, with sufficient number of samples of g, accurate failure probability estimate can always be obtained. The percentage of samples of g becomes very small when the gPC order is high enough.

Next we examine the iterative algorithm. In Table 5.2 the numerical results are tabulated for various orders of gPC surrogate g_N. And the same results are also graphed in Fig. 5.5. Again we observe the fast decay of errors and required number of samples of g at higher orders. Note that the iterative algorithm can produce exactly the reference MCS results at higher orders, with the help of a few hundreds of samples of g, even though g_N has finite accuracy and will not converge to g. As expected, the percentages of the samples of g are lower than those in Fig. 5.4, which are upper bounds.

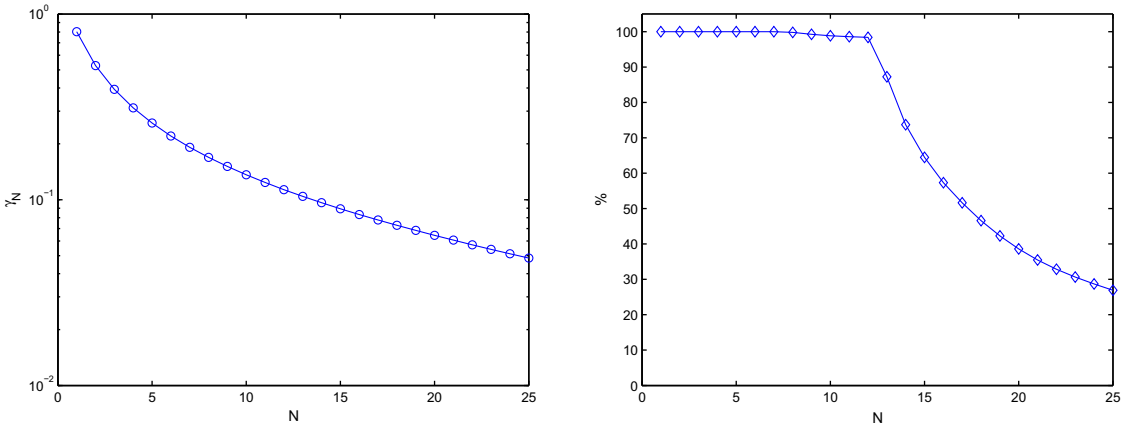

Fig. 5.6. Example 3, direct algorithm. Decay of the critical threshold γ_N (left) and the percentage of the required number of samples of g (right) for different orders of gPC expansion N. (Accuracy requirement is $\epsilon = 10^{-3}$.)

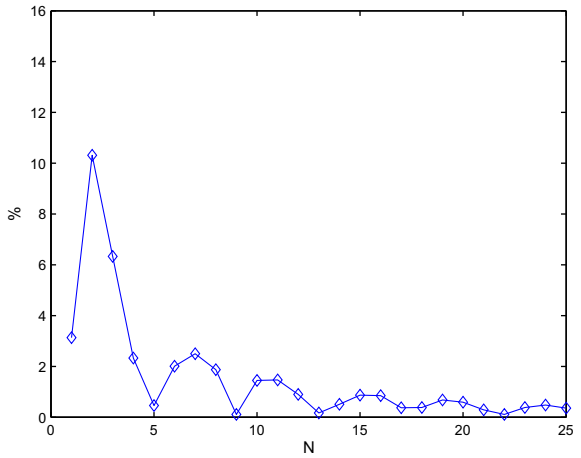

Fig. 5.7. Example 3, iterative algorithm. Percentage of the number of samples of g for different order of gPC expansions. (Note the errors are zero for all orders.)

5.3. Partial differential equation

Consider viscous Burgers' equation

$$\begin{aligned} &u_t + uu_x = vu_{xx}, \quad x \in [-1, 1], \\ &u(-1) = 1 + \delta, \quad u(1) = -1. \end{aligned} \tag{5.7}$$

where δ is a small perturbation to the left boundary condition ($x = -1$) and $v > 0$ is the viscosity. The solution has a transition layer, whose location z, defined as the zero of the solution profile at steady state $u(z) = 0$, can be super sensitive to the uncertainty δ at the left boundary condition. Detailed discussion and simulations of this problem can be found in [35]. Let $\delta \in (0, e)$ be a uniformly distributed random variable with $e \ll 1$, and the failure function be defined as:

$$g(z(\delta)) = -z(\delta) + z_0. \tag{5.8}$$

Note z can be found as the solution of a nonlinear system of equations ([35])

$$A \tanh\left[\frac{A}{2v}(1+z)\right] = 1 + \delta, \quad A \tanh\left[\frac{A}{2v}(1-z)\right] = 1.$$

Let $g_N(\delta) = \sum_{i=0}^{N} a_i \Psi_i(\delta)$ be the gPC approximation of $g(z(\delta))$, where $\Psi_i(\delta)$ are Legendre polynomials and the coefficients are computed via the pseudo spectral stochastic collocation method [30]. The error $\|g(\delta) - g_N(\delta)\|_{L^2}$ can then be computed to estimate γ_N. We evaluate the failure probability when δ has a uniform distribution in $(0, 0.1)$ and $z_0 = 0.75$. The MCS reference solution is $P_f^{mc} = 0.127170$.

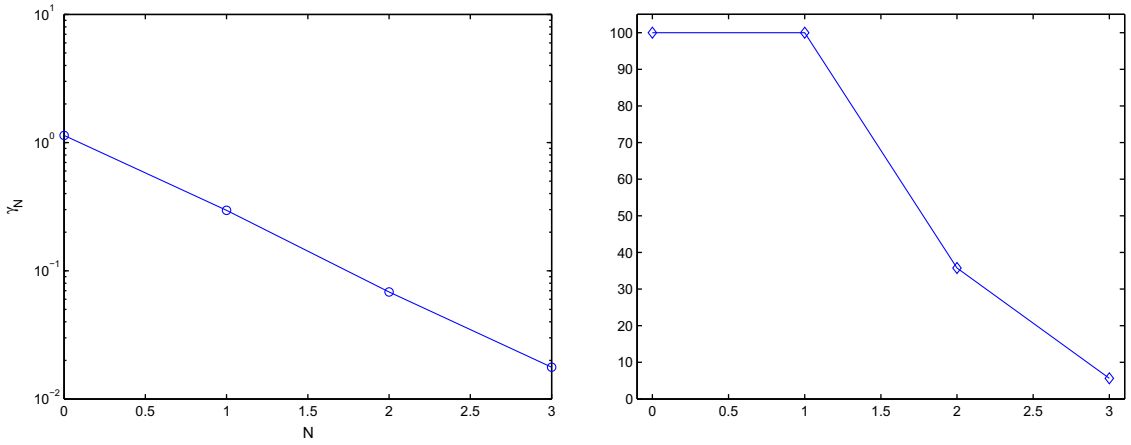

Fig. 5.8. Example 4, direct algorithm. Decay of the critical threshold γ_N (left) and the percentage of the required number of samples of g (right) for different orders of gPC expansion N. (Accuracy requirement is $\epsilon = 10^{-3}$.)

Table 5.3
Example 4, iterative algorithm. Number of samples (#) of the original model g (left) and the error of failure probability estimates (right), for different order (N) of gPC expansions.

N	1	2	3
#	46,400	5100	3600
Error	1.9(−5)	1.6(−5)	6.0(−6)

The results by the direct algorithm are in Fig. 5.6, where the decay of the critical threshold γ_N is shown on the left and the percentage of the number of samples of g required to reach the error control ($\epsilon = 10^{-3}$) is shown on the right. The required number of samples of g is relatively higher than the previous two examples. This is not totally surprising due to the highly nonlinear and supersensitive nature of the problem.

For the iterative algorithm, the numerical results show that, at all orders, the converges results are exactly the same as the reference MCS results. Therefore, the errors are zero for all orders. The percentage of the number of samples of g required to produce such converged (and exact) solutions are plotted in Fig. 5.7 for all orders. Again we observe the fast decay of the number of the true model samples. Note that the percentage of g samples are much lower than that from the direct algorithm estimate in Fig. 5.6. Thus the iterative algorithm is a more practical implementation.

5.4. System of equations: cell-signaling cascade

Finally we consider a mathematical model of autocrine cell-signaling loop described in [28]. Suppose e_{1p}, e_{2p}, and e_{3p} are the dimensionless concentrations of the active form of the enzymes and satisfying the following equations:

$$\begin{aligned}
\frac{de_{1p}}{dt} &= \frac{I(t)}{1+G_4 e_{3p}}\frac{V_{max,1}(1-e_{1p})}{K_{m,1}+1-e_{1p}} - \frac{V_{max,2}e_{1p}}{K_{m,2}+e_{1p}},\\
\frac{de_{2p}}{dt} &= \frac{V_{max,3}e_{1p}(1-e_{2p})}{K_{m,3}+(1-e_{2p})} - \frac{V_{max,4}e_{2p}}{K_{m,4}+e_{2p}},\\
\frac{de_{3p}}{dt} &= \frac{V_{max,5}e_{2p}(1-e_{3p})}{K_{m,5}+(1-e_{3p})} - \frac{V_{max,6}e_{3p}}{K_{m,6}+e_{3p}}.
\end{aligned} \tag{5.9}$$

We choose the same parameters and random variables as in [30], where $K_{m,1-6} = 0.2$, $\langle V\rangle_{max,1} = 0.5$, $\langle V\rangle_{max,2} = 0.15$, $\langle V\rangle_{max,3} = 0.15$, $\langle V\rangle_{max,4} = 0.15$, $\langle V\rangle_{max,5} = 0.25$, and $\langle V\rangle_{max,6} = 0.05$, and for $i = 1,\ldots,6$, $V_{max,i} = \langle V\rangle_{max,i}(1+\sigma\ Z_i)$, where $\sigma = 0.1$, Z_i has a uniform distribution in (-1,1). Here, we use Legendre polynomials as the basis to expand the steady state response of e_{3p} with the fixed input signal $I(t) = 1$ and $G_4 = 0$. Then the random space is six-dimensional, and the sparse grids collocation method ([33]) is used. In this case, we define failure function as:

$$g(Z) = e_{3p}(Z) - e_{3p,0} \tag{5.10}$$

where $e_{3p,0} = 0.8$. The reference MCS solution is $P_f^{mc} = 0.040455$.

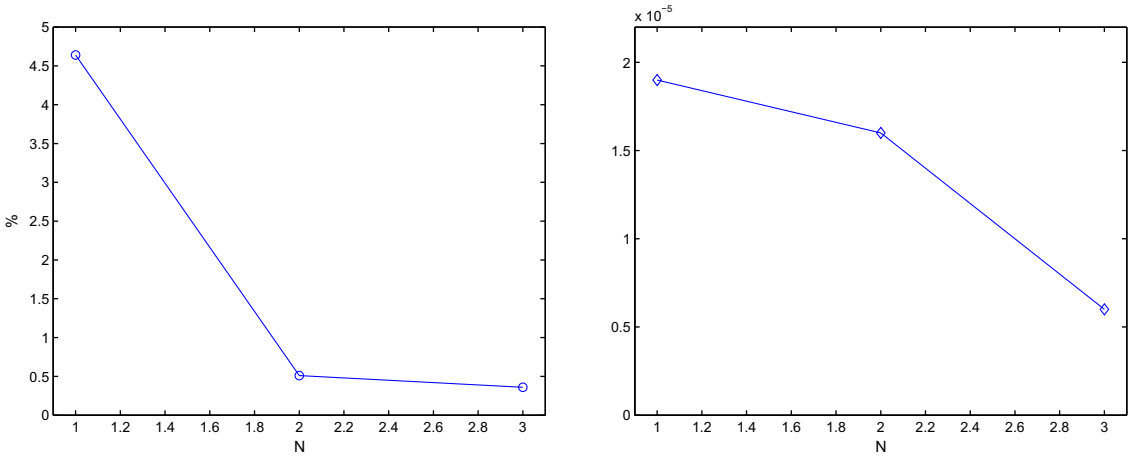

Fig. 5.9. Example 4, iterative algorithm. Percentage of the number of samples of g (left) and the error of failure probability estimates (right), for different order of gPC expansions.

For this high-dimensional problem, we conduct the hybrid algorithm simulations for gPC order up to 3. The results by the direct algorithm are shown in Fig. 5.8, where the error control is set at $\epsilon = 10^{-3}$. Even at modest orders of $N = 2, 3$, the critical threshold γ_N drops to very small values and results in modestly small percentage of samples of g.

For the iterative algorithm, we tabulate and plot the required number of true model sampling of g and the corresponding error in the failure probability estimate in Table 5.3 and Fig. 5.9, respectively. We observe that the iterative algorithm produces highly accurate result even at these low order gPC expansion. The percentages of the samples of g are much smaller than those required by the direct algorithm.

6. Summary

In this paper we identified a fundamental issue in estimating failure probability of a system by the response surface method (RSM). It was shown that the lacking of robustness of the traditional RSM, which seeks to sample the response surface directly, can not be addressed by improving the accuracy of the response surface. Direct sampling of even the most accurate response surface, which is termed surrogate model here, can still lead to erroneous estimate of the failure probability. We then developed a hybrid approach that combines the Monte Carlo sampling (MCS) and the surrogate method. We employed high order generalized polynomial chaos (gPC) expansion as an accurate surrogate, though this certainly is not the only choice. The failure probability is then estimated by conducting a large number of samples in the probability space, where a large portion of the samples is drawn from the surrogate (gPC) model and only a small portion from the original system. The hybrid method is then much efficient than the traditional MCS method and much more accurate than the RSM method. Convergence proof was established and different implementation algorithms were presented in details. Various numerical examples were presented to closely examine the performance and demonstrate the efficiency of the hybrid algorithms.

Acknowledgment

This work is in part supported by AFOSR, DOE/NNSA, NSF.

References

[1] S.K. Au, J. Beck, A new adaptive importance sampling scheme for reliability calculations, Struct. Safety 21 (2) (1999) 135–158.
[2] S.K. Au, J. Beck, Estimation of small failure probabilities in high dimensions by subset simulation, Prob. Eng. Mech. 16 (2001) 263–277.
[3] P. Bjerager, Probability integration by directional simulation, J. Eng. Mech. 114 (1988) 1285–1302.
[4] C.G. Bucher, U. Bourgund, A fast and efficient response surface approach for structural reliability problems, Struct. Safety 7 (1990) 57–66.
[5] G.Q. Cai, I. Elishakoff, Refined second-order reliability analysis, Struct. Safety 14 (1994) 235–246.
[6] A. Der Kiureghian, M. De Stefano, Efficient algorithm for second-order reliability analysis, J. Eng. Mech. 117 (12) (1991) 2904–2923.
[7] A. Der Kiureghian, H.Z. Lin, S.J. Hwang, Second-order reliability approximations, J. Eng. Mech. 113 (8) (1987) 1208–1225.
[8] O. Ditlevsen, P. Bjerager, Methods of structural systems reliability, Struct. Safety 3 (1986) 195–229.
[9] S. Engelund, R. Rackwitz, A benchmark study on importance sampling techniques in structural reliability, Struct. Safety 12 (1993) 255–276.
[10] L. Faravelli, Response surface approach for reliability analysis, J. Eng. Mech. 115 (12) (1989) 2763–2781.
[11] B. Fiessler, H.-J. Neumann, R. Rackwitz, Quadratic limit states in structural reliability, J. Eng. Mech. 105 (4) (1979) 661–676.
[12] N. Gayton, J.M. Bourinet, M. Lemaire, CQ2RS: a new statistical approach to the response surface method for reliability analysis, Struct. Safety 25 (2003) 99–121.
[13] R.G. Ghanem, P. Spanos, Stochastic Finite Elements: A Spectral Approach, Springer-Verlag, 1991.
[14] S. Gupta, C.S. Manohar, An improved response surface method for the determination of failure probability and importance measures, Struct. Safety 26 (2004) 123–139.

[15] M. Hohenbichler, S. Gollwitzer, W. Kruse, R. Rackwitz, New light on first- and second-order reliability methods, Struct. Safety 4 (1987) 267–284.
[16] M. Hohenbichler, R. Rackwitz, First-order concepts in system reliability, Struct. Safety 1 (1983) 177–188.
[17] M. Hohenbichler, R. Rackwitz, Improvement of second-order reliability estimates by importance sampling, J. Eng. Mech 114 (12) (1988) 2195–2199.
[18] P.S. Koutsourelakis, H.J. Pradlwarter, G.I. Schueller, Reliability of structures in high dimensions part I: algorithms and applications, Prob. Eng. Mech. 19 (2004) 409–417.
[19] H.U. Koyluoglu, S.R.K. Nielsen, New approximations for SORM integrals, Struct. Safety 13 (1994) 235–246.
[20] R.E. Melchers, Structural system reliability assessment using directional simulation, Struct. Safety 16 (1994) 23–37.
[21] J. Nie, B.R. Ellingwood, Directional methods for structural reliability, Struct. Safety 22 (2000) 233–249.
[22] M. Paffrath, U. Wever, Adapted polynomial chaos expansion for failure detection, J. Comput. Phys. 226 (1) (2007) 263–281.
[23] R. Pulch. Polynomial chaos for the computation of failure probabilities in periodic problems. In J. Roos and L. Costa, editors, Scientific Computing in Electrical Engineering SCEE 2008, 2010.
[24] R. Rackwitz, Reliability analysis – a review and some perspectives, Struct. Safety 23 (2001) 365–395.
[25] M.R. Rajashekhar, B.R. Ellingwood, A new look at the response surface approach for reliability analysis, Struct. Safety 12 (1993) 205–220.
[26] G.I. Schueller, H.J. Pradlwarter, P.S. Koutsourelakis, A critical appraisal of reliability estimation procedures for high dimensions, Prob. Eng. Mech 19 (2004) 463–474.
[27] G.I. Schueller, R. Stix, A critical appraisal of methods to determine failure probabilities, Struct. Safety 4 (1987) 293–309.
[28] S.Y. Shvartsman, M.P. Hagan, A. Yacoub, P. Dent, H.S. Wiley, D.A. Lauffenburger, Autocrine loops with positive feedback enable context-dependent cell signaling, Am. J. Physiol. Cell Physiol. 282 (3) (2002) C545–559.
[29] F.S. Wong, Slope reliability and response surface method, J. Geo. Eng. 111 (1985) 32–53.
[30] D. Xiu, Efficient collocational approach for parametric uncertainty analysis, Comm. Comput. Phys. 2 (2) (2007) 293–309.
[31] D. Xiu, Fast numerical methods for stochastic computations: a review, Comm. Comput. Phys. 5 (2009) 242–272.
[32] D. Xiu, Numerical Methods For Stochastic Computations: A Spectral Method Approach, Princeton University Press, Princeton, New Jersey, 2010.
[33] D. Xiu, J.S. Hesthaven, High-order collocation methods for differential equations with random inputs, SIAM J. Sci. Comput. 27 (3) (2005) 1118–1139.
[34] D. Xiu, G.E. Karniadakis, The Wiener–Askey polynomial chaos for stochastic differential equations, SIAM J. Sci. Comput. 24 (2) (2002) 619–644.
[35] D. Xiu, G.E. Karniadakis, Supersensitivity due to uncertain boundary conditions, Int. J. Numer. Meth. Engng. 61 (12) (2004) 2114–2138.
[36] Y.-G. Zhao, T. Ono, A general procedure for first/second-order reliability method (FORM/SORM), Struct. Safety 21 (1999) 95–112.