

An Efficient Multifidelity Model for Assessing Risk Probabilities in Power Systems under Rare Events

Yijun Xu
Virginia Tech
yijunxu@vt.edu

Mert Korkali
Lawrence Livermore
National Laboratory
korkali1@llnl.gov

Lamine Mili
Virginia Tech
lmili@vt.edu

Xiao Chen
Lawrence Livermore
National Laboratory
chen73@llnl.gov

Abstract

Risk assessment of power system failures induced by low-frequency, high-impact rare events is of paramount importance to power system planners and operators. In this paper, we develop a cost-effective multi-surrogate method based on multifidelity model for assessing risks in probabilistic power-flow analysis under rare events. Specifically, multiple polynomial-chaos-expansion-based surrogate models are constructed to reproduce power system responses to the stochastic changes of the load and the random occurrence of component outages. These surrogates then propagate a large number of samples at negligible computation cost and thus efficiently screen out the samples associated with high-risk rare events. The results generated by the surrogates, however, may be biased for the samples located in the low-probability tail regions that are critical to power system risk assessment. To resolve this issue, the original high-fidelity power system model is adopted to fine-tune the estimation results of low-fidelity surrogates by reevaluating only a small portion of the samples. This multifidelity model approach greatly improves the computational efficiency of the traditional Monte Carlo method used in computing the risk-event probabilities under rare events without sacrificing computational accuracy.

1. Introduction

Faced with many independent and correlated system conditions, variables, parameters, and events, evaluating the risk of a bulk power system to extreme events is prohibitively involved and very complex for large-scale systems. A particularly daunting challenge in this process is to select from a sufficiently large set of low-probability extreme contingencies since it entails the analysis of an impossibly large number of fault samples in a power system subject to uncertain stochastic loading conditions. The other challenge is to find an efficient way to quantify the risk brought

by these rare events and to enhance the resilience of the system. Hence, novel algorithmic capabilities and analysis tools are needed to arrive at credible and timely risk-informed decisions under various uncertainties and threats in complex power grids.

Motivated by these challenges, some researchers have attempted to simulate cascading outages in power grids to better understand, predict, mitigate and restore the power systems [1, 2] while others have developed probabilistic methods to account for the uncertainties in power system operation and planning [3–9]. Among the proposed methods, those based on probabilistic power-flow (PPF) analysis are able to propagate the uncertainties of system inputs through a nonlinear ac power-flow solver to obtain the probability density functions (pdfs) of the output variables, e.g., power flow and voltage magnitude. The obtained pdfs can provide a full statistical description of the quantity of interest (QoI). Even though many methods focus on deriving information from the first two moments of the QoI [10, 11], the probabilities of the risk events that are located at the tail regions of the pdfs are of great interest to system planners and operators [3, 12–14]. This motivates us to conduct the risk-probability assessment for these rare events in PPF analysis. To solve this problem, Monte Carlo (MC) simulations are typically implemented because of their high accuracy and implementation flexibility [15]. However, the impediment arises from the prohibitively high computational burden. It turns out that in practice, tens of thousands of MC simulations are required to achieve a crude estimation of the pdfs, not to mention to quantify the risks of power system failures subject to low-frequency, high-impact rare events. For example, for a failure probability of 10^{-4} , it is not uncommon to use 10^5 (or even 10^6) as the number of samples for a desired accuracy [13, 16, 17], processing of which is too time-consuming for realistic power system applications. Even by drawing only 10^4 samples, there is no guarantee that the MC method finds rare events with large impact, not to mention accurate estimation of their occurrence

probability. Even with the use of variance-reduction techniques (e.g., importance sampling), which were proposed to reduce the sample size based on the biased prior pdfs, the computing time still necessitates significant reduction [13, 15, 18]. Therefore, a few alternative methods (e.g., the Cornish-Fisher, Edgeworth, and Gram-Charlier methods) to further improve the computational efficiency have been proposed. However, the accuracy of these approaches cannot be guaranteed at either tail regions or under heavily loaded conditions, which makes them unreliable [3, 12, 14].

To overcome the shortcomings of the aforementioned methods, a response-surface (a.k.a. *surrogate* or *reduced-order*) model based on polynomial chaos expansion (PCE) has been advocated by [3, 19–21]. This PCE-based surrogate model can very closely capture the behavior of the complicated, high-fidelity simulation model of a power system while being computationally inexpensive to evaluate [22], allowing the efficient propagation of a large amount of samples. However, the straightforward sampling of a surrogate model can lead to biased results in the low-probability tail regions that are essential for the assessment of rare failure events [16, 17, 23]. Furthermore, the construction of the PCE-based surrogate model relies on the assumption of the smoothness of the original system model equations [24]. However, this assumption is violated when the system topology changes following random branch outages. Indeed, none of the aforementioned PCE-based methods by itself is able to account for the uncertainties of branch outages [3, 19–21, 25]. These uncertainties are structural uncertainties that have much larger influences on system states than those of nodal power injections so that they should not be neglected in assessing system risks of rare events [26]. Thus far, only a handful of attempts have been made to account for random branch outages in PPF analysis owing to the complex nature of this problem [26, 27]. Finally, the crude pdf estimates obtained from only several thousands of samples do not adequately provide a statistical description of long tails in the pdfs, which are mainly induced by the low-frequency, high-impact samples associated with rare events [26].

To address the aforementioned issues, this paper develops a novel PCE-based hybrid multi-surrogate (HMS) method for the risk assessment of power system subject to rare events in PPF analysis. More specifically, this HMS method consists of two stages:

- The first stage is the “sample propagation” stage. Here, multi-PCE surrogates are first constructed to model power system responses

under a non-contingency condition and each of the single-outage (or “ $N - 1$ ”) contingencies. These low-fidelity surrogates models are used to propagate all the samples associated with “ $N - 0$ ” and “ $N - 1$ ” system states under uncertain stochastic nodal power injections. Besides, a small amount of samples related to higher-order contingencies under uncertain nodal power injections are further propagated through the high-fidelity power system model without using any surrogate.

- The second stage is the “sample reevaluation” stage, which is implemented to overcome the inaccuracy of the low-fidelity surrogates in the tails of the target pdfs. Using the aforementioned multi-surrogate models, the samples associated with the high-risk rare events can be screened out efficiently. Then, the high-fidelity power system model is further adopted to fine-tune the low-fidelity estimation results by reevaluating only a small portion of the samples.

2. Problem Formulation

This section formulates the problem of failure-probability assessment of rare risk events in the PPF analysis considering random branch outages. Let us first formulate the power system forward model as

$$z = f(\mathbf{m}). \quad (1)$$

Here, z denotes the QoI, e.g., voltage magnitude, voltage stability margin, and line flow; $\mathbf{m} = [m_1, m_2, \dots, m_N]$ is a vector of uncertain model parameters described by some distribution functions with finite variance. In our work, the active power and reactive power of the loads are considered to follow the Gaussian distribution and the branch states are modeled in the form of 0-1 binomial distributions [3, 26]. The nonlinear function, $f(\cdot)$, denotes the power system model, which maps the model parameters, \mathbf{m} , to the QoI, z . Since the framework developed in this paper is in the context of PPF analysis, $f(\cdot)$ represents the nonlinear ac power-flow model.

Due to the uncertainties in the model input parameters, the QoI will follow some unknown pdf $q(z)$. With that, the probability for target events can be represented by

$$P_f = \Pr(Z \in \Omega_f) = \int_{\Omega_f} q(z) dz = \int \chi_{\Omega_f}(z) q(z) dz, \quad (2)$$

where χ is the characteristic function satisfying

$$\chi_{\Omega_f}(z) = \begin{cases} 1 & \text{if } z \in \Omega_f, \\ 0 & \text{if } z \notin \Omega_f, \end{cases} \quad (3)$$

and Ω_f is the target domain defined as

$$\Omega_f \triangleq \{Z : g(Z) < 0\}. \quad (4)$$

Here, $g(z)$ is a limit state function which is also called the ‘‘performance function’’ that defines the target domain for failure-probability assessment [16]. More specifically, the domain where $g < 0$ stands for the failure domain, and the domain where $g \geq 0$ stands for the safe domain [16, 17]. Based on (2), the probability of failure can be expressed as

$$P_f = \int \chi_{\{g(z)<0\}} q(z) dz. \quad (5)$$

Now, we have completed the formulation of the failure-probability assessment of a power system. For the rare events considered in this paper, the failure probability is typically very low, with a value of $P_f < 10^{-3}$ or $P_f < 10^{-4}$. This entails high requirements to achieve a good accuracy in the long tail of $q(z)$.

Remark. *In this paper, we choose, as our risk estimate (e.g., probability of occurrence), the failure probability exceeding the operational limit in PPF analysis by following the practice suggested by Leite da Silva and de Castro [13].*

3. Algorithm Preliminaries

This section will first introduce the relationship between the MC and the response-surface methods. Then, a cost-effective way to construct a response surface of the power system model via the PCE is illustrated.

3.1. Relationship Between MC Simulations and the Response Surface Method

The most straightforward way to conduct a risk assessment is through MC simulations, where a set of N_m samples are drawn from the multivariate probability distribution of \mathbf{m} , yielding $\{\mathbf{m}^{(j)}\}_{j=1}^{N_m}$. Then, for each $\mathbf{m}^{(j)}$, $j = 1, \dots, N_m$, the QoI is realized through $z^j = f(\mathbf{m}^{(j)})$, yielding a set of $\{z^{(j)}\}_{j=1}^{N_m}$. The probability of failure is calculated as

$$P_f^{\text{MC}} = \sum_{j=1}^{N_m} \frac{1}{N_m} \chi_{\{g(z)<0\}}(z^j). \quad (6)$$

Despite its easy implementation, the computational burden for executing the complicated solver $f(\cdot)$ at N_m parameter values is too heavy for practical applications. This motivates us to use an accurate surrogate model, $\tilde{f}(\mathbf{m})$, that holds a relationship with the original complicated power system model, $f(\mathbf{m})$, as $\tilde{f}(\mathbf{m}) \approx f(\mathbf{m})$. By this way, the uncertainties of the system model can be propagated through the surrogate model, $\tilde{f}(\mathbf{m})$, at little or no extra computational cost. The limit state function for the response surface model described as $\tilde{g}(z)$ also holds for $\tilde{g}(Z) \approx g(Z)$. This enables us to estimate the probability of rare risk events in PPF analysis using the response surface method via

$$P_f^{\text{RS}} = \sum_{j=1}^{N_m} \frac{1}{N_m} \chi_{\{\tilde{g}(z)<0\}}(z^j). \quad (7)$$

By this way, the risk assessment in PPF analysis can be conducted in the surrogate model efficiently. Next, we will present the basic knowledge of the PCE-based surrogate model

3.2. Review of the PCE-based Response Surface

Introduced by Wiener and further developed by Xiu and Karniadakis [22, 24], the generalized polynomial chaos expansion has been shown to be a cost-effective tool in modeling response surfaces [3, 19–21]. In this method, the stochastic outputs are represented as a weighted sum of a given set of orthogonal polynomial chaos basis functions constructed from the probability distribution of the input random variables. Let $\boldsymbol{\xi} = [\xi_1, \xi_2, \dots, \xi_N]$ be a vector of random variables following a standard probability distribution (e.g., the Gaussian or beta distribution), to which, as shown in Table 1 [24], a unique orthogonal polynomial is associated.

Table 1. Univariate gPC Polynomial Bases

Random Variable	Polynomial Basis Function	Support
Gaussian	Hermite	$(-\infty, +\infty)$
Gamma	Laguerre	$[0, +\infty)$
Beta	Jacobi	$[0, 1]$
Uniform	Legendre	$[-1, 1]$

Let $\Phi_i(\xi_1, \xi_2, \dots, \xi_N)$ denote this procedure’s corresponding polynomial chaos basis and a_i denote the i th polynomial chaos coefficient. Formally, we have

$$z = \sum_{i=0}^{N_P} a_i \Phi_i(\boldsymbol{\xi}), \quad (8)$$

where $N_P = (N + P)!/(N!P!) - 1$; N is the total number of the random variables involved in the gPC; and P is the maximum order of the polynomial chaos basis functions, for which a relatively low number (typically 2) is found to provide output results with enough accuracy [3, 19, 21, 28]. From the polynomial chaos coefficients, the mean, μ , and the variance, σ^2 , of the output z can be determined as

$$\mu = a_0, \quad (9)$$

$$\sigma^2 = \sum_{i=1}^{N_P} a_i^2 \mathbb{E}[\phi_i^2], \quad (10)$$

where $\mathbb{E}[\cdot]$ is the expectation operator.

3.2.1. The Orthogonal Polynomial Chaos Basis A set of one-dimensional polynomial chaos basis functions $\{\phi_i(\xi), i = 0, 1, 2, 3, \dots\}$ with respect to some real positive measure should satisfy the following relations:

$$\int_{\mathbb{R}} \phi_r(\xi) \phi_s(\xi) d\lambda \begin{cases} = 0 & \text{if } r \neq s, \\ > 0 & \text{if } r = s. \end{cases} \quad (11)$$

Here, λ is a probability measure defined as the cumulative distribution function (cdf) of ξ . For every cdf, the associated orthogonal polynomials are unique.

Similarly, any set of multi-dimensional polynomial chaos basis functions, $\{\phi_i(\boldsymbol{\xi}), i = 1, 2, 3, \dots\}$, is orthogonal to each other with respect to their joint probability measure.

3.2.2. Three-Term Recurrence Relation Although several representative probability distributions can be easily addressed by the presented gPC framework using their corresponding polynomial chaos bases function (see Table 1 for example), it is difficult to handle all types of probability distributions. For example, the well-known Weibull distribution used for wind power modeling is missing. To deal with that, we propose to adopt the Stieltjes procedure [29] that is able to handle arbitrary distributions. Its main idea is to generate polynomial chaos bases of an arbitrary pdf through the three-term recurrence relationship of the orthogonal polynomials so that they satisfy (11).

The orthogonal polynomials satisfy a three-term recurrence relation given by

$$\begin{aligned} \phi_{k+1}(\xi) &= (\xi - \alpha_k) \phi_k(\xi) - \beta_k \phi_{k-1}(\xi), \\ \phi_{-1} &= 0, \quad \phi_0 = 1; \\ k &= 0, 1, 2, \dots, K, \end{aligned} \quad (12)$$

where $\phi_k(\xi)$ is a set of orthogonal polynomials defined as

$$\phi_k(\xi) = \xi^k + \text{lower-degree terms}, \quad k = 0, 1, \dots, K, \quad (13)$$

and α_k and β_k are the coefficients of the orthogonal polynomials of the k th order, which are uniquely determined by a probability measure.

3.2.3. The Stieltjes Procedure Several methods exist in the literature to calculate the coefficients α_k and β_k of an orthogonal polynomial chaos basis for an arbitrary probability measure. In this paper, the Stieltjes procedure is chosen as an accurate and a cost-effective method [21, 30]. It is given by

$$\alpha_k = \frac{\int_{\mathbb{R}} \xi \phi_k^2(\xi) d\lambda(\xi)}{\int_{\mathbb{R}} \phi_k^2(\xi) d\lambda(\xi)}, \quad k = 0, 1, 2, \dots, K, \quad (14)$$

$$\beta_k = \frac{\int_{\mathbb{R}} \phi_k^2(\xi) d\lambda(\xi)}{\int_{\mathbb{R}} \phi_{k-1}^2(\xi) d\lambda(\xi)}, \quad k = 1, 2, \dots, K. \quad (15)$$

Here, β_0 is arbitrary and can be conveniently chosen as $\beta_0 = \int_{\mathbb{R}} d\lambda(\xi)$ and K is the highest order of the polynomials. If the measure consists of n discrete points, the integrals in (14) and (15) become summations.

3.2.4. Construction of the Polynomial Chaos Basis A set of multi-dimensional polynomial chaos basis functions can be constructed as the tensor product of the one-dimensional polynomial chaos basis associated with each input random variable. Formally, we have

$$\phi(\boldsymbol{\xi}) = \phi(\xi_1) \otimes \phi(\xi_2) \otimes \dots \otimes \phi(\xi_N), \quad (16)$$

where $\Phi(\xi_i)$ denotes the one-dimensional polynomial chaos basis for the i th random variable.

3.2.5. Collocation Points They can be regarded as a finite sample of $\boldsymbol{\xi} = [\xi_1, \xi_2, \dots, \xi_N]$ that are chosen to approximate the polynomial chaos coefficients. The elements of the collocation points are generated by using the union of the zeros and the roots of one higher-order, one-dimensional polynomial for every random variable [3, 24]. For example, for a 2nd-order Hermite polynomial, its one higher-order polynomial is $\phi_3(\xi) = \xi^3 - 3\xi$. The elements of the collocation points are $\{\sqrt{3}, -\sqrt{3}, 0\}$. With these 3 collocation point elements, if there are N random variables, the number

of possible combinations is 3^N . Since there are $N_P + 1$ unknown coefficients, at least $N_P + 1$ independent combinations should be chosen randomly from the 3^N possible ones [3]. Using the detailed procedure described in [3], the polynomial chaos coefficients can be approximated at selected collocation points.

4. The Proposed Hybrid Multi-Surrogate (HMS) Method

This section will introduce the proposed HMS method with two hybrid procedures involved. The hybrid PCE-based method is first introduced to improve the sampling accuracy along the tails of the QoI. Then, a hybrid multi-PCE model is further developed to handle the topology uncertainties.

4.1. Hybrid PCE-based Method

Motivated by the fact that the direct surrogate-based MC simulations may introduce a significant error in the small-probability tail regions of the pdfs, Li and Xiu [16] first developed the hybrid PCE method, which is shown to be a cost-effective tool in handling small failure probabilities [17, 23]. The main idea is to combine the sampling of the surrogate and original system models. For most of samples, which are located “away” from the limit state for the QoI, the samples of the surrogate models are used. For the samples located “close” to the limit state of the QoIs, the samples are reevaluated through the original system model. By doing so, only a small portion of the samples are obtained from the original system model; thus, the overall cost is much cheaper than the one obtained by sampling using the power system model. Furthermore, this reevaluation stage, also known as a two-stage MC method [23], prevents loss of accuracy coming from the direct usage of the surrogate-based method. Therefore, based on (6) and (7), the risk of rare events via the hybrid approach can be obtained from

$$P_f^H = \sum_{j=1}^{N_m} \frac{1}{N_m} \chi_{\tilde{\Omega}_f}(z^j), \quad (17)$$

where the approximate target domain is defined as

$$\tilde{\Omega}_f \triangleq \{\tilde{g}(Z) < -\gamma\} \cup \{|\tilde{g}(Z)| \leq \gamma\} \cap \{g(Z) < 0\}. \quad (18)$$

Here, γ denotes a threshold, typically set to a small positive number. It is worth noting that γ determines the efficiency of this algorithm. A larger value of γ leads to more reevaluations of the original model, yet with a high accuracy. If $\gamma = 0$, the two-stage MC

method is equivalent to a one-stage, direct surrogate method. Choosing a proper value of γ can enable the results of this two-stage MC method to converge to those of the MC simulations based on the original model [16]. Therefore, the hybrid approach enjoys both the accuracy of the MC method and the efficiency of the surrogate-based method.

4.2. Description of the HMS Method

Although we have obtained the accurate surrogate model with the aforementioned hybrid PCE-based surrogate model, we have not yet been able to fully handle the topological uncertainties in the PPF analysis.

First, it is well known that the PCE-based surrogate model can be inadequate for problems involving model discontinuities [30, 31]. This is because the construction of the surrogate model is based on the smoothness assumption of the system response. However, when the topology of a power system changes due to random branch outages, system responses tend to have abrupt changes that violate this assumption. This is especially true for the rare events involving high-order contingencies. Facing this challenge and motivated by the multi-element generalized polynomial chaos (MEgPC) method that has been widely used in the uncertainty quantification problem comprising model discontinuities or long-term dynamic simulations [7, 30, 31]; instead, we propose to use a multi-surrogate method as explained next. Similar to the main idea underlying the MEgPC method that decomposes the random space of single element to construct multiple new elements, we treat every system topology as a single element; therefore, this results in multiple elements associated with different system topologies. Within every single element, we can construct its corresponding surrogate model that only considers the nodal power-injection uncertainties.

However, due to the complexity of the power system model, simply adopting the idea of “multi-element” will not be sufficient. The number of possible topologies caused by random branch outages is directly proportional to the number of system branches. What is worse, even for a very-small-scale power system, modeling all possible combinations of these topologies can be computationally infeasible. Assuming that the branch states are modeled to follow the pdfs of 0-1 binomial distributions and the number of transmission lines is denoted by N_{branch} , the total number of the combinations becomes $2^{N_{\text{branch}}}$. For example, for a power system with $N_{\text{branch}} = 30$, the total number of the possible topologies reaches $2^{30} \approx 1.07 \times 10^9$. However, construction of that many surrogate

models for these different topologies goes far beyond the existing computational capabilities, thus rendering this method highly impractical.

To overcome this problem, we further consider a hybrid approach to handle the numerous topologies with a much higher computational efficiency. Following the power system planning tradition, we classify different system topologies into $N - 1$, $N - 2$, $N - 3$, etc. criteria. For the original system model without any branch outages and for the $N - 1$ criterion, we maintain the idea of “multi-element” and construct $N_{\text{branch}} + 1$ surrogate models for them separately. For the topologies corresponding to $N - 2$ and higher-order contingencies, we retain the traditional MC method without constructing any surrogate model.

This is because we need to consider the balance between the construction of the surrogate model and the usage of the direct MC method. As is shown in Section 3.2, constructing the PCE-based surrogate model depends on the realizations at a small number of collocation points. That is nearly all the computational cost associated with using the surrogate model. Once the surrogate model is constructed, the realizations of a large number of samples through this surrogate model can be computationally negligible. However, if the number of samples that need to be propagated through the surrogate model is very small, e.g., 2, 1, or even 0, then the construction of the surrogate model itself already becomes computationally more expensive than directly using the original system model. In these cases, using direct MC simulations is more cost-effective.

Even though system failures may cascade in numerous ways, it is well known that the probability of occurrence of every topology can vary dramatically. Apart from the original system model without branch outages, the contingencies of the $N - 1$ criterion occur more frequently than those of the higher-order criteria, e.g., the $N - 2$ and the $N - 3$ criteria, as shown in Table 2 taken from [32].

Table 2. An IEEE Survey of Total Number of North America Overhead Transmission Outages at 230 kV and above (1965-1985)

Type	N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8
Count	10143	951	143	36	8	2	4	2

Furthermore, for all the contingencies of the $N - 1$ criterion, we only need to construct N_{branch} surrogate models. For the risk assessment aiming at rare events associated with a very low probability, e.g., 10^{-4} , it is common to choose a sample size

of at least 10^6 samples for achieving the desired accuracy [13, 16, 17]. If, for instance, the probability of each $N - 1$ topology equals approximately 0.01%, this implicates a 100-sample realization for such topologies. Constructing a PCE-based surrogate model still becomes computationally far more efficient than the direct MC sampling with original system model. However, when the same logic is applied to the $N - 2$ criterion or higher-order contingencies, the number of surrogates increases combinatorially (e.g., $\binom{N_{\text{branch}}}{2}$ and $\binom{N_{\text{branch}}}{3}$ for the $N - 2$ and $N - 3$ criteria, respectively), contributing to diminishing returns in computational efficiency, or, even worse, making this model less efficient than the direct MC sampling.

The above discussions prompt us to merge the aforementioned hybrid PCE method into the framework of the HMS method. The associated framework is depicted in Fig. 1. Note that the computational efficiency of the HMS method can be further improved if all the blocks highlighted in orange are processed with parallel computing. The detailed procedure for implementing the HMS algorithm is illustrated in Algorithm 1.

5. Simulation Results

Using the framework we have established in the preceding section, we perform extensive studies on the IEEE 24-bus Reliability Test System (RTS) whose

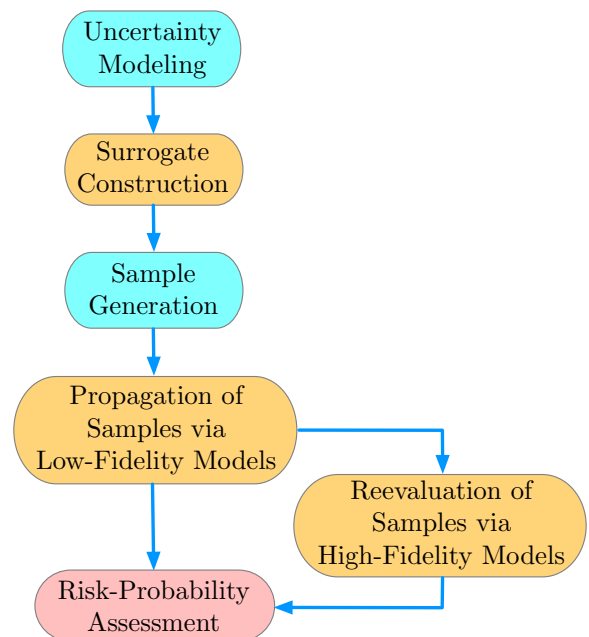


Figure 1. Flowchart of risk-probability assessment via the HMS method.

Algorithm 1 Power System Risk-Probability Assessment via the HMS Algorithm

- 1: Model the uncertainties with their associated pdfs, including those of nodal power injections and branch states; choose the QoI for failure-probability assessment;
 - 2: Considering only the uncertainties of nodal power injections, construct the PCE-based surrogate model of the system without branch outage;
 - 3: **for** $k = 1, \dots, N_{\text{branch}}$ **do**
 - 4: Set Branch k in an “off” state and all the other branches in “on” states; then, construct a PCE-based surrogate model by only considering the nodal power-injection uncertainties;
 - 5: **end for**
 - 6: Generate N_m samples as $\{\mathbf{m}^{(j)}\}_{j=1}^{N_m}$, including uncertainties for nodal power injections and branch states;
 - 7: **for** $j = 1, \dots, N_m$ **do**
 - 8: Classify $\mathbf{m}^{(j)}$ into different types of contingencies, i.e., $N - 0$, $N - 1$, $N - 2$, etc.
 - 9: **if** $\mathbf{m}^{(j)}$ belongs to the cases of the $N - 0$ or the $N - 1$ criteria **then**
 - 10: Find the corresponding surrogate model for $\mathbf{m}^{(j)}$ and compute z^j via its surrogate model $\tilde{f}(\mathbf{m}^{(j)})$;
 - 11: **if** z^j located in target domain $\tilde{\Omega}_f$ **then**
 - 12: Mark sample $\mathbf{m}^{(j)}$ for reevaluation;
 - 13: **end if**
 - 14: **else**
 - 15: Compute z^j via original system model $f(\mathbf{m}^{(j)})$;
 - 16: **end if**
 - 17: **end for**
 - 18: Reevaluate all the samples marked in Step 12 via $f(\mathbf{m}^{(j)})$;
 - 19: Perform risk-probability assessment via (17)
-

network and reliability data are extracted from [33]. The algorithms are tested with MATPOWER package using MATLAB® R2018a version on a laptop with 2.60-GHz Intel® Core™ i7-6600U processors and a 16 GB of main memory.

For this test system, we build a setup to capture the low-probability risk events resulting from topology changes and the variations of the active and reactive power of the loads. Here, it is assumed that the loads follow a Gaussian distribution with mean values equal to the original bus loads and standard deviations equal to 15% of their means. The probabilities of

transmission-line outages are obtained via

$$P_\ell = \frac{\alpha \lambda_p t}{8,760}, \quad (19)$$

where λ_p is the permanent outage rate [outages/yr]; t is the permanent outage duration [h]; and α is the scaling factor [26]. The data for λ_p , t , and α are provided in [33].

5.1. Case Study 1

In this case study, the proposed HMS method is verified in the risk-probability assessment simulations under different scaling factors. The simulation results of the HMS method are compared with the MC method with 1,000,000 samples of power-flow cases. Under different scaling factors, the samples are associated with different numbers of contingencies. Here, we increase α from 1 to 3 to 5, in order to generate samples provided in Table 3.

Table 3. Sample Sizes for Random Branch Outages

Group A ($\alpha = 1$)						
Type	N-0	N-1	N-2	N-3	N-4	N-5
Count	974948	24743	308	1	0	0
Group B ($\alpha = 3$)						
Type	N-0	N-1	N-2	N-3	N-4	N-5
Count	926468	70838	2645	47	2	0
Group C ($\alpha = 5$)						
Type	N-0	N-1	N-2	N-3	N-4	N-5
Count	879625	113230	6868	265	11	1

Let us also denote the number of samples for reevaluation as N_{re} . It is easy to verify that the number of higher-order contingencies increases with an increase in the scaling factor. To make a fair comparison, the samples are propagated through both the MC and HMS methods. It should be noted that 2 CPUs are used in parallel to execute the simulation runs with the MC and HMS methods as displayed in the orange blocks shown in Fig. 1. Let us select the voltage magnitude at Bus 3 as the QoI. Now, we seek to assess the probability of the QoI falling below its minimum operating limit, i.e., 0.95 pu. It is easy to infer that as α increases, the tails of the pdfs become thicker, leading to a higher risk probability, P_f^{MC} . However, this probability is still too low, making an accurate assessment with the traditional response-surface-based method computationally demanding. This motivates

Table 4. Validation of the HMS Method with Different Values of P and γ under Different Scaling Factors α

Group A Data Validation					
P	γ	N_{re}	$P_f^H(\%)$	$P_f^{MC}(\%)$	Time (s)
2	10^{-3}	1479	0.782	0.845	27.3
2	$1.5 \cdot 10^{-3}$	2259	0.826	0.845	31.4
2	$2 \cdot 10^{-3}$	3059	0.843	0.845	34.1
2	$3 \cdot 10^{-3}$	4676	0.844	0.845	40.7
4	10^{-5}	13	0.859	0.845	25.7
4	10^{-4}	161	0.856	0.845	26.2
4	10^{-3}	1643	0.846	0.845	31.7
4	$1.5 \cdot 10^{-3}$	2454	0.845	0.845	34.9
Group B Data Validation					
P	γ	N_{re}	$P_f^H(\%)$	$P_f^{MC}(\%)$	Time (s)
2	10^{-3}	1718	1.25	1.31	43.0
2	$1.5 \cdot 10^{-3}$	2573	1.29	1.31	50.2
2	$2 \cdot 10^{-3}$	3496	1.31	1.31	53.1
2	$3 \cdot 10^{-3}$	5333	1.31	1.31	63.6
4	10^{-5}	17	1.33	1.31	40.5
4	10^{-4}	188	1.32	1.31	41.7
4	10^{-3}	1886	1.32	1.31	48.6
4	$1.5 \cdot 10^{-3}$	2760	1.31	1.31	54.0
Group C Data Validation					
P	γ	N_{re}	$P_f^H(\%)$	$P_f^{MC}(\%)$	Time (s)
2	10^{-3}	2049	1.75	1.82	66.7
2	$1.5 \cdot 10^{-3}$	3076	1.80	1.82	72.9
2	$2 \cdot 10^{-3}$	4092	1.81	1.82	78.1
2	$3 \cdot 10^{-3}$	6121	1.81	1.82	90.2
4	10^{-5}	28	1.84	1.82	56.8
4	10^{-4}	205	1.83	1.82	59.3
4	10^{-3}	2085	1.82	1.82	70.5
4	$1.5 \cdot 10^{-3}$	3162	1.82	1.82	73.5

us to conduct extensive simulations to validate our proposed HMS method using varying PCE orders and tuning parameters, γ , and under varying scaling factors, α , thereby producing the results displayed in Table 4.

The following conclusions can be drawn from these results:

- Compared to the MC method execution that requires approximately 1.5 h to complete all the tests, the HMS method execution requires only about 1 min.

- Compared to the MC results, the HMS method can provide equally accurate simulation results even for risk events with a very low probability, e.g., 10^{-4} or 10^{-5} .
- A larger γ will lead to a larger state space required for reevaluation, and consequently, to an increasing number of reevaluations and a larger computing cost, yet it brings further improvement in simulation accuracy.
- With a higher PCE order P , the simulation results tend to be more accurate under the same γ values. For a relatively low P value, a larger γ value is required to achieve the same level of accuracy.
- As α increases, the computational efficiency gradually decreases. This is partly because there are more samples associated with higher-order contingencies that are propagated through the power system solver $f(\cdot)$. On the other hand, the tails of pdfs become thicker, which potentially leads to more reevaluations for the same target domain.

5.2. Case Study 2

To evaluate the performance of HMS method in a relatively larger power system, the IEEE 118-bus system with 186 branches is selected. Regarding the MC method, we use a sample size of 10^6 to generate a collection of $N - k$ events that fit a negative binomial distribution [34]. Of these, 829,909 are $N - 0$, 154,720 are $N - 1$, 14,452 are $N - 2$, 870 are $N - 3$, and 47 are $N - 4$, and 2 are $N - 5$ event samples, while the samples required to construct multiple surrogates in the HMS method are significantly less. All loads follow a Gaussian distribution with mean values equal to the original bus loads and standard deviations equal to 5% of their means [3]. Since line-outage rates have not been provided for the IEEE 118-bus system in [33], we set the failure rate of each transmission line to 0.1% by multiplying the average failure rate reported in [26] for the IEEE RTS 24-bus system by a scaling factor of $\alpha = 2.5$ so as to increase the complexity of the assessment derived from the increased likelihood of outage occurrence. In this case, the QoI is chosen to be the apparent power of Line 117 connecting Buses 74 and 75 [13]. We validate the accuracy and computational efficiency of the HMS method in dealing with different limit settings for the QoI, and the results in Table 5 are obtained.

These results lead to the following conclusions:

- For the IEEE 118-bus test system, the HMS

method can always complete the simulations within 5 min, much faster than the MC method that spends approximately 3.5 h to complete the simulations.

- The simulation results obtained using the HMS method are very accurate compared to the benchmark results obtained via the MC method. Even if the PCE order is very low, the desired accuracy can be maintained.
- The QoI of the HMS method is not limited to bus

Table 5. Validation of the HMS Method with Different Values of P , γ , and QoI Limit Settings on the IEEE 118-Bus System

Group 1 with limit set to 60 MVA					
P	γ	N_{re}	$P_f^H(\%)$	$P_f^{MC}(\%)$	Time (s)
2	10^{-3}	35	3.02	3.02	254.6
2	10^{-2}	271	3.02	3.02	256.8
2	10^{-1}	2872	3.02	3.02	269.0
3	10^{-3}	26	3.02	3.02	256.1
3	10^{-2}	269	3.02	3.02	258.7
3	10^{-1}	2859	3.02	3.02	282.1
Group 2 with limit set to 65 MVA					
P	γ	N_{re}	$P_f^H(\%)$	$P_f^{MC}(\%)$	Time (s)
2	10^{-3}	16	0.538	0.538	251.9
2	10^{-2}	135	0.538	0.538	254.5
2	10^{-1}	1532	0.538	0.538	265.7
3	10^{-3}	15	0.538	0.538	258.2
3	10^{-2}	136	0.538	0.538	260.7
3	10^{-1}	1537	0.538	0.538	276.9
Group 3 with limit set to 70 MVA					
P	γ	N_{re}	$P_f^H(\%)$	$P_f^{MC}(\%)$	Time (s)
2	10^{-1}	38	0.207	0.207	253.0
2	1	415	0.207	0.207	258.6
3	10^{-1}	35	0.207	0.207	259.6
3	1	414	0.207	0.207	265.8
Group 4 with limit set to 80 MVA					
P	γ	N_{re}	$P_f^H(\%)$	$P_f^{MC}(\%)$	Time (s)
2	10^{-1}	19	0.111	0.111	251.6
2	1	223	0.111	0.111	257.4
3	10^{-1}	19	0.111	0.111	258.7
3	1	224	0.111	0.111	261.9

voltage and can be used for other quantities, such as the power flow on transmission lines.

- Even in a relatively large test system, the usage of the surrogate model is still cost-effective since we construct surrogate models only for high-probability contingency events.

6. Conclusions

In the presence of low-frequency, high-impact rare events in power systems, this paper develops a novel HMS method, which leverages the low-fidelity surrogate model and the high-fidelity power system model, to quantify the likelihood of rare risk events in probabilistic power-flow analysis in a cost-effective manner. Simulation results show that the HMS method with a multifidelity model can accurately capture the low-probability risk events with a much better computational efficiency than the traditional MC method.

Acknowledgment

This work was supported, in part, by the U.S. National Science Foundation under EPAS Grant 1917308 and by the U.S. Department of Energy Office of Electricity Advanced Grid Modeling Program, and performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. Document released as LLNL-CONF-776981.

References

- [1] R. Baldick *et al.*, "Initial review of methods for cascading failure analysis in electric power transmission systems," in *IEEE Power and Energy Soc. Gen. Meeting*, 2008.
- [2] J. Bialek *et al.*, "Benchmarking and validation of cascading failure analysis tools," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4887–4900, Nov. 2016.
- [3] Z. Ren, W. Li, R. Billinton, and W. Yan, "Probabilistic power flow analysis based on the stochastic response surface method," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 2307–2315, May 2016.
- [4] J. D. E. Haesen, C. Bastiaensen and R. Belmans, "A probabilistic formulation of load margins in power systems with stochastic generation," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 951–958, May 2009.
- [5] N. D. Hatziaegyriou and T. S. Karakatsanis, "Probabilistic load flow for assessment of voltage instability," *IEE Proc.-Gener. Transm. Distrib.*, vol. 145, no. 2, pp. 196–202, Mar. 1998.

- [6] X. Li, X. Zhang, L. Wu, P. Lu, and S. Zhang, "Transmission line overload risk assessment for power systems with wind and load-power generation correlation," *IEEE Trans. Smart Grid*, vol. 6, no. 3, pp. 1233–1242, May 2015.
- [7] Y. Xu, L. Mili, A. Sandu, M. R. v. Spakovsky, and J. Zhao, "Propagating uncertainty in power system dynamic simulations using polynomial chaos," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 338–348, Jan. 2019.
- [8] Y. Xu *et al.*, "Response-surface-based Bayesian inference for power system dynamic parameter estimation," *IEEE Trans. Smart Grid*, 2019. [Online]. Available: <https://doi.org/10.1109/TSG.2019.2892464>
- [9] Y. Xu, X. Chen, L. Mili, C. Huang, and M. Korkali, "Polynomial-chaos-based decentralized dynamic parameter estimation using Langevin MCMC," *IEEE Power and Energy Soc. Gen. Meeting*, 2019.
- [10] M. Aien, M. Fotuhi-Firuzabad, and F. Aminifar, "Probabilistic load flow in correlated uncertain environment using unscented transformation," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2233–2241, Nov. 2012.
- [11] S. Peng, J. Tang, and W. Li, "Probabilistic power flow for AC/VSC-MTDC hybrid grids considering rank correlation among diverse uncertainty sources," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 4035–4044, Sep. 2017.
- [12] M. Fan, V. Vittal, G. T. Heydt, and R. Ayyanar, "Probabilistic power flow studies for transmission systems with photovoltaic generation using cumulants," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2251–2261, Nov. 2012.
- [13] A. M. Leite da Silva and A. M. de Castro, "Risk assessment in probabilistic load flow via Monte Carlo simulation and cross-entropy method," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1193–1202, Mar. 2019.
- [14] T. Williams and C. Crawford, "Probabilistic load flow modeling comparing maximum entropy and Gram-Charlier probability density function reconstructions," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 272–280, Feb. 2013.
- [15] Q. Chen and L. Mili, "Composite power system vulnerability evaluation to cascading failures using importance sampling and antithetic variates," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2321–2330, Aug. 2013.
- [16] J. Li and D. Xiu, "Evaluation of failure probability via surrogate models," *J. Comput. Phys.*, vol. 229, no. 23, pp. 8966–8980, Nov. 2010.
- [17] J. Li, J. Li, and D. Xiu, "An efficient surrogate-based method for computing rare failure probability," *J. Comput. Phys.*, vol. 230, no. 24, pp. 8683–8697, Oct. 2011.
- [18] A. M. Leite da Silva, R. A. G. Fernández, and C. Singh, "Generating capacity reliability evaluation based on Monte Carlo simulation and cross-entropy methods," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 129–137, Feb. 2010.
- [19] H. Wu, Y. Zhou, S. Dong, and Y. Song, "Probabilistic load flow based on generalized polynomial chaos," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 820–821, Jan. 2017.
- [20] F. Ni, P. H. Nguyen, and J. F. G. Cobben, "Basis-adaptive sparse polynomial chaos expansion for probabilistic power flow," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 694–704, Jan. 2017.
- [21] Y. Xu, L. Mili, and J. Zhao, "Probabilistic power flow calculation and variance analysis based on hierarchical adaptive polynomial chaos-ANOVA method," *IEEE Trans. Power Syst.*, vol. 34, no. 5, pp. 3316–3325, Sep. 2019.
- [22] D. Xiu and G. E. Karniadakis, "The Wiener-Askey polynomial chaos for stochastic differential equations," *SIAM J. Sci. Comput.*, vol. 24, no. 2, pp. 619–644, 2002.
- [23] J. Zhang, W. Li, G. Lin, L. Zeng, and L. Wu, "Efficient evaluation of small failure probability in high-dimensional groundwater contaminant transport modeling via a two-stage Monte Carlo method," *Water Resour. Res.*, vol. 53, no. 3, pp. 1948–1962, Mar. 2017.
- [24] D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*. Princeton, NJ: Princeton University Press, 2010.
- [25] Y. Xu and L. Mili, "Probabilistic power flow analysis based on the adaptive polynomial chaos-ANOVA method," *IEEE Power and Energy Soc. Gen. Meeting*, 2018.
- [26] Z. Hu and X. Wang, "A probabilistic load flow method considering branch outages," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 507–514, May 2006.
- [27] A. M. Leite da Silva, R. N. Allan, S. M. Soares, and V. L. Arienti, "Probabilistic load flow considering network outages," *IEE Proc. C-Gener. Transm. Distrib.*, vol. 132, no. 3, pp. 139–145, May 1985.
- [28] Y. Xu and L. Mili, "A novel polynomial-chaos-based Kalman filter," *IEEE Signal Process. Lett.*, vol. 26, no. 1, pp. 9–13, Jan. 2019.
- [29] W. Gautschi, "On generating orthogonal polynomials," *SIAM J. Sci. Stat. Comput.*, vol. 3, no. 2, pp. 289–317, Sep. 1982.
- [30] X. Wan and G. E. Karniadakis, "Multi-element generalized polynomial chaos for arbitrary probability measures," *SIAM J. Sci. Comput.*, vol. 28, no. 3, pp. 901–928, 2006.
- [31] X. Wan and G. E. Karniadakis, "An adaptive multi-element generalized polynomial chaos method for stochastic differential equations," *J. Comput. Phys.*, vol. 209, no. 2, pp. 617–642, 2005.
- [32] R. B. Adler *et al.*, "An IEEE survey of US and Canadian overhead transmission outages at 230 kV and above," *IEEE Trans. Power Del.*, vol. 9, no. 1, pp. 21–39, Jan. 1994.
- [33] University of Washington, Power Systems Test Case Archive. (accessed: 6/5/2019). [Online]. Available: <https://labs.ece.uw.edu/pstca/>
- [34] Q. Chen, "The probability, identification, and prevention of rare events in power systems," Ph.D. dissertation, Iowa State University, 2004.