Risk Assessment of Rare Events in Probabilistic Power Flow via Hybrid Multi-Surrogate Method

Yijun Xu, *Member, IEEE*, Mert Korkali, *Senior Member, IEEE*, Lamine Mili, *Life Fellow, IEEE*, Xiao Chen, and Liang Min, *Senior Member, IEEE*

***Abstract*—The risks associated with rare events threatening the security of power system operation are of paramount importance to power system planners and operators. To analyze the risks caused by high-impact, low-frequency rare events, an immensely large number of samples are typically required for the Monte-Carlo (MC) method on the high-fidelity power system model to achieve a sufficient accuracy, thereby rendering this approach computationally prohibitive. To handle this problem efficiently, it is desirable to construct a surrogate model for the power system response. However, the straightforward MC sampling of the low-fidelity surrogate can lead to biased results in the low-probability tail regions that are vital to risk assessment. Moreover, a single surrogate is unable to handle the topology uncertainties caused by random branch outages. To overcome these issues, we propose a hybrid multi-surrogate (HMS) method based on the polynomial chaos expansion (PCE) with low-probability tail events reevaluated by the high-fidelity model through a probabilistic analysis. This method improves the computational efficiency of the MC method for rare-event risk assessment by leveraging multi-fidelity models while retaining the desired accuracy. Simulations conducted in three test systems verify the excellent performances of the HMS method.**

***Index Terms*—Risk assessment, rare events, branch outage, Monte-Carlo sampling, probabilistic power flow, polynomial chaos expansion.**

I. Introduction

POWER system operators regularly perform security analysis to ensure that the network does not operate outside tolerable limits that contribute to, e.g., line overloads, voltage limit excursions, and voltage instability. However, power system is inherently stochastic. Sources of stochasticity include continuous load variations over time, intermittency of renewable energy resources, and random outages of transmission lines and transformers, to cite a few [1]. These uncertainties can bring formidable challenges to power system planning and operation. Ignoring them will produce inappropriate planning strategies or control actions, which, in turn, may result in unprecedented system failures.

Facing these challenges, several probabilistic methods to account for the uncertainties in power system operation and planning have been advocated in the literature over the past decades [2]–[7]. Among them, methods based on probabilistic power flow (PPF) are developed 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 flows and voltage magnitudes. The obtained pdf 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 [8], [9], the probabilities associated to risk events at the tail regions of the pdfs are of more interest among system planners and operators [2], [10]–[12]. This motivates us to conduct the risk assessment for these rare events in PPF analysis.

To solve this problem, the MC simulations are typically adopted for its high accuracy and implementation flexibility [13]. However, the impediment arises from the prohibitively large 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 for 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 [11], [14], [15], processing of which is too time-consuming for realistic power system applications. 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 cannot still be reduced significantly [11], [13]. 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, Cornish-Fisher and Edgeworth expansions can exhibit poor tail behavior while the Gram-Charlier series may yield negative cumulative probability values [2], [10], which makes it impractical. To address this issue, Williams and Crawford [12] propose the combined use of the maximum entropy and the Gram-Charlier expansion to

Manuscript received March 28, 2019; revised July 3, 2019; accepted September 8, 2019. Date of publication September 11, 2019; date of current version February 19, 2020. This work was supported in part by the U.S. National Science Foundation under Grant 1917308, and in part by the United States 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 (document released as LLNL-JRNL-770044) under Contract DE-AC52-07NA27344. Paper no. TSG-00465-2019. *(Corresponding author: Xiao Chen.)*

Y. Xu and L. Mili are with the Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Northern Virginia Center, Falls Church, VA 22043 USA (e-mail: yijunxu@vt.edu; lmili@vt.edu).

M. Korkali is with the Engineering Division, Lawrence Livermore National Laboratory, Livermore, CA 94550 USA (e-mail: korkali1@llnl.gov).

X. Chen is with the Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, Livermore, CA 94550 USA (e-mail: chen73@llnl.gov).

L. Min is with the Bits & Watts Initiative at Stanford University, Stanford, CA 94305 USA (e-mail: liangmin@stanford.edu).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSG.2019.2940928

construct pdfs based on cumulant arithmetic treatment of the linearized power flow. While this would seemingly be a cost-effective method, the linearization assumption for power-flow models might not hold for systems that are highly stressed under rare events.

To overcome the shortcomings of the aforementioned methods, a polynomial-chaos-expansion (PCE)-based response-surface model, also known as a *surrogate* (or *reduced-order*) model, has been advocated by [2], [16]–[18]. This PCE-based surrogate model can capture the behavior of the complicated, high-fidelity simulation model of a power system very closely while being computationally inexpensive to evaluate [19], 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 vital in risk assessments [14], [15], [20]. Furthermore, the construction of the PCE-based surrogate model relies on the assumption of the smoothness of the original system function [21]. However, this assumption cannot be guaranteed when the topology changes caused by random branch outages are considered. Indeed, none of the aforecited PCE-based methods is able to account for the uncertainties of branch outages [2], [16]–[18]. However, these uncertainties may have much greater influence on system states than those of nodal power injections. Therefore, these structural uncertainties should not be neglected in assessing risks under rare events [22]. 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 [22], [23]. The crude pdf obtained from only several thousands of samples cannot 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 [22].

To address the abovementioned issues, this paper proposes a novel PCE-based HMS method for the risk assessment of rare events in PPF analysis, resulting in the following contributions:

- This paper demonstrates the first attempt to use a PCE-based method to conduct risk assessment in PPF analysis. Not only are the uncertainties from the active power and reactive power of loads considered, but also the structural uncertainties associated with random branch outages, including a number of extreme events, i.e., $N-2$, $N-3$, and higher-order contingencies, are accounted for.
- To handle the inaccuracy of the PCE method in sampling the low-probability tail regions, a more accurate hybrid-PCE method is developed [14], [15], [20].
- To overcome the incapability of the traditional PCE-based method in handling topology changes, this paper, for the first time, proposes a "multi-surrogate" model. This multi-surrogate model is further combined with a second-stage, *hybridized* procedure to deal with extreme-event cases. Amenability of the resultant hybrid multi-surrogate (HMS) method to parallel processing is also discussed.
- Simulation results reveal the excellent accuracy and the computing efficiency of the proposed method.

This paper is organized as follows: after formulating the problem of risk assessment in Section II, we briefly review the MC and response-surface methods in Section III. The hybrid surrogate method and the proposed HMS method are presented in Section IV. Section V presents the results for test cases. Conclusions and future work are highlighted in Section VI.

II. Problem Formulation

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

$$z = f(\boldsymbol{m}). \tag{1}$$

Here, z stands for the QoI, e.g., voltage magnitude, voltage stability margin, and line flow; $\boldsymbol{m} = [m_1, m_2, \ldots, m_N]$ 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 [2], [22]; the $f(\cdot)$ is the nonlinear function that represents the power system model, which maps the model parameters, $\boldsymbol{m}$, to the QoI, z. In this paper, $f(\cdot)$ denotes the ac power-flow model. Furthermore, $f(\cdot)$ can represent other power system models, e.g., voltage stability model [3], centralized power system dynamic model [6], or decentralized synchronous generator model [7].

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

$$P_f = \Pr\big(Z \in \Omega_f\big) = \int_{\Omega_f} q(z)\mathrm{d}z = \int \chi_{\Omega_f}(z)q(z)\mathrm{d}z, \tag{2}$$

where χ is the characteristic function satisfying

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

and Ω_f is the target domain defined as

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

Here, $g(z)$ is a limit state function, also called performance function that defines the target domain for risk assessment. More specifically, the domain where $g(z) < 0$ stands for the domain with risk, and the domain where $g(z) \geq 0$ stands for the safe domain [14], [15]. Here, based on (2), the probability of failure can be represented as

$$P_f = \int \chi_{\{g(z)<0\}}(z)q(z)\mathrm{d}z. \tag{5}$$

Now, we have completed the formulation of the risk assessment in PPF analysis. For the rare events considered in this paper, the failure probability is typically very small, 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 1: As Aven *et al.* have summarized in [24], the "risk" terminology has multiple qualitative definitions, e.g., the possibility of an unfortunate occurrence, or the potential for realization of unwanted, negative consequences of an event, or the consequences of the activity and associated uncertainties.

In this article, we choose, as our measure of risk, the failure probability exceeding the operational limit in PPF analysis by following the practice suggested by da Silva and de Castro [11]. Furthermore, the concept of "risk" can be easily extended to other quantitative metric, e.g., the average amount of QoI that exceeds the target. Then, (3) is reformed as

$$\chi_{A\Omega_f}(z) = \begin{cases} \Delta A & \text{if } z \in \Omega_f, \\ 0 & \text{if } z \notin \Omega_f. \end{cases} \tag{6}$$

Here, ΔA is the value of the QoI falling below or above the allowed limit.

III. Algorithm Preliminaries

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

A. Relationship Between the MC Simulation and the Response Surface Method

The most straightforward way to conduct a risk assessment is the MC simulation, where a set of N_m samples are drawn from the multivariate probability distribution of $\boldsymbol{m}$, yielding $\{\boldsymbol{m}^{(j)}\}_{j=1}^{N_m}$. Then, for each $\boldsymbol{m}^{(j)}$, $j = 1, \ldots, N_m$, the QoI is realized through $z^j = f(\boldsymbol{m}^{(j)})$, yielding a set of $\{z^{(j)}\}_{j=1}^{N_m}$, which includes the system realizations both including and excluding branch outages. The failure probability is calculated as

$$P_f^{\mathrm{MC}} = \sum_{j=1}^{N_m} \frac{1}{N_m} \chi_{\{g(z)<0\}}\left(z^j\right). \tag{7}$$

Despite its easy implementation, the computational burden for realizing 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 $\widetilde{f}(\boldsymbol{m})$ that holds a relationship with the original complicated power system model $f(\boldsymbol{m})$ as $\widetilde{f}(\boldsymbol{m}) \approx f(\boldsymbol{m})$. By this way, the uncertainties of the system model can be propagated through the surrogate model $\widetilde{f}(\boldsymbol{m})$ at little or no extra computational cost. The limit state function for the response surface model described as $\widetilde{g}(z)$ also holds for $\widetilde{g}(Z) \approx g(Z)$. This enables us to estimate the risk probability of the rare events by response surface method via

$$P_f^{\mathrm{RS}} = \sum_{j=1}^{N_m} \frac{1}{N_m} \chi_{\{\widetilde{g}(z)<0\}}\left(z^j\right). \tag{8}$$

By this way, the risk assessment problem can be conducted in the surrogate model efficiently. Next, we will present the way to construct the surrogate model via PCE.

B. Review of PCE-Based Response Surface

Introduced by Wiener and further developed by Xiu and Karniadakis [19] and Xiu [21], the generalized polynomial chaos expansion has been shown to be a cost-effective tool in modeling response surfaces [2], [16]–[18]. 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 $\boldsymbol{\xi} = [\xi_1, \xi_2, \ldots, \xi_N]$ be a vector of random variables following a standard probability distribution (e.g., the Gaussian or the beta distribution), to which, as shown in [21], a unique orthogonal polynomial is associated. Let $\Phi_i(\xi_1, \xi_2, \ldots, \xi_N)$ denote this procedure's corresponding polynomial chaos basis and a_i denote the ith polynomial chaos coefficient. Formally, we have

$$z = \sum_{i=0}^{N_P} a_i \Phi_i(\boldsymbol{\xi}), \tag{9}$$

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 [2], [16], [18]. From the polynomial chaos coefficients, the mean, μ, and the variance, σ^2, of the output z can be determined as

$$\mu = a_0, \tag{10}$$

$$\sigma^2 = \sum_{i=1}^{N_P} a_i^2 \mathrm{E}\left[\phi_i^2\right], \tag{11}$$

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

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

$$\int_{\mathbb{R}} \phi_r(\xi)\phi_s(\xi)\mathrm{d}\lambda \begin{cases} = 0 & \text{if } r \neq s, \\ > 0 & \text{if } r = s. \end{cases} \tag{12}$$

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, $\{\phi_i(\boldsymbol{\xi}), i = 1, 2, 3, \ldots\}$, is orthogonal to each other with respect to their joint probability measure.

2) Three-Term Recurrence Relation: The orthogonal polynomials satisfy a three-term recurrence relation given by [25]

$$\begin{aligned} &\phi_{k+1}(\xi) = (\xi - \alpha_k)\phi_k(\xi) - \beta_k \phi_{k-1}(\xi), \\ &\quad \phi_{-1} = 0, \quad \phi_0 = 1; \quad k = 0, 1, 2, \ldots, K, \end{aligned} \tag{13}$$

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

$$\phi_k(\xi) = \xi^k + \text{lower-degree terms}, \quad k = 0, 1, \ldots, K, \tag{14}$$

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

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 [18], [26]. It is given by

$$\alpha_k = \frac{\int_{\mathbb{R}} \xi \phi_k^2(\xi)\mathrm{d}\lambda(\xi)}{\int_{\mathbb{R}} \phi_k^2(\xi)\mathrm{d}\lambda(\xi)}, \quad k = 0, 1, 2, \ldots, K, \tag{15}$$

$$\beta_k = \frac{\int_{\mathbb{R}} \phi_k^2(\xi)\mathrm{d}\lambda(\xi)}{\int_{\mathbb{R}} \phi_{k-1}^2(\xi)\mathrm{d}\lambda(\xi)}, \quad k = 1, 2, \dots, K. \tag{16}$$

Here, β_0 is arbitrary and can be conveniently chosen as $\beta_0 = \int_{\mathbb{R}} \mathrm{d}\lambda(\xi)$ and K is the highest order of the polynomials. If the measure consists of n discrete points, the integrals in (15) and (16) become summations.

4) Construction of the Polynomial Chaos Basis: A set of multidimensional 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

$$\phi(\boldsymbol{\xi}) = \phi(\xi_1) \otimes \phi(\xi_2) \otimes \cdots \otimes \phi(\xi_N), \tag{17}$$

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

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 [2], [21]. 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 [2].

C. Building a PCE-Based Surrogate for Power System Response

Here, the uncertain parameters $\boldsymbol{m}$ in the power system model are viewed as random variables following certain types of distributions. By mapping the parameters $\boldsymbol{m}$ into $\boldsymbol{\xi}$, we can build a PCE as the response surface of power-flow solutions. The detailed PCE procedure is as follows:

(1) Map the ith random parameter, m_i, to a given random variable, ξ_i, as follows:

$$m_i = F_i^{-1}(T(\xi_i)), \tag{18}$$

where F_i^{-1} is the inverse cdf of m_i and T is the cdf of ξ_i;

(2) Construct the polynomial chaos basis; then, express the output z in the gPC expansion form of (9);

(3) Construct M combinations of collocation points and put them into the polynomial chaos basis $(M \times (N_P + 1))$ matrix $\mathbf{H_{pc}}$. Formally, we have

$$\mathbf{H_{pc}} = \begin{bmatrix} \Phi_0(\boldsymbol{\xi_1}) & \Phi_1(\boldsymbol{\xi_1}) & \cdots & \Phi_{N_P}(\boldsymbol{\xi_1}) \\ \Phi_0(\boldsymbol{\xi_2}) & \Phi_1(\boldsymbol{\xi_2}) & \cdots & \Phi_{N_P}(\boldsymbol{\xi_2}) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_0(\boldsymbol{\xi_M}) & \Phi_1(\boldsymbol{\xi_M}) & \cdots & \Phi_{N_P}(\boldsymbol{\xi_M}) \end{bmatrix}; \tag{19}$$

(4) Compute the power-flow output for the selected collocation points to get the $(M \times 1)$ output matrix $\mathbf{z}$ given by

$$\mathbf{z} = \begin{bmatrix} z(\boldsymbol{\xi_1}) & z(\boldsymbol{\xi_2}) & \dots & z(\boldsymbol{\xi_M}) \end{bmatrix}^{\mathsf{T}}; \tag{20}$$

(5) Estimate the unknown coefficients $\mathbf{a}$ based on the collocation points that are selected and the model output from

$$\mathbf{z} = \mathbf{H_{pc}}\mathbf{a}, \tag{21}$$

where $\mathbf{a}$ is the $(N_P \times 1)$ coefficient vector expressed as

$$\mathbf{a} = \begin{bmatrix} a_0 & a_1 & \dots & a_{N_P} \end{bmatrix}^{\mathsf{T}}; \tag{22}$$

(6) Let $\mathbf{r}$ denote the residual vector defined as $\mathbf{r} = \mathbf{z} - \mathbf{H_{pc}}\mathbf{a}$. An estimated $\hat{\mathbf{a}}$ can be obtained by minimizing the 2-norm of the residual vector, i.e., $J(\hat{\mathbf{a}}) = \arg\min_{\hat{\mathbf{a}}} \mathbf{r}^{\mathsf{T}}\mathbf{r}$, which yields

$$\hat{\mathbf{a}} = \left(\mathbf{H_{pc}^{\mathsf{T}}}\mathbf{H_{pc}}\right)^{-1}\mathbf{H_{pc}^{\mathsf{T}}}\mathbf{z}. \tag{23}$$

With the coefficients estimated and the bases selected, we can build the PCE for the target output. The power system response surface can now be represented in a polynomial form as the surrogate model.

Remark 2: When renewable energy generation is considered, correlations among input variables cannot be ignored. To handle these correlations, several methods have been proposed in the literature, e.g., a whitening transformation [2], [3], [18], the Nataf transformation [17], and the Karhunen-Loève expansion [21]. In this work, we solely focus on the topics of risk assessment and topology uncertainties. The handling of correlations is beyond the scope of this paper.

IV. The Proposed 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.

A. Hybrid PCE-Based Method

Motivated by the fact that the direct surrogate-based MC simulation may introduce a significant error in the small-probability tail regions of the pdfs, Li and Xiu [14] first developed the hybrid PCE method, which is shown to be a cost-effective tool in handling small failure probabilities [15], [20]. 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 [20], prevents loss of accuracy coming from the direct usage of the surrogate-based method. Therefore, based on (7) and (8), the risk probability of rare events via the hybrid approach can be obtained from

$$P_f^{\mathrm{H}} = \sum_{j=1}^{N_m} \frac{1}{N_m} \chi_{\tilde{\Omega} f}\left(z^j\right), \tag{24}$$

where the approximate target domain is defined as

$$\widetilde{\Omega}_f \triangleq \{\widetilde{g}(Z) < -\gamma\} \cup \{\{|\widetilde{g}(Z)| \leq \gamma\} \cap \{g(Z) < 0\}\}. \quad (25)$$

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 will lead to more reevaluations of the original model, yet with a higher 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 simulation based on the original model [14]. Therefore, the hybrid approach enjoys both the accuracy of the MC method and the efficiency of the surrogate-based method.

B. Hybrid Multi-Surrogate (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 [26], [27]. 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 [6], [26], [27], we propose a multi-surrogate method. 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 and, 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 adopt the idea of using 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. cases. For the original system model without any branch outages and for the $N-1$ topologies, 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.

TABLE I
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

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 (20), 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 simulation 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 $N-1$ cases occur more frequently than the higher-order contingencies, e.g., $N-2$, $N-3$, as shown in Table I taken from [28].

Furthermore, for all the $N-1$ cases, we only need to construct N_{branch} surrogate models. For the risk assessment targeted at the rare events associated with 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 [11], [14], [15]. 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 $N-2$ or higher-order contingencies, the number of surrogates increases combinatorially (e.g., $\binom{N_{\text{branch}}}{2}$ and $\binom{N_{\text{branch}}}{3}$ for $N-2$ and $N-3$ cases, 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 yellow are

Algorithm 1 Risk Assessment via the HMS Algorithm

1: Model the uncertainties with their associated pdfs, including the uncertainties from load variations and branch states; choose the QoI for risk assessment;
2: Only considering the uncertainties of nodal power injections, construct the PCE-based surrogate model of the system without branch outage;
3: **for** $k = 1, \ldots, N_{\text{branch}}$ **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 $\{\boldsymbol{m}^{(j)}\}_{j=1}^{N_m}$, including uncertainties for nodal power injections and branch states;
7: **for** $j = 1, \ldots, N_m$ **do**
8:     Classify $\boldsymbol{m}^{(j)}$ into different types of contingencies, i.e., $N-0$, $N-1$, $N-2$, etc.
9:     **if** $\boldsymbol{m}^{(j)}$ belongs to $N-0$ or $N-1$ cases **then**
10:         Find the corresponding surrogate model for $\boldsymbol{m}^{(j)}$ and compute z^j via its surrogate model $\widetilde{f}(\boldsymbol{m}^{(j)})$;
11:         **if** z^j located in target domain $\widetilde{\Omega}_f$ **then**
12:             Mark sample $\boldsymbol{m}^{(j)}$ for reevaluation;
13:         **end if**
14:     **else**
15:         Compute z^j via original system model $f(\boldsymbol{m}^{(j)})$;
16:     **end if**
17: **end for**
18: Reevaluate all the samples marked in Step 12 via $f(\boldsymbol{m}^{(j)})$;
19: Perform risk assessment via (24)

processed with parallel computing. The detailed procedure for implementing the HMS algorithm is illustrated in Algorithm 1. Note that if the outages create islands in the system, we first identify each island and solve for the ac power flow separately to find the corresponding QoI.

Remark 3: It is worth pointing out the current HMS approach is still based on the sampled data. However, this method can be extended to historical utility outage data as the HMS method improves the computational efficiency through reduced-order modeling which has no limitation on the sources of the samples. In case the utility outage data is available, data generation introduced in Step 6 of Algorithm 1 is not necessary. True data are evaluated through this hybrid model directly.

V. 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) and the 118-bus system whose network and reliability data are extracted from [29]. 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. We further extend the proposed method to a very-large-scale power system, the European 1,354-bus system. The data can be obtained form [30], [31]. For this very large power system test case, the simulation results are obtained on a desktop with 3.50-GHz Intel Xeon CPU E5-1650 v2 processors and a 32 GB of main memory.

A. Results for IEEE 24-Bus RTS System

For this test system, we build a setup to capture the low-probability 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

$$P_\ell = \frac{\alpha \lambda_p t}{8,760}, \tag{26}$$

where λ_p is the permanent outage rate [outages/yr]; t is the permanent outage duration [h]; and α is the scaling factor [22]. The data for λ_p, t, and α are provided in [29]. The proposed HMS method is verified in the risk 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, with the generated samples provided in Table II. 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 to parallelize the executions of the MC and of the HMS methods as displayed in the yellow blocks shown in Fig. 1. Let us select the voltage magnitude at Bus 3 as the QoI. Now, we seek to assess the risk of the QoI falling below its minimum operating limit, i.e., 0.95 pu. The simulation results obtained with the MC method are provided in Fig. 2. As α increases, the tails of the pdfs become thicker, leading to a higher risk probability, P_f^{MC}. However, this risk probability is still too low to make an accurate assessment with the traditional response-surface-based method. This motivates 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 III.

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 very-low-probability events, e.g., 10^{-4} or 10^{-5}.
- A larger γ will lead to a larger state space required for reevaluation and therefore to an increasing number of model reevaluations and a larger computing cost, but further improving the simulation accuracy.
- With a higher PCE order P, the simulation results tend to be more accurate under the same γ values. For a relatively

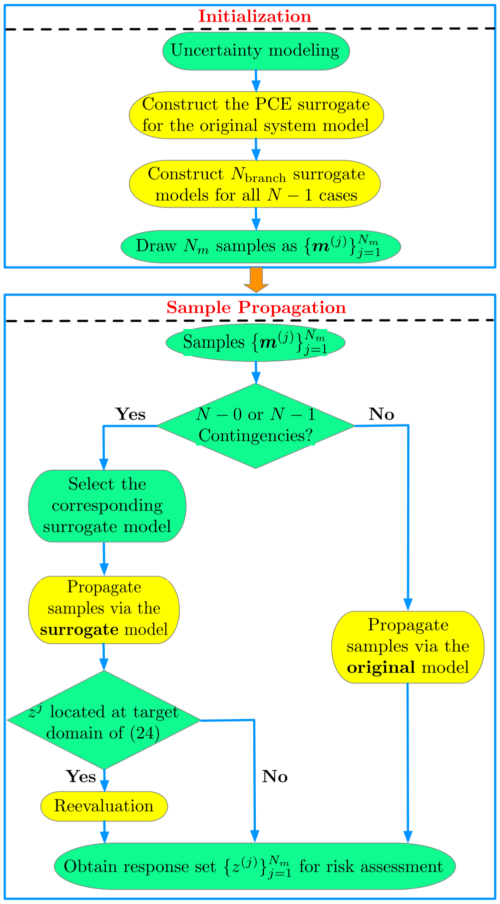

Fig. 1. Flowchart of risk assessment via the HMS method.

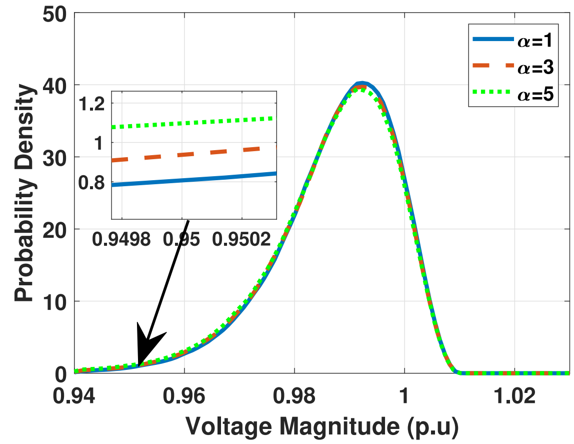

Fig. 2. Pdf of voltage magnitude at Bus 3 with different scaling factors α.

TABLE II
SAMPLE SIZES OF RANDOM BRANCH OUTAGES USED IN THE IEEE RTS 24-BUS SYSTEM

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

TABLE III
VALIDATION OF THE HMS METHOD WITH DIFFERENT VALUES OF P AND γ FOR SAMPLE GROUPS A–C ON THE IEEE RTS 24-BUS SYSTEM

Group A Data Validation					
Order P	γ	**Resamples**	$P_f^{\text{H}}(\%)$	$P_f^{\text{MC}}(\%)$	**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	$2\cdot10^{-3}$	3059	0.843	0.845	34.1
2	$3\cdot10^{-3}$	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					
Order P	γ	**Resamples**	$P_f^{\text{H}}(\%)$	$P_f^{\text{MC}}(\%)$	**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	$2\cdot10^{-3}$	3496	1.31	1.31	53.1
2	$3\cdot10^{-3}$	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					
Order P	γ	**Resamples**	$P_f^{\text{H}}(\%)$	$P_f^{\text{MC}}(\%)$	**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	$2\cdot10^{-3}$	4092	1.81	1.82	78.1
2	$3\cdot10^{-3}$	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

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 due to two major reasons. First, this is because when α increases, more samples related to higher-order contingencies are generated. These samples are all propagated through the time-consuming ac power-flow model $f(\cdot)$ instead of the computationally inexpensive surrogate model $\widetilde{f}(\cdot)$, thereby increasing the computing time at the first-stage MC period. Second, the tails of pdfs become thicker as shown in Fig. 2. This potentially leads to more reevaluations for the same target domain in the second stage of the MC period. In sum, when the number of the higher-order contingencies increases, the computational efficiency of the proposed method decreases proportionately. This is the tradeoff between the PCE-based surrogate model and the original power system model.

Remark 4: It is worth noting that the HMS method can be extended to include failures of other system components, such as substations, the failure of which will also cause topology

TABLE IV
VALIDATION OF THE HMS METHOD CONSIDERING BRANCH AND SUBSTATION OUTAGES ON THE IEEE RTS 24-BUS SYSTEM

Order P	γ	**Resamples**	$P_f^{\mathrm{H}}(\%)$	$P_f^{\mathrm{MC}}(\%)$	**Time (s)**
2	10^{-3}	1664	1.04	1.10	36.5
2	$3 \cdot 10^{-3}$	5032	1.09	1.10	56.4

TABLE V
VALIDATION OF THE HMS METHOD WITH DIFFERENT VALUES OF P, γ, AND QOI LIMIT SETTINGS ON THE IEEE 118-BUS SYSTEM

Order P	γ	**Resamples**	$P_f^{\mathrm{H}}(\%)$	$P_f^{\mathrm{MC}}(\%)$	**Time (s)**
Group 1 with limit set to 60 MVA					
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					
Order P	γ	**Resamples**	$P_f^{\mathrm{H}}(\%)$	$P_f^{\mathrm{MC}}(\%)$	**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					
Order P	γ	**Resamples**	$P_f^{\mathrm{H}}(\%)$	$P_f^{\mathrm{MC}}(\%)$	**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					
Order P	γ	**Resamples**	$P_f^{\mathrm{H}}(\%)$	$P_f^{\mathrm{MC}}(\%)$	**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

changes and threaten the security of the power grid. The following part will provide a demo case considering substation outages.

Here, due to the unavailability of outage data for substations, we assume all the PQ buses to be substations and set the failure rate of each substation to be 0.1% while adopting a scaling factor $\alpha = 1$. In this case, we generate samples with the size of 10^6, among which, $962,231$ are $N-0$, $37,099$ are $N-1$, 659 are $N-2$, and 11 are $N-3$ event samples. Compared to Table II, it is shown that the number of samples associated with higher-order contingencies increases because more component failures are considered in the case study. 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 QoI remains the same as in the previous case. The simulation results are depicted in Table IV.

As is shown in Table IV, the HMS method can provide accurate estimation results at a much cheaper computing cost, compared to 1.5 h required by the traditional MC method. This also demonstrates the capability of the HMS method to be extended to include other component outages.

B. Results for IEEE 118-Bus System

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. 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 [2]. Since line-outage rates have not been provided for the IEEE 118-bus system in [29], we set the failure rate of each transmission line to 0.1% by multiplying the average failure rate reported in [22] 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 [11]. 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 V 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 accuracy of the simulation results is not highly sensitive to γ as shown in Section V-A. The assessment of a low probability, $P_f^{\mathrm{H}}(\%)$, of the risk of multiple contingencies remains very accurate even if γ is set to a very small value.

To illustrate the last point above, let us check the pdf plots for the tails in different QoI limit settings as shown in Fig. 3. It can be seen that the HMS method can provide almost the same pdfs as those of the MC method. Furthermore, as observed in Fig. 3, the pdf of the QoI, which ranges from 10^{-4} to 10^{-2}, is much lower than that shown in Fig. 2, which is around 1. This means that for the same domain range determined by γ, a pdf with lower magnitudes is obtained when using a much lower number of samples for the same sample size. In this case, even with a small γ, the risk probability is unlikely to be heavily influenced by such a small amount of samples in this region. On the other hand, it is because that the system response in the target domain for apparent power is so smooth that it can be accurately approximated even with second-order polynomial chaos basis functions. Therefore, a small γ with a low-order polynomial chaos basis function can readily ensure a good performance of the HMS method.

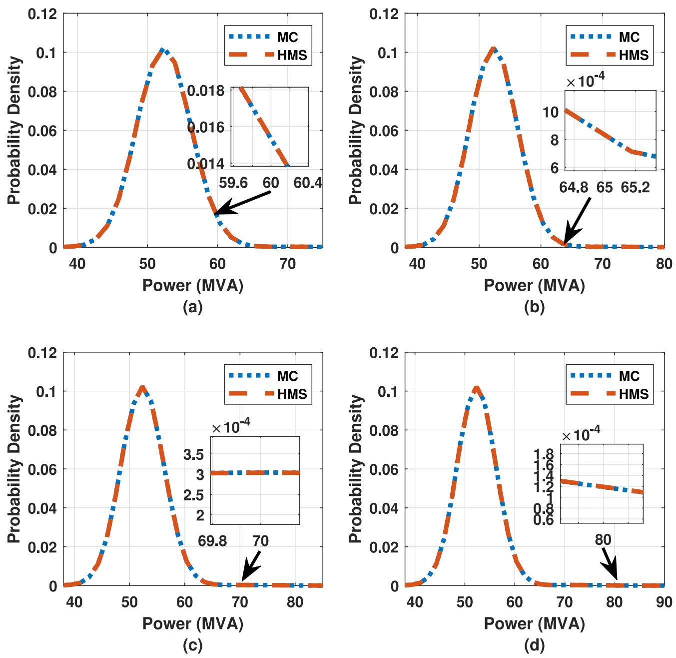

Fig. 3. Pdfs and their target region of the apparent power in Line 117 for (a) Group 1 with $P = 2$, $\gamma = 0.1$; (b) Group 2 with $P = 2$, $\gamma = 0.1$; (c) Group 3 with $P = 2$, $\gamma = 0.1$; and (d) Group 4 with $P = 2$, $\gamma = 0.1$.

TABLE VI
VALIDATION OF THE HMS METHOD USING EXTREME-VALUE ANALYSIS METRICS

Flow Limit	60 **MVA**	70 **MVA**	80 **MVA**
Δ**Flow** (HMS)	0.1168	0.0035	0.0011
Δ**Flow** (MC)	0.1168	0.0035	0.0011

We further test our proposed method using some extreme-value analysis metrics instead of simply adopting a failure probability. We quantify the average amount of apparent power that exceeds different limit settings for the transmission lines. In this case, the characteristic function is chosen as (6) rather than (3). We set $\gamma = 0.1$ and $P = 2$ for the HMS method. The comparative results of the HMS and MC methods are shown in Table VI. It can be seen that the proposed method can provide very accurate estimation results even using some extreme-value analysis metrics. Furthermore, adopting these metrics makes no difference in the computational efficiency of the HMS method compared to the metrics for calculating the failure probability.

C. Results on the European 1,354-Bus System

Here, we demonstrate the proposed HMS method in a much larger power system. This case study is conducted on the European high-voltage transmission system, which contains 1,354 buses, 260 generators, and 1,991 branches, and operates at both 380 and 220 kV [30], [31]. For MC, we use a sample size of 10^6 to generate a collection of $N-k$ events that fit a negative binomial distribution. Of these, $451,023$ are $N-0$, $358,887$ are $N-1$, $142,994$ are $N-2$, $38,066$ are $N-3$, $7,617$ are $N-4$, $1,253$ are $N-5$, 143 are $N-6$, 15 are $N-7$, and 2 are $N-8$ event samples, while the samples required to construct as large as 1,992 surrogates in the HMS method are still 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 [2]. Since line-outage rates are unavailable for the European 1,354-bus system, we set the failure rate of each transmission line to 0.04%, which is the average failure rate reported in [22] for the IEEE RTS 24-bus system. In this case, the QoI is chosen to be the apparent power of Line 1005 between Buses 5420 and 5658. We validate the accuracy and computational efficiency of the HMS method in dealing with different limit settings for the QoI, and the results in Tables VII and VIII are obtained. It should be emphasized that in order to carry out the simulation on such a large-scale power system, we use a desktop environment running 6 CPUs in parallel to execute the simulation runs conducted with the MC and HMS methods (as displayed in the yellow blocks shown in Fig. 1) with detailed settings described above. Nonetheless, we expect that running simulations on a different machine would not dramatically alter the speedup gain for this system.

These results lead to the following conclusions:

- For the European 1,354-bus system, the simulation results obtained using the HMS method are very accurate compared to the benchmark data obtained via the MC method. Even with a PCE order of 2, the desired accuracy can be maintained. In other words, the accuracy of the proposed method is not influenced by the system size.
- The accuracy of the simulation results is not highly sensitive to γ as shown in Section V-A. The failure probability $P_f^{\mathrm{H}}(\%)$ remains very accurate even if γ is set to a very small value.
- From the point of computing performance, simulations carried out with the HMS method run to completion in around 1.4 h, which is much faster than that of the

TABLE VII
VALIDATION OF THE HMS METHOD WITH DIFFERENT VALUES OF P, γ, AND QOI LIMIT SETTINGS ON THE EUROPEAN 1,354-BUS SYSTEM

Group 1 **with limit set to** 550 **MVA**				
Order P	γ	**Resamples**	$P_f^{\mathrm{H}}(\%)$	$P_f^{\mathrm{MC}}(\%)$
2	0.1	14	0.31	0.32
2	1	106	0.32	0.32
Group 2 **with limit set to** 600 **MVA**				
Order P	γ	**Resamples**	$P_f^{\mathrm{H}}(\%)$	$P_f^{\mathrm{MC}}(\%)$
2	0.1	5	1.35	1.36
2	1	56	1.36	1.36

TABLE VIII
CPU TIMES OF THE HMS AND DIRECT MC METHODS FOR THE EUROPEAN 1,354-BUS SYSTEM

Group 1 **with limit set to** 550 **MVA**			
Method	**HMS** ($\gamma = 0.1$)	**HMS** ($\gamma = 1$)	**MC**
CPU Time	≈ 1.4 h	≈ 1.4 h	≈ 5 h
Group 2 **with limit set to** 600 **MVA**			
Method	**HMS** ($\gamma = 0.1$)	**HMS** ($\gamma = 1$)	**MC**
CPU Time	≈ 1.4 h	≈ 1.4 h	≈ 5 h

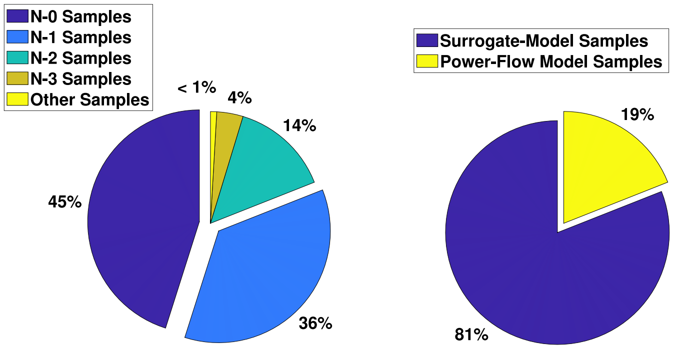

Fig. 4. Pie charts for sample types (left) and model types (right) for the European 1,354-bus system.

MC method that spends approximately 5 h to run the simulations.

- The speedup achieved by the HMS method reduces remarkably when compared to the ones obtained for the IEEE RTS 24- and 118-bus systems.

Here, let us further illustrate this last point. Since the number of samples used in the reevaluation stage is quite small, their influence on the total computing time is considered subtle. Nevertheless, the main reason for the reduced speedup factor is because the European 1,354-bus system has almost 2,000 branches, all of which are subjected to failure. This notably increases the portion of samples associated with $N-2$ and higher-order contingencies as shown in Fig. 4. These samples account for almost 20% of the total sample size, which must be evaluated through the time-intensive ac power-flow model instead of the computationally inexpensive surrogate model. The surrogate models can only be used to accelerate the samples associated with $N-0$ and $N-1$ cases, which account for around 80% of the sample size. This is a limitation of the proposed HMS method. On the other hand, similar to many existing methods (see [11], [13]) whose computational efficiency decreases when applied to larger, more complicated systems, the computational performance of the HMS method in a very-large-scale power system also decreases. Further improvement in the HMS method's performance in very large power system deserves further exploration.

VI. Conclusion and Future Work

In this paper, we propose a novel PCE-based HMS method in conducting a risk assessment for rare events in the PPF analysis. The proposed method overcomes the inefficacy of the surrogate-based method in handling structural uncertainties and the inaccuracy of the surrogate method in the tails of the QoI. Simulation results show that the HMS method can accurately capture the low-probability risks associated with rare events with a much better computational efficiency as compared to the traditional MC method.

However, the proposed method still has some limitations. When the scale of the system and hence the number of higher-order contingencies increases, the speedup gained by the proposed method reduces to a great extent. Further algorithmic enhancement is still needed to ameliorate the performance of the proposed method on larger-scale test cases.

Furthermore, the proposed method has the potential to be extended to solve other related research problems, such as composite system reliability analysis and resilience assessment [32], [33], which merit their own lines of research. Future work will tackle these problems by accommodating historical utility outage data (see [34]), additional sources of uncertainty, and long-term consequences (e.g., over 5-10 years) of high-impact, low-probability extreme events.

References

[1] M. Schwalbe, *Mathematical Sciences Research Challenges for the Next-Generation Electric Grid: Summary of a Workshop* (National Academies of Sciences, Engineering, and Medicine). Washington, DC, USA: Nat. Acad. Press, 2015.

[2] 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.

[3] E. Haesen, C. Bastiaensen, J. Driesen, 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.

[4] N. D. Hatziargyriou 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.

[5] 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.

[6] Y. Xu *et al.*, "Response-surface-based Bayesian inference for power system dynamic parameter estimation," *IEEE Trans. Smart Grid*, to be published. doi: 10.1109/TSG.2019.2892464.

[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] 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.

[9] 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.

[10] 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.

[11] A. M. L. 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.

[12] 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.

[13] 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.

[14] J. Li and D. Xiu, "Evaluation of failure probability via surrogate models," *J. Comput. Phys.*, vol. 229, no. 23, pp. 8966–8980, Nov. 2010.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] 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 Resources Res.*, vol. 53, no. 3, pp. 1948–1962, Mar. 2017.

[21] D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*. Princeton, NJ, USA: Princeton Univ. Press, 2010.

[22] 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.

[23] A. M. L. 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.

[24] *Society for Risk Analysis Glossary*. Accessed: Jun. 2019. [Online]. Available: https://sra.org/sites/default/files/pdf/SRA%20Glossary%20-%20FINAL.pdf

[25] W. Gautschi, "On generating orthogonal polynomials," *SIAM J. Sci. Stat. Comput.*, vol. 3, no. 2, pp. 289–317, Sep. 1982.

[26] 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.

[27] 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.

[28] R. B. Adler, S. L. Daniel, C. R. Heising, M. G. Lauby, R. P. Ludorf, and T. S. White, "An IEEE survey of U.S. and Canadian overhead transmission outages at 230 kV and above," *IEEE Trans. Power Del.*, vol. 9, no. 1, pp. 21–39, Jan. 1994.

[29] *Power Systems Test Case Archive*, Univ. Washington, Seattle, WA, USA. Accessed: Feb. 2019. [Online]. Available: https://labs.ece.uw.edu/pstca/

[30] C. Josz, S. Fliscounakis, J. Maeght, and P. Panciatici. *AC Power Flow Data in MATPOWER and QCQP Format: iTesla, RTE Snapshots, and PEGASE*. Accessed: Jul. 2019. [Online]. Available: https://arxiv.org/pdf/1603.01533.pdf

[31] S. Fliscounakis, P. Panciatici, F. Capitanescu, and L. Wehenkel, "Contingency ranking with respect to overloads in very large power systems taking into account uncertainty, preventive, and corrective actions," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4909–4917, Nov. 2013.

[32] M. Panteli and P. Mancarella, "Modeling and evaluating the resilience of critical electrical power infrastructure to extreme weather events," *IEEE Syst. J.*, vol. 11, no. 3, pp. 1733–1742, Sep. 2017.

[33] M. Panteli, C. Pickering, S. Wilkinson, R. Dawson, and P. Mancarella, "Power system resilience to extreme weather: Fragility modeling, probabilistic impact assessment, and adaptation measures," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3747–3757, Sep. 2017.

[34] M. R. Kelly-Gorham, P. Hines, and I. Dobson. *Using Historical Utility Outage Data to Compute Overall Transmission Grid Resilience*. Accessed: Jul. 2019. [Online]. Available: http://arxiv.org/pdf/1906.06811.pdf

Yijun Xu (M'19) received the Ph.D. degree from the Bradley Department of Electrical and Computer Engineering, Virginia Tech, Falls Church, VA, USA, in 2018. He is currently a Post-Doctoral Associate with Virginia Tech-Northern Virginia Center, Falls Church. He did the computation internship with Lawrence Livermore National Laboratory, Livermore, CA, USA, and power engineer internship with ETAP—Operation Technology, Inc., Irvine, CA, USA, in 2018 and 2015, respectively. His research interests include uncertainty quantification, risk assessment, power system dynamic state, and parameter estimation.

Mert Korkali (SM'18) received the Ph.D. degree in electrical engineering from Northeastern University, Boston, MA, USA, in 2013.

He is currently a Research Staff Member with Lawrence Livermore National Laboratory, Livermore, CA, USA. From 2013 to 2014, he was a Post-Doctoral Research Associate with the University of Vermont, Burlington, VT, USA. His current research interests lie at the broad interface of robust state estimation and fault location in power systems, extreme event modeling, cascading failures, uncertainty quantification, and probabilistic grid planning. He is currently serving as an Associate Editor for the *Journal of Modern Power Systems and Clean Energy* (Springer), and has been listed among the exceptional reviewers of the IEEE TRANSACTIONS ON POWER DELIVERY.

Lamine Mili (LF'17) received the Ph.D. degree from the University of Liège, Belgium, in 1987.

He is a Professor of electrical and computer engineering with Virginia Tech, Blacksburg. He has five years of industrial experience with the Tunisian electric utility, STEG. At STEG, he worked in the Planning Department from 1976 to 1979 and then at the Test and Meter Laboratory from 1979 till 1981. He was a Visiting Professor with the Swiss Federal Institute of Technology, Lausanne, the Grenoble Institute of Technology, the École Supérieure D'électricité, France, and the École Polytechnique de Tunisie, Tunisia, and did consulting work for the French Power Transmission Company, RTE. His research has focused on power system planning for enhanced resiliency and sustainability, risk management of complex systems to catastrophic failures, robust estimation and control, nonlinear dynamics, and bifurcation theory. He was a recipient of several awards, including the U.S. National Science Foundation (NSF) Research Initiation Award and the NSF Young Investigation Award. He is the Co-Founder and Co-Editor of the *International Journal of Critical Infrastructure*. He is the Chairman of the IEEE Working Group on State Estimation Algorithms.

Xiao Chen received the Ph.D. degree in applied mathematics from Florida State University, Tallahassee, FL, USA, in 2011. He is a Computational Scientist and a Project Leader with the Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, Livermore, CA, USA. He works primarily on the development and application of advanced computational and statistical methods and techniques to power engineering, reservoir simulation, subsurface engineering, and seismic inversion. His research interests include uncertainty quantification, data assimilation, and machine learning.

Liang Min (SM'12) received the Ph.D. degree in electrical engineering from Texas A&M University in 2007. He was a Senior Project Manager and a Research Scientist with the Electric Power Research Institute. He was the Associate Program Leader of Cyber & Infrastructure Resilience Program and the Group Leader of Energy Delivery and Utilization with Lawrence Livermore National Laboratory, Livermore, CA, USA, from November 2011 to June 2019. He is currently the Managing Director of the Bits & Watts Initiative at Stanford University, Stanford, CA, USA. He works primarily on the simulation and analysis of national critical infrastructure with a particular focus on energy infrastructure. He holds two U.S. patents in the area of measurement-based voltage stability assessment, which has been commercialized and deployed at Entergy's Control Center through DOE's Smart Grid Investment Grant.