Uncertainty Quantification, State, and Parameter Estimation in Power Systems using Polynomial Chaos Based Methods

Yijun Xu

Dissertation submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Electrical Engineering

Lamine Mili. Chair
Saifur Rahman. Co-Chair
Adrian Sandu
Vassilis Kekatos
Guoqiang Yu

December 6, 2018
Falls Church, Virginia

Keywords: Uncertainty Quantification, Dynamic State Estimation, Generalized Polynomial Chaos, Multi-Element Polynomial Chaos, ANOVA, Polynomial-Chaos-Based Kalman Filter, Response Surface, Bayesian Inference, Markov Chain Monte Carlo.

Uncertainty Quantification, State, and Parameter Estimation in Power Systems using Polynomial Chaos Based Methods

Yijun Xu

(ABSTRACT)

It is a well-known fact that a power system contains many sources of uncertainties. These uncertainties coming from the loads, the renewables, the model and the measurement, etc, are influencing the steady state and dynamic response of the power system. Facing this problem, traditional methods, such as the Monte Carlo method and the Perturbation method, are either too time consuming or suffering from the strong nonlinearity in the system.

To solve these, this Dissertation will mainly focus on developing the polynomial chaos based method to replace the traditional ones. Using it, the uncertainties from the model and the measurement are propagated through the polynomial chaos bases at a set of collocation points. The approximated polynomial chaos coefficients contain the statistical information. The method can greatly accelerate the calculation efficiency while not losing the accuracy, even when the system is highly stressed.

In this dissertation, both the forward problem and the inverse problem of uncertainty quantification will be discussed. The forward problems will include the probabilistic power flow problem and statistical power system dynamic simulations. The generalized polynomial chaos method, the adaptive polynomial chaos-ANOVA method and the multi-element polynomial chaos method will be introduced and compared. The case studies show that the proposed methods have great performances in the statistical analysis of the large-scale power systems. The inverse problems will include the state and parameter estimation problem. A novel polynomial-chaos-based Kalman filter will be proposed. The comparison studies with other traditional Kalman filter demonstrate the good performances of the proposed Kalman filter. We further explored the area dynamic parameter estimation problem under the Bayesian inference framework. The polynomial-chaos-expansions are treated as the response surface of the full dynamic solver. Combing with hybrid Markov chain Monte Carlo method, the proposed method yields very high estimation accuracy while greatly reducing the computing time.

For both the forward problem and the inverse problems, the polynomial chaos based methods haven shown great advantages over the traditional methods. These computational techniques can improve the efficiency and accuracy in power system planning, guarantee the rationality and reliability in power system operations, and, finally, speed up the power system dynamic security assessment.

Uncertainty Quantification, State, and Parameter Estimation in Power Systems using Polynomial Chaos Based Methods

Yijun Xu

(GENERAL AUDIENCEABSTRACT)

It is a well-known fact that a power system state is inherently stochastic. Sources of stochasticity include load random variations, renewable energy intermittencies, and random outages of generating units, lines, and transformers, to cite a few. These stochasticities translate into uncertainties in the models that are assumed to describe the steady-sate and dynamic behavior of a power system. Now, these models are themselves approximate since they are based on some assumptions that are typically violated in practice. Therefore, it does not come as a surprise if recent research activities in power systems are focusing on how to cope with uncertainties when dealing with power system planning, monitoring and control.

This Dissertation is developing polynomial-chaos-based method in quantifying, and managing these uncertainties. Three major topics, including uncertainty quantification, state estimation and parameter estimation are discussed. The developed method can improve the efficiency and accuracy in power system planning, guarantee the rationality and reliability in power system operations in dealing with the uncertainties, and, finally, enhancing the resilience of the power systems.

Acknowledgments

Here, I would like to give my sincere thanks to all who have helped me in my three and half years Ph.D studies.

First and foremost, I would like to express my greatest gratitude to my adviser, Dr. Lamine Mili. Besides his patient guidance on my research and his consistent support to my career, Dr. Mili has entirely reshaped my research vision: he has taught me to be an independent researcher. His passion for research and life would also inspire me in my future life.

I am thankful to all the members of my Ph.D. committee, namely Dr. Saifur Rahman, Dr. Adrian Sandu, Dr. Vassilis Kekatos, and Dr. Guoqiang Yu, for their encouraging and constructive feedback. I would like to give special thanks to Dr. Adrian Sandu, who brings me into the world of polynomial chaos, the fantastic area that I explored in my Ph.D studies.

Next, I would like to thank my dear friends at Northern Virginia Center, Dr. Junbo Zhao and Marcos Netto, who gave me lots of help in research and life. I would also like to thank my friends in Blacksburg, Dr. Duotong Yang, Chen Wang and Xiawen Li. Even one year is short, the memory is precious.

I would also like to show my sincere thanks to Dr. Xiao Chen, Dr. Can Huang and Dr. Mert Korkali for offering me the summer internship opportunities at Lawrence Livermore National Lab and giving me a wonderful summer in California.

Here, I would like to thank my girlfriend for her kindness, accompany, encouragement and consistent support through these years.

Finally and more deeply, I would like to delicate my dissertation to my parents in China. Your unconditional love and complete support carry me along the way. Thank you with all my heart and soul.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Motivation

1.1 Motivation

It is a well-know fact that a power system is inherently stochastic. Sources of stochasticity include load random variations, renewable energy intermittencies, and random outages of generating units, lines, and transformers, to cite a few. These stochasticities translate into uncertainties in the models that are assumed to describe the steady-sate and dynamic behavior of a power system. Now, these models are themselves approximate since they are based on some assumptions that are typically violated in practice. Therefore, it does not come as a surprise if recent research activities in power systems are focusing on how to cope with uncertainties when dealing with power system planning, monitoring and control. Examples include the development of new tools for probabilistic power flow calculations [1], probabilistic small-signal stability analysis [2], and stochastic economic dispatch, [3], among others.

To address these challenges, an accurate and fast method to conduct uncertainty quantification (UQ) is a prerequisite. Currently, two traditional approaches have been utilized to account for model uncertainties in power systems, namely, Monte Carlo (MC) simulations and the Perturbation Method [4]. MC simulations apply a straightforward procedure that consists in randomly sampling the assumed probability distribution functions of the input random variables of a power system model. When the pseudo-random generator being used is not biased, the sample means and variances of the power system model outputs provided by MC simulations tend to the true values, asymptotically. It turns out that in practice, tens of thousands of MC simulations are required to achieve sufficient accuracy in the results, which typically require prohibitive computational times for power systems with detailed generator models , especially when dynamical simulations are involved [5].To cut down on these computing times, the use of the perturbation method has been suggested in [6]. In this method, the random components of the input variables are assumed to be small perturbations around

stable equilibria, which allows us to carry out a linearization of the power system model and to perform small-signal stability analysis. Obviously, this assumption may not hold true in practice since it may produce large errors in the presence of large system disturbances.

To overcome the shortcomings of these foregoing methods, a generalized Polynomial Chaos (gPC)-based method has been proposed by Xiu [7]. In this part, the uncertainties are treated as random variables following known probability distributions. The derived orthogonal polynomials can be weighted and summed to represent the stochastic output of the system, whose sample mean and sample variance can be directly obtained from the gPC weights. Then, the sample standard deviation is used as an index for quantifying the uncertainties in the stochastic outputs. This method has two main advantages. Compared with the MC method, the calculation efficiency is significantly improved, typically ranging from two to three magnitude. Compared with the perturbation method, the gPC-based method can be applied to systems with strong non-linearities, and thereby avoiding to conduct the linearization for the system model. Due to the above advantages, this method has been successfully applied to fluid dynamics [8] and to vehicle dynamics modeling [9], to cite a few. However, the gPC-based method suffers from the following drawbacks. It assumes that the uncertainties follow standard distributions such as the Gaussian distribution, or the beta distribution, or the uniform distribution, to cite a few. But these distributions may provide poor models for the uncertainties found in power systems. Even though the calculation speed is greatly improved when compared to the MC method, the increase of the number of the input random variables significantly affect the calculation efficiency of the method. because in power systems, the number of uncertain inputs is extremely large, the gPC method suffer from the "curse of dimensionality". Thirdly, the gPC method becomes very inaccurate for problems involving random space discontinuities and for long-term dynamical system analysis, which can not be ignored in power system dynamic security analysis.

To address the foregoing problems, this dissertation will introducing and developing several polynomial chaos based methods. These polynomial chaos based methods will be involved in several applications in the power systems, both steady state problems and dynamic problems are included.

1.2 Uncertainty Quantification

UQ is the science of the characterization and reduction of uncertainties present in computational and real world problems. It tries to determine how likely certain outcomes are if some aspects of the system are not exactly known. It has been widely used in many areas, such as prediction, sensitivity analysis, calibration and history matching, optimization and design [4].

1.2.1 Forward and Inverse Problems

In [64], UQ has been classified into two essential problems: the forward problem and the inverse problem. They are defined as:

The Forward Problem: The uncertainty in the inputs is propagated through the simulator to the outputs to learn about the input-output relationship of the physical process;

The Inverse Problem: Observations of the physical process are used to constrain the simulator to match reality, under the presence of uncertainty. This inverse problem requires iteratively solving many forward problems. The developed fast and accurate forward solvers can boost the solving of the inverse problem.

In this dissertation, we will focus on developing fast solvers for both forward problems and inverse problems. Chapter 3 is the forward steady state problem and Chapter 4 is the forward dynamic problem. Chapter 5 is the inverse problem about dynamic state estimation and Chapter 6, 7 are the inverse problems about dynamic parameter estimation.

1.2.2 Uncertainties in the Power systems

The power system model, estimation and validation are subjected to many uncertainties. Ignoring these uncertainties may lead to poorly decision making in power system dynamic security assessment, operation and planning [5], [10], [11]. In [12], the sources of uncertainty are classified into six types. Here, we explain them in the power system model:

(1) *Parameter Uncertainty*: All parameter values are uncertain to some extent. The load, the renewable energy generations, the parameters for the transmission and generation models are not exactly known to power engineers [11].

(2) *Model Discrepancy*: No model is perfect. The simplified models for generators and loads, the π equivalence circuits for transmission lines, etc, are all based on a assumptions that are approximated for the convenience, the simplicity and the lack of knowledge in actual process involved in power systems.

(3) *Residual Variability*: In reality, the true states vary every time the state estimator algorithms executes. This variation may be due to the fact that the process is actually stochastic, or because the set of inadequacy of the measure inputs and the types of the estimator.

(4) *Parametric Variability*: When there is not enough information to specify the inputs, the state estimator output depends on some uncontrolled and unspecified conditions. This induces an extra source of uncertainty into the predicted output which is known as parametric variability.

(5) *Observational Error*: The measurements in power system cannot be fully trusted. Both PMU and SCADA contain noise or even large gross errors.

(6) *Code Uncertainty*: Due to the complexity of the power system dynamic simulations and the time it takes to run, it is not practical to run the simulations at every available input configuration. This indeed brings the code uncertainties.

Due to the large amount of uncertainties in power system, the report by the National Academy of Engineering and Science members claimed that it is important to carry out uncertainty quantification as it would improve the confidence associated with an answer [10]. That is why we are focusing on developing the polynomial chaos based uncertainty quantification method to account these uncertainties in the power system modelling, planning, operation and control.

1.3 Contributions

This thesis focuses on the development of efficient forward solvers and inverse solver to accelerate the uncertainty quantification in the power systems. The novel contributions include three parts.

Probabilistic Power Flow

In this part, we develop an ANOVA-based algorithm that adaptively selects the most effective polynomial chaos basis functions to solve probabilistic power flow problems. Specifically, because of geographical distances, topology features and environmental factors of power systems, the inputs variables of the model can be separated into groups of coupled subsets, which in turn translate into corresponding groups of coupled output variables. The latter can be reliably identified by the ANOVA decomposition, which in turn allows us to greatly reduce their number, thereby alleviating the curse of dimensionality of the gPC method. Furthermore, the equivalence between the polynomial chaos expansion and the ANOVA decomposition is exploited to conduct a variance analysis based on the polynomial chaos expansion without additional time cost.

Statistical Power System Dynamic Simulations

The main contributions of this part include the following points. The gPC based method for the complex power system dynamic simulation with detailed generator model is presented for the first time. Both its strength in short-time simulation and inadequacy in long-time simulations are demonstrated. Second, to solve this issue, a MEgPC method is first time

introduced and implemented. It shows good accuracy and calculation efficiency for the long-time simulations.

Polynomial-Chaos-Based Kalman Filter

This part proposes a new framework for the polynomial-chaos-based Kalman filter (PCKF). By analysis of the variance, the "curse of dimensionality" is greatly overcome. Compared by traditional generalized-polynomial-chaos-based extended Kalman filter, and Ensemble Kalman filter, the proposed PCKF can achieve very high estimation accuracy at much less computing time.

Response-Surface-Based Bayesian Inference for Power System Dynamic Parameter Estimation

This part proposes a novel method for dynamic parameter estimation. Under the Bayesian inference framework, the Markov Chain Monte Carlo method is adopted to provide the full descriptions of the posterior distributions of the parameters estimated. The Maximum-a-posterior estimator is further used to provide the estimated parameter values. The time-consuming dynamic solver is accelerated by polynomial-chaos-based response-surface. The proposed method achieves significant improvement in the computing time while maintaining a high accuracy.

Real-Time Dynamic Parameter Estimation using Polynomial-Chaos-Expansion-Based Bayesian Inference with Hybrid MCMC

This part is following the work above. Facing the challenges that traditional Metropolis-Hastings algorithm cannot overcome strong correlations between parameters and show a slow convergence rate when the dimension increases or the prior is bad, we proposed a hybrid MCMC algorithm to replace the Metropolis-Hastings algorithm. The proposed algorithm combines the advantages of the Langevin MCMC algorithm and the adaptive MCMC algorithm. The Langevin MCMC can accelerate the convergence rate of the Markov Chain in dealing with bad priors and the adaptive MCMC can overcomes the strong correlations between parameters. This hybrid MCMC method can greatly reduce sample size that required by the Metropolis-Hastings algorithm.

1.4 Publications

Journal

Please refer to [13, 15, 16]

- Y. Xu, L. Mili *et al.* "Propagating uncertainty in power system dynamic simulations using polynomial chaos ", *IEEE Trans. Power Syst.*, 2019. (Accepted)
- Y. Xu, L. Mili and J. Zhao, "A novel polynomial-chaos-based Kalman filter ", *IEEE Signal Processing Letters*, 2019. (Accepted)
- Y. Xu, C. Huang, X. Chen, L. Mili *et al.* "Response-surface-based bayesian inference for power system dynamic parameter estimation using PMU measurement", *IEEE Trans. Smart Grid.*, 2019. (Accepted)
- 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.*, 2018. (2nd round review)
- Y. Xu, L. Mili *et al.* "Real-Time dynamic parameter estimation using polynomial-chaos-expansion-based Bayesian inference with hybrid MCMC", *IEEE Trans. Power Syst.*, 2018. (Under Review)

Conference

- Y. Xu, and L. Mili, "Probabilistic power flow analysis based on the adaptive polynomial chaos-ANOVA method ", *Power and Energy Society General Meeting (PESGM)*, 2018. (Accepted)
- Y. Xu, X. Chen, L. Mili *et al.*," Polynomial-chaos-based decentralized dynamic parameter estimation using langevin MCMC", *Power and Energy Society General Meeting (PESGM)*, 2019. (Under Review)

1.5 Report Outline

This thesis is organized as follows:

- Chapter 2 The generalized polynomial chaos theory is introduced.
- Chapter 3 The generalized-ANOVA method in solving probabilistic power flow and variance analysis is proposed.

- Chapter 4 The generalized polynomial chaos and multi-element polynomial chaos method in solving the statistical power system dynamic simulations are demonstrated.
- Chapter 5 A novel polynomial chaos based Kalman filter is proposed.
- Chapter 6 The Response-Surface-Based Bayesian Inference for Power System Dynamic Parameter Estimation is proposed.
- Chapter 7 Real-Time dynamic parameter estimation using polynomial-chaos-expansion-based Bayesian inference with hybrid MCMC is presented.
- Chapter 8 Future work is listed in this chapter.
- Chapter 9 Conclusion is given in this chapter.

Chapter 2

Polynomial Chaos Theory

2.1 Theoretical Background

2.1.1 The Generalized Polynomial Chaos

In the gPC 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 z denote an output, which may be a bus voltage magnitude or a bus voltage angle or any other variable. Let $\boldsymbol{\xi} = [\xi_1, \xi_2, ...\xi_N]$ be a vector of random variables following a known probability distribution, to which is associated a unique orthogonal polynomial as shown in Table 2.1. For the non-standard probability distributions or a standard probability distribution not on the list, such as the Weibull distribution, the Stieltjes procedure is proposed in this part to generate the associated orthogonal polynomial functions. Let $\phi_i(\xi_1, \xi_2, ...\xi_N)$ denote this procedure's corresponding polynomial chaos basis and let a_i denote the ith polynomial chaos coefficient. Formally, we have

$$z = \sum_{i=0}^{\infty} a_i \phi_i(\boldsymbol{\xi}). \tag{2.1}$$

In practice, a truncated expansion is used such that

$$z = \sum_{i=0}^{N_P} a_i \phi_i(\boldsymbol{\xi}), \tag{2.2}$$

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. Note that for a steady-state problem, such as the PPF calculation, a relatively low maximum polynomial chaos order, typically 2, is found to provide output results with enough accuracy [17], [18],

[19], [20]. Here, the N-variate gPC basis functions is constructed as the tensor product of the one-dimensional polynomial chaos basis associated with each input random variable of total degree less than or equal to P, that is:

$$\phi(\boldsymbol{\xi}) = \phi(\xi_1) \times \phi(\xi_2) \otimes \cdots \otimes \phi(\xi_N), \tag{2.3}$$

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

From the polynomial chaos coefficients, the mean, μ, and the variance, σ^2, of the output z can be determined as follows:

$$\mu = a_0, \tag{2.4}$$

$$\sigma^2 = \sum_{i=1}^{N_P} a_i^2 E[\phi_i^2], \tag{2.5}$$

where $E[.]$ is the expectation operator.

Table 2.1: UNIVARIATE GPC POLYNOMIAL BASIS OF DIFFERENT RANDOM VARIABLES

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

The Orthogonal Polynomial Chaos Basis

A set of one-dimensional polynomial chaos basis functions given by $\{\phi_i(\xi), i = 0, 1, 2, 3....\}$ with respect to some real positive measure satisfy the following relations [21]:

$$\int \phi_r(\xi)\phi_s(\xi)d\lambda = \gamma_r \delta_{rs}. \tag{2.6}$$

Here, λ is a probability measure defined as the Cumulative Probability Distribution Function (CPDF) of ξ, for which, the associated orthogonal polynomials are unique; δ_{rs} is the Kronecker delta function, that is, $\delta_{rs} = 0$ if $s \neq r$ and $\delta_{rs} = 1$ if $s = r$; γ_r is a positive constant termed the normalization constant for the rth order polynomial chaos basis. Obviously,

$$\gamma_r = \int \phi_r^2(\xi)d\lambda \tag{2.7}$$

If $\gamma_r = 1$, the polynomial chaos basis is orthonormal. By defining

$$\widetilde{Q_r} = Q_r/\sqrt{\gamma_r}, \tag{2.8}$$

the $\widetilde{Q_r}$ is orthonormal.

Similarly, any set of multi-dimensional polynomial chaos basis functions given by $\{\phi_i(\boldsymbol{\xi}), i = 1, 2, 3....\}$, are orthogonal to each other with respect to their joint probability measure [4].

Collocation Points

collocation points can be regarded as a finite sample of $\boldsymbol{\xi} = [\xi_1, \xi_2, ...\xi_N]$ 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 [4, 17]. 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.

Three-Term Recurrence Relation

The orthogonal polynomials satisfy a three-term recurrence relation given by [21]

$$\begin{aligned} \phi_{k+1}(\xi) &= (\xi - \alpha_k)\phi_k(\xi) - \beta_k\phi_{k-1}(\xi), \\ \phi_{-1} &= 0, \quad \phi_0 = 1; \quad k = 0, 1, 2 \ldots, K, \end{aligned} \tag{2.9}$$

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

$$\phi_k(\xi) = \xi^k + lower\ degree\ terms, \quad k = 0, 1, \ldots, K, \tag{2.10}$$

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.

For the standard normal distribution, its corresponding orthogonal polynomials are the Hermite polynomials, for which $\alpha_k = 0$ and $\beta_k = k$. After orthonormalization, the orthonormal Hermite polynomials become:

$$He_0 = 1, \quad He_1 = \xi, \quad He_2 = (\xi^2 - 1)/\sqrt{2} \quad \ldots \tag{2.11}$$

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 part, the Stieltjes

procedure is chosen as an accurate and cost-effective method. It is given by

$$\alpha_k = \frac{\int_{\mathbb{R}} \xi \phi_k^2(\xi) d\lambda(\xi)}{\int_{\mathbb{R}} \phi_k^2(\xi) d\lambda(\xi)}, \quad k = 0, 1, 2, \ldots, K, \tag{2.12}$$

$$\beta_k = \frac{\int_{\mathbb{R}} \phi_k^2(\xi) d\lambda(\xi)}{\int_{\mathbb{R}} \phi_{k-1}^2(\xi) d\lambda(\xi)}, \quad k = 1, 2, \ldots, K. \tag{2.13}$$

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 (4.15), (4.16) become summations. The Stieltjes procedure ensures the orthogonality of the basis functions with good precision when the polynomial chaos order K is not large. However, when K approaches a large n, such as 100, the coefficients may be calculated with poor accuracy. However, for the PPF problem, only a polynomial chaos basis with low order is used, implying that the Stieltjes procedure is very accurate in that case [21].

Chapter 3

Probabilistic Power Flow and Variance Analysis Based on Hierarchical Adaptive Polynomial Chaos-ANOVA method

3.1 Introduction

It is a well-known fact that a power system is inherently stochastic. Sources of stochasticity include load continuous variations with time, renewable energy intermittencies, and random outages of generating units, lines, and transformers, to cite a few [17]. These stochasticities translate into uncertainties in the dynamical responses of a power system. To address this problem, research activities have focused on uncertainty modeling in power system planning, monitoring and control. Examples include the development of new tools for probabilistic small-signal stability analysis [2], probabilistic load margins formulation [19], stochastic economic dispatch [3], and uncertainty quantification for power system dynamics [6].

This chapter investigates the probabilistic power flow problem [1], which quantifies the uncertainties induced in the bus voltages and line power flows by the uncertainties of the loads, power generations, and network parameters. Two traditional approaches have been utilized to address this problem, namely, Monte Carlo (MC) simulations and the analytical Method. MC simulations apply a straightforward procedure that consists in randomly sampling the assumed probability distribution functions of the input random variables of a power system model and executing a power flow calculation for each sample. When the pseudo-random generator being used is not biased and the assumed probability distributions are exact, the power system model outputs provided by MC simulations tend to the true values, asymptotically. It turns out that in practice, tens of thousands of MC simulations are required

to achieve sufficiently accurate results, which is too time-consuming for large- scale power systems.

Instead of using the time-consuming and heuristic sampling-based MC simulations, computational efficient analytic methods that derive the statistical relationship between the input random variables and the outputs explicitly have been proposed, such as the Cornish-Fisher, Edgeworth and Gram-Charlier-based ones. However, Cornish-Fisher and Edgeworth expansions have the problems of bad tail behavior while the Gram-Charlier series may yield negative cumulative probability values [17], [22], which is impractical. To address this issue, Williams and Curran proposed to use the Maximum Entropy and the Gram-Charlier expansion together to construct PDFs based on cumulant arithmetic treatment of the linearized power flow [23]. Although this is a cost-effective method, linearization assumption on the power flow solutions might not always hold for true in practice, especially under heavily loaded operation conditions.

To overcome the shortcomings of the foregoing methods, a generalized-polynomial-chaos (gPC)-based power flow method has been advocated by [17], [18] and [24]. Introduced by Wiener [25] and further developed by Xiu and Karniadakis [7], this method assumes that the random inputs of a system follow a set of standard probability distributions while the outputs of interest are expressed as a weighted sum of orthogonal polynomial chaos basis. This method has been shown to be accurate and efficient when the number of the input variables are not large. In addition, it suffers from the "curse of dimensionality" and can only be applied to systems with input random variables following the limited types of probability distributions. To mitigate that issue, the sparse polynomial chaos expansion (SPCE) method is advocated [24, 26]. Although SPCE reduces the number of terms for polynomial chaos expansion (PCE), it is unable to do so for the deterministic calculations required to construct the PCE surrogate. In fact, the latter turns out to be the most time-consuming procedure for large-scale power system PPF problem. As a result, the capability of SPCE for computational efficiency improvement is quite limited.

To deal with the "curse of dimensionality" while maintaining high accuracy, we propose to integrate a hierarchical polynomial chaos with the Analysis of Variance (ANOVA) decomposition together with the Stieltjes procedure. Introduced by Fisher [27] and further extended by Hoeffing [28], the ANOVA decomposition has been shown to be an effective method for dimensionality reduction of large-scale problems [29–31]. In this part, we develop an ANOVA-based algorithm that adaptively selects the most effective polynomial chaos basis functions to solve probabilistic power flow problems. Specifically, because of geographical distances, topology features and environmental factors of power systems, the inputs variables of the model can be separated into groups of coupled subsets, which in turn translate into corresponding groups of coupled output variables. The latter can be reliably identified by the ANOVA decomposition, which allows us to greatly reduce their number, thereby alleviating the curse of dimensionality of the gPC method. Furthermore, the equivalence between the polynomial chaos expansion and the ANOVA decomposition is exploited to conduct a variance analysis based on the polynomial chaos expansion without additional time cost.

Furthermore, the Stieltjes procedure allows the proposed method to be applied to random inputs following arbitrary probability distributions [32]. Finally, renewable energy and load injections typically follow time-series patterns, indicating temporal correlations. They may also show temporal correlations in the same geographic locations. To this end, we advocate the approach shown in [17] to deal with the correlations between the input random variables. Extensive simulation results carried out on the IEEE 118-bus system and the 1354-bus European high voltage system with correlated renewable energy generations reveal that the developed method outperforms the generalized polynomial chaos method and the Monte Carlo method.

3.2 Theoretical Background

3.2.1 ANOVA Decomposition

The ANOVA expansion represents a function f with N random variable in the form of [31], [29], [33]:

$$\begin{aligned} f(\xi_1, \ldots, \xi_N) = f_0 + \sum_{1 \le j_1 \le N} f_{j_1}(\xi_{j_1}) \\ + \sum_{1 \le j_1 < j_2 \le N} f_{j_1, j_2}(\xi_{j_1}, \xi_{j_2}) + \ldots + f_{1,2,\ldots,N}. \end{aligned} \tag{3.1}$$

The functions $f_{j_k}(\xi_{j_k})$, $1 \le j_k \le N$, $f_{j_k, j_l}(\xi_{j_k}, \xi_{j_l})$, $1 \le j_k \le j_l \le N$, etc., are called the first-order, second-order,... ANOVA components, respectively. In general, a function f is an ANOVA decomposition if the following three properties are satisfied:

1) The constant term is the mean of function, i.e.,

$$f_0 = \int_{\Gamma^N} f(\boldsymbol{\xi}) d\lambda(\boldsymbol{\xi}), \tag{3.2}$$

implying that all higher-order components have zero mean, that is,

$$\int_{\Gamma^N} f_{j_1, \ldots, j_s} d\lambda(\xi_{j_1}, \xi_{j_2}, \ldots, \xi_{j_N}) = 0, \quad 1 \le s \le N. \tag{3.3}$$

2) The other important property of ANOVA expansion is the orthogonality among its terms, which is expressed as

$$\int_{\Gamma^N} f_{j_1, \ldots, j_s} f_{k_1, \ldots, k_l} d\lambda(\xi_{j_1}, \xi_{j_2}, \ldots, \xi_{j_N}) = 0, \tag{3.4}$$

if $(j_1, \ldots, j_s) \neq (k_1, \ldots, k_l)$. This is the direct consequence of (3.3).

3) The variance of f is the sum of the variances of all te component functions, that is,

$$\sigma^2(f) = \sum_{1\leq j_1\leq N} \sigma^2(f_{j_1}) + \sum_{1\leq j_1<j_2\leq N} \sigma^2(f_{j_1,j_2}) + \ldots + \sigma^2(f_{1,2,\ldots,N}). \tag{3.5}$$

Note that (3.5) holds true only when the c.p.d.f used in calculating the variance is the same as that in the ANOVA decomposition given by (8.4).

The terms in ANOVA decomposition are computed as follows:

$$f_S = \int_{\Gamma^{N-|S|}} f(\boldsymbol{\xi}) d\lambda(\boldsymbol{\xi}_{S^c}) - \sum_{T\subset S} f_T(\boldsymbol{\xi}_T), \tag{3.6}$$

where $S = \{j_1, j_2, \ldots, j_s\}$, $|S|$ is the number of elements in S, C is the compliment of S, T is a subset of S and $f_T = f_{j_1,j_2,\ldots,j_t}(\xi_{j_1}, \xi_{j_2}, \ldots, \xi_{j_t})$. Note that the computation of f_S requires first computing the lower-order components f_S. So f_0 is computed via equation (3.2) first. Then the higher-order component functions are computed by equation (3.6) recursively. Let $T_{j_1,\ldots,j_s}$ denote the fraction of the variance $\sigma^2(f)$ that is contributed by $f_{j_1,\ldots,j_s}$, which is defined as

$$T_{j_1,\ldots,j_s} = \frac{\sigma^2(f_{j_1,\ldots,j_s})}{\sigma^2(f)}. \tag{3.7}$$

Here T_{j_1} denotes the variance contributed by the single variable ξ_{j_1}, T_{j_1,j_2} denotes the variance contributed by the coupling effect of the two random inputs, (ξ_{j_1}, ξ_{j_2}), and so on for the higher-order terms. $T_{j_1,\ldots,j_s}$ satisfies the relation given by

$$\sum_{i=1}^{N} T_{j_1} + \sum_{1\leq j_1<j_2\leq N} T_{j_1,j_2} + \ldots T_{j_1,\ldots,j_N} = 1. \tag{3.8}$$

3.2.2 Relation between the Polynomial Chaos Expansion and the ANOVA Decomposition

The polynomial chaos expansion is a function of N random variables, which can be reorganized as

$$z = a_0 + \sum_{i=1}^{N} a_i\phi(\xi_i) + \sum_{i=1}^{N}\sum_{j=1}^{i} a_{i,j}\phi(\xi_i, \xi_j) + \sum_{i=1}^{N}\sum_{j=1}^{i}\sum_{r=1}^{j} a_{i,j,r}\phi(\xi_i, \xi_j, \xi_r) + \ldots. \tag{3.9}$$

This equation can be further reorganized as a truncated form of the ANOVA expansion as follows. As an example, let us consider the PPF calculation where $P = 2$. The relationship (3.9) can be truncated and reorganized as

$$\begin{aligned} z = a_0 + (\sum_{i=1}^{N} a_i \phi(\xi_i) + \sum_{i=1}^{N} \sum_{j=i} a_{i,j} \phi(\xi_i, \xi_j)) \\ + \sum_{i=1}^{N} \sum_{1 \leq j < i} a_{i,j} \phi(\xi_i, \xi_j). \end{aligned} \tag{3.10}$$

Here, the first two summations and the third summations are the first-order and the second-order ANOVA components, respectively. The "curse of dimensionality" comes from the third summation, which has a number of terms equal to $\frac{N!}{(N-2)!(2)!}$. It is observed that the PC expansion is in the same form as an ANOVA expansion since the three relationships given by (3.2), (3.3), (3.4), (3.5) are satisfied. Indeed, we have

1) The mean of PC expansion is a_0, a constant term;

2) The relationships (3.3) and (3.4) hold true because of the orthogonality property of the PC basis;

3) Let $z_i = a_i \phi_i(\boldsymbol{\xi})$. Its variance is expressed as

$$\sigma^2(z_i) = \int_{\Gamma^N} (a_i \phi_i(\boldsymbol{\xi}) - \mu_i)^2 d\lambda(\boldsymbol{\xi}), \tag{3.11}$$

where μ_i is the mean of z_i expressed as

$$\mu_i = \int_{\Gamma^N} (a_i \phi_i(\boldsymbol{\xi})) d\lambda(\boldsymbol{\xi}) = a_i \int_{\Gamma^N} \phi_i(\boldsymbol{\xi}) d\lambda(\boldsymbol{\xi}). \tag{3.12}$$

Since $\int_{\Gamma^N} \phi_i(\boldsymbol{\xi}) d\lambda(\boldsymbol{\xi}) = 0$, z_i has zeros mean, implying that (5.12) is given by

$$\sigma^2(z_i) = \int_{\Gamma^N} (a_i \phi_i(\boldsymbol{\xi}))^2 d\lambda(\boldsymbol{\xi}) = a_i^2 E[\phi_i^2]. \tag{3.13}$$

From (2.5) and (5.14), we infer (3.5).

From the above derivations, we conclude that the PC expansion with Pth-order has the same form as a truncated form of an ANOVA expansion. Therefore, the terms in the PC expansion can be viewed as ANOVA components.

3.3 The Proposed gPC-ANOVA method

In [17] and [18], it is shown that the coupled PC basis functions for the input random variables are neglected in the 1st-order gPC expansion, which may result in inaccurate system outputs, while they are included in the 2nd-order gPC expansion. Interestingly, it is found that in large-scale power systems, only limited coupling effects between the input random variables are important. The latter are identified by the ANOVA decomposition as the important random inputs. When setting a criterion, only the 2nd-order basis functions will be selected and the higher order basis functions will be ignored. Hence, the calculation efficiency will be significantly improved while the accuracy barely deteriorates. The detailed steps for the gPC-ANOVA method that we propose is described next.

Normalizing the Random Variables

We observe from Table 2.1 that the ranges of the random variables may be different. To make the gPC expansion compatible with the ANOVA decomposition, we need to map all the random variables into new random variables defined on $[-1, 1]^N$. To this end, every random variable is first rescaled via the following linear transformation:

$$\xi^i = \frac{b_i - a_i}{2} Y_i + \frac{b_i + a_i}{2}, \tag{3.14}$$

where a_i, b_i are the lower and upper bounds of the ith random variable, and Y_i is the ith new random variable defined on $[-1, 1]$. The PDF of the Y_i is then obtained by

$$f(Y_i) = f(\xi(Y_i)) \frac{b_i - a_i}{2}; \tag{3.15}$$

Note that for a Gaussian random variable defined on $(-\infty, +\infty)$, the $[a_i, b_i]$ can be set over the interval [-6, 6], which is associated with a probability of more than 99.9999%.

Approximating the 1st-Order gPC Coefficients

To do that we apply the following procedure:

(a) Construct the polynomial chaos basis using the Stietjies procedure and the relationship (2.3).

(b) Choose appropriate $N+1$ combinations of the collocation points and every combination satisfies $CP_i = [\boldsymbol{\xi_i}]$, where the covariance matrix of the input variables is denoted as C. To handle the correlations among input random variables, the strategy proposed in [17]

is used. Specifically, by applying the Cholesky decomposition to C, we get the lower matrix, L. Then, CP_i is transformed into CP_i^c by

$$CP_i^c = LCP_i. \tag{3.16}$$

(c) Put the collocation points of the polynomial chaos functions into the $((N+1)\times(N+1))$ matrix H_1 to get

$$H_1 = \begin{pmatrix} \phi_0(\boldsymbol{\xi}_0) & \phi_1(\boldsymbol{\xi}_0) & \dots & \phi_N(\boldsymbol{\xi}_0) \\ \phi_0(\boldsymbol{\xi}_1) & \phi_1(\boldsymbol{\xi}_1) & \dots & \phi_N(\boldsymbol{\xi}_1) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_0(\boldsymbol{\xi}_N) & \phi_1(\boldsymbol{\xi}_N) & \dots & \phi_N(\boldsymbol{\xi}_N) \end{pmatrix}; \tag{3.17}$$

(d) Compute the power flow output for the selected collocation points to get the $((N+1)\times 1)$ output vector, Z_1, given by

$$Z_1 = \begin{pmatrix} z_0 & z_1 & \dots & z_N \end{pmatrix}^T; \tag{3.18}$$

(e) Based on the selected collocation points and the model output, calculate the coefficients of the $((N+1)\times 1)$ vector, A_1, given by

$$A_1 = \begin{pmatrix} a_0 & a_1 & \dots & a_N \end{pmatrix}^T = {H_1}^{-1} Z_1; \tag{3.19}$$

Applying the Adaptive ANOVA Procedure

The 1st-order gPC expansion is a function of N random variable expressed as

$$z = a_0 + \sum_{i=1}^{N} a_i \phi(\xi_i) = f(\xi_1, \dots, \xi_N). \tag{3.20}$$

The variance for each 1st-order ANOVA component can be calculated by

$$\sigma^2(f(\xi_i)) = \sigma^2(a_i\phi(\xi_i)) = a_i^2 E[\phi_i^2]. \tag{3.21}$$

Here, we assume that $\sigma^2(f(\xi_i, \xi_j))$ is related to the variance of the $1st$-order ANOVA component $\sigma^2(f(\xi_i))$ given by

$$\sigma^2(f(\xi_i, \xi_j)) \approx \eta\sigma^2(f(\xi_i))\sigma^2(f(\xi_j)) = \eta a_i^2 a_j^2 E[\phi_i^2]E[\phi_j^2], \tag{3.22}$$

where η is the ratio coefficient [33]. It should be noted that η does not affect the derivation of the proposed approach as it will be canceled out in equation (3.23) The relationship (3.22) is based on the following assumption [31]: a dimension pair (i, j) is likely to have a larger coupling effect (measured by $\sigma^2(f(\xi_i, \xi_j))$) compared with other pairs if the dimensions i and j are likely to have large individual effects (measured by $\sigma^2(f(\xi_i))$, $\sigma^2(f(\xi_j))$), and vice versa. Here, the steps to select the important basis functions are as follows:

(i) Estimate all the $\sigma^2(f(\xi_i, \xi_j))$ by (3.22) and calculate the sum of them as $\sum \sigma^2(f(\xi_i, \xi_j))$;

(ii) Sort $\sigma^2(f(\xi_i, \xi_j))$ in a descending order and put them in a set U. Set U_i as the ith element in the U;

(iii) Defining θ as a variance ratio indicator given by

$$\theta = \frac{\sum_{i=1}^{M} U_i}{\sum \sigma^2(f(\xi_i, \xi_j)}, \tag{3.23}$$

find the smallest positive integer for M that makes θ larger than a given threshold ϵ, *e.g.*, 0.99. Mark these M pairs as the effective basis functions, which are chosen for the 2nd-order gPC calculation.

Computing Selected 2nd-Order gPC Basis Functions

Using the M important basis functions estimated above, M additional deterministic power flow calculations are executed to approximate the new gPC coefficients. The main steps are summarized as follows:

(1) Choose M additional collocation points and transform them into correlated ones;

(2) Put (N+1+M) collocation points into the $((N+1+M) \times (N+1+M))$-matrix H given by

$$H = \begin{pmatrix} \phi_0(\boldsymbol{\xi}_0) & \dots & \phi_{N+M}(\boldsymbol{\xi}_0) \\ \vdots & \ddots & \vdots \\ \phi_0(\boldsymbol{\xi}_{N+M}) & \dots & \phi_{N+M}(\boldsymbol{\xi}_{N+M}) \end{pmatrix}; \tag{3.24}$$

(3) Compute the system power flows for the M new collocation points to get the $(M \times 1)$-output vector Z_2 given by

$$Z_2 = \begin{pmatrix} z_{N+1} & z_{N+2} & \dots & z_{N+M} \end{pmatrix}^T; \tag{3.25}$$

Combining Z_2 with Z_1, we get

$$Z = \begin{pmatrix} z_0 & \dots & z_N & z_{N+1} & \dots & z_{N+M} \end{pmatrix}^T; \tag{3.26}$$

(4) Calculate the $((N+1+M) \times 1)$-vector A defined as

$$A = \begin{pmatrix} a_0 & \dots & a_N & \dots & a_{N+M} \end{pmatrix}^T = H^{-1}Z; \tag{3.27}$$

(5) Represent the power flow responses z by the polynomial chaos expansions with (N+1+M) terms.

Conducting Variance Analysis

Based on the equivalence relationship between the polynomial chaos expansion and the ANOVA decomposition, we get without additional effort the variance fraction by directly using (8.5) and (7.21) as

$$T_{j_1} = \frac{\sigma^2(f_{j_1)}}{\sigma^2(f)} = \frac{a_i^2 E[\phi(\xi_i)^2] + a_{i,j}^2 E[\phi(\xi_i, \xi_j)^2]_{i=j}}{\sigma^2}, \tag{3.28}$$

$$T_{j_1,j_2} = \frac{\sigma^2(f_{j_1,j_2})}{\sigma^2(f)} = \frac{a_{i,j}^2 E[\phi(\xi_i, \xi_j)^2]_{i\neq j}}{\sigma^2}. \tag{3.29}$$

Chapter 4

Propagating Uncertainty in Power System Dynamic Simulations Using Polynomial Chaos

4.1 Introduction

Uncertainty quantification (UQ) of the renewable energy power injections and uncontrollable and elastic loads has become an important research topic [10]. For example, several probabilistic methods have been developed for power system operation and analysis by accounting their uncertainties, such as probabilistic power flow [1], [17], probabilistic small-signal analysis [2] and probabilistic power system dynamic analysis [5]- [11]. The latter is critical to ensure the dynamic security of the next-generation power systems, which is the main scope of this part.

To date, two approaches have been utilized to account for model uncertainties in power systems, namely, the Monte Carlo (MC) simulations [5] and the Perturbation Method [4]. MC simulations typically apply a straightforward procedure to propagate the uncertainties of random inputs through a massive number of samples obtained from its assumed probability distribution function. It is well-known in the literature that when the number of samples tends to very large, the propagated uncertainties tend to be the ground truth, asymptotically. For practical applications, such as power systems, tens of thousands of MC simulations are required to achieve sufficiently accurate results, which typically require prohibitive computational times for power systems with detailed generator models, especially when dynamical simulations are involved [10]. To improve its computational efficiency, the perturbation-based method has been proposed [6, 37]. In this method, the random inputs are assumed to be small perturbations, which permit the linearization of the power system model around a stable equilibrium point. However, this assumption does not hold true in practice and may

produce large errors in the presence of large system disturbances. This is precisely the case for power system dynamic security analysis where unstable cases have to be identified and analyzed.

To address the aforementioned issues, two uncertainty quantification approaches using polynomial-chaos-based methods are proposed and investigated. It yields the following contributions:

- We extend the generalized Polynomial Chaos (gPC) method [7, 25] to quantify the uncertainties of multi-machine power system dynamic simulations. It is able to reduce the computing time by three orders of magnitude compared with Monte Carlo methods while achieving the same accuracy. Although this approach has been used in other areas, such as power electronics and voltage sensor validation [38–40], we believe this is the first time for multi-machine power system dynamics with detailed generator model. It is worth pointing out that the approach using probabilistic collocation method [41] is actually one special case of gPC method and it did not consider the multi-machine power system with multiple random variables. Furthermore, we find that the gPC method is very useful for short-term power system dynamic simulations, but it may produce unreliable results for long-term simulations.

- To address the weakness of the gPC method, we present the Multi-Element generalized Polynomial Chaos(MEgPC) method [32, 42]. It is a self-adaptive method that decomposes the random space into small elements to guarantee the desired accuracy of the statistics of the output variables for the long-term simulations. Note that, the value of MEgPC method has been demonstrated in the areas of the fluid dynamics, the vehicle dynamics modeling and the single-machine system on a ship [8]- [43]. In this part, we extend it to multi-machine power systems and demonstrate its values for long-term dynamic simulations.

- Since the uncertainties of the renewable-based generations and loads can follow very different distributions, we extend the Stieltjes' procedure to address that. Note that this procedure is able to handle any type of probability distributions of the input random variables. It is worth emphasizing that the extensions of gPC and MEgPC to address arbitrary distributions is one of the important contributions of the paper. This distinguishes the approaches proposed in the existing literature that is working under the Wiener-Askey polynomial families [7] or simply assumes uniform distributions for input random variables [43].

- Extensive simulation results show that our proposed approaches are able to produce comparable accuracy as the Monte Carlo-based method while achieving significantly improved computational efficiency for both stable and unstable operation conditions.

4.2 Problem Formulation

Following Sauer and Pai in [44], we formulate the power system dynamic model with uncertainty as:

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{\xi}), \quad \mathbf{0} = \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi}). \tag{4.1}$$

where $\boldsymbol{f}(\cdot)$ and $\boldsymbol{g}(\cdot)$ are vector-valued nonlinear functions; $\boldsymbol{x}$ is the state vector containing the dynamic state variables of the generators; $\boldsymbol{y}$ denotes the algebraic vector containing the set of algebraic variables; $\boldsymbol{u}$ denotes the input vector for the generators and $\boldsymbol{\xi}$ denotes random variable vector whose cumulative distribution functions are equal to the cumulative distribution functions of the power injections of the loads and the renewable energy, which are considered to be random. A set of M samples are drawn from the multivariate probability distribution of $\boldsymbol{\xi}$, yielding $\{\boldsymbol{\xi}^{(j)}\}_{j=1}^{M}$. Then, for each $\boldsymbol{\xi}^{(j)}$, $j = 1, ..., M$,

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{\xi}^{(j)}), \quad \mathbf{0} = \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi}^{(j)}), \tag{4.2}$$

are solved to get M dynamical solutions, from which the sample means and variances of the state and of the algebraic variables of the power system model can be determined. The detailed way mapping of the random inputs to a standard random variable ξ is illustrated in Section IV-A.

4.3 The Generalized Polynomial Chaos Method

The detailed gPC procedure as applied to power system dynamic analysis is described next.

(a) Map the ith random input variable, such as the power injection P_i, to a given random variable ξ_i as follows:

$$P_i = F_i^{-1}(T(\xi_i)), \tag{4.3}$$

where F_i^{-1} is the inverse cumulative probability function of P_i and T is the cumulative probability function of ξ_i.

(b) Construct the polynomial chaos basis, then express the output z in the gPC expansion form of (7.9).

(c) Choose M appropriate combinations of collocation points and put them into the polynomial chaos basis $(M \times N_P)$ matrix H such that

$$H = \begin{pmatrix} \phi_0(\xi_0) & \phi_1(\xi_0) & \dots & \phi_{N_P}(\xi_0) \\ \phi_0(\xi_1) & \phi_1(\xi_1) & \dots & \phi_{N_P}(\xi_1) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_0(\xi_M) & \phi_1(\xi_M) & \dots & \phi_{N_P}(\xi_M) \end{pmatrix}; \tag{4.4}$$

(d) Compute the power system dynamic model output for the selected collocation points to get the $(M \times 1)$ output Z matrix given by

$$Z = \begin{pmatrix} z(t,\xi_0) & z(t,\xi_1) & \dots & z(t,\xi_M) \end{pmatrix}^T; \tag{4.5}$$

(e) Estimate the unknown coefficients A based on the collocation points selected and the model output calculated by

$$A = H^{-1}Z, \tag{4.6}$$

A is the $(N_P \times 1)$ coefficient matrix expressed as

$$A = \begin{pmatrix} a_0(t) & a_1(t) & \dots & a_{N_P}(t) \end{pmatrix}^T; \tag{4.7}$$

(f) Calculate the sample mean and variance of the output variables directly from the polynomial chaos coefficients using (5.6) and (2.5).

In total, the gPC method needs to solve the power system dynamical equations for just $N_P + 1$ samples at chosen collocation points, a much smaller number than the larger number of samples required by the MC method, greatly reducing the computational burden. The algorithm implementing this method is summarized in Algorithm 1.

Algorithm 1 The gPC for a dynamic power systems model

```
Build the power systems dynamic model;
Select the input random variables and the statistical output of the model;
Construct the polynomial chaos basis for the input random variables;
Choose a set of collocation points;
Calculate M solution of the power system dynamical equations;
for time step i = 1 to t do
    Calculate the coefficients of the matrix A by (6.28);
    Calculate the sample mean and the sample variance via (5.6) and (2.5);
end for
Plot the graph of the sample mean of the output variables and indicate the standard
deviations by error bars.
```

4.4 Multi-Element Generalized Polynomial Chaos Method

4.4.1 Multi-Element Generalized Polynomial Chaos Theory

While the gPC can greatly reduce the calculation burden of the MC method, it becomes very inaccurate for long-term dynamical simulations. Increasing the polynomial order only

postpones the error growth. Therefore, a more accurate and robust method called MEgPC is proposed by [42], [32]. In MEgPC, the domains of the input random variables are adaptively decomposed into smaller intervals, whose associated random variables are termed elements. For every element, its corresponding set of orthogonal polynomials is generated by the Stieltjes procedure. The gPC frame is then applied to every element to determine the local sample mean and variance. Finally, the global sample mean and variance are determined. This method greatly reduces the error at each time step, significantly enhancing the accuracy over long-term integration. Furthermore, a relatively low-order polynomial can be used for each element since the degree of perturbation in each element has been reduced.

This adaptive MEgPC method includes the following steps: *1)* decompose the domain of the random input variables; *2)* map the domain over the interval [-1, 1]; and *3)* construct a set of orthogonal polynomials and apply the adaptive procedure to every local element to choose the best decomposition of the domain of the input random variables. A more detailed description of each step is as follows:

Decomposition of the Domain of the Input Random Variables

Let $\boldsymbol{\xi} = [\xi_1, \xi_2, ...\xi_N]$ denote an N-dimensional random input vector, where ξ_i for $i = 1, ..., N$ is a set of independent identically distributed (i.i.d.) random variables. Let B_k denote the domain of the kth element and let B denote the union of all of the B_k. Formally, this is written as

$$\begin{aligned} B_k = [a_1^k, b_1^k) \times [a_2^k, b_2^k) \times \ldots \times [a_N^k, b_N^k], \\ B = \textstyle\bigcup_{k=1}^m B_k, \\ B_\kappa \textstyle\bigcap B_j = \varnothing, \end{aligned} \tag{4.8}$$

where $\kappa = 1, ..., m$; $j = 1, ..., m$, $\kappa \neq j$; $k = 1, 2, ...m$, and a_i^k, b_i^k are the lower and upper bounds of the ith random variable in the kth element $\zeta_i^k (i = 1, 2, \ldots, N$ and $k = 1, 2, \ldots, m)$, respectively. Let $f_i(x)$ denote the PDF of the ith random variable ξ_i. Let I_{B_k} denote the indicator function defined as

$$I_{B_k} = \begin{cases} 1 & \text{if } \boldsymbol{\zeta} \in B_k, \\ 0 & \text{otherwise ;} \end{cases} \tag{4.9}$$

Thus, for the kth random element, its probability is expressed as

$$Pr(I_{B_k}) = \prod_{i=1}^{N} \int_{a_i^k}^{b_i^k} f_i(x) dx; \tag{4.10}$$

Mapping the Local Random Vectors

All of the local random vectors $\boldsymbol{\zeta}^k$ are mapped to a new random vector defined in $[-1,1]^N$. To this end, every local random variable is first rescaled via the following linear transformation:

$$\zeta_k^i = \frac{b_i^k - a_i^k}{2} Y_i^k + \frac{b_i^k + a_i^k}{2}, \tag{4.11}$$

where Y_i^k is the ith new random variable in the kth element defined in $[-1,1]^N$.

The PDF of the Y_i^k is then obtained by truncating the domains of all of the probability distribution functions with infinite range to a finite range that accounts for most of the probabilities. Each is then mapped to the interval $[-1,1]$. For instance, the Gaussian distribution defined over the interval [-6, 6], which is associated with a probability of more than 99.999%, is transformed into a new function over the interval [-1, 1] as follows:

$$f(Y_i^k) = \frac{f(x(Y_i^k))}{Pr(x \in [a_i^k, b_i^k])} \frac{b_i^k - a_i^k}{2}; \tag{4.12}$$

Construction of Orthogonal Polynomials

Once a random variable is transformed as described in the previous step, the PDF in every element is no longer orthogonal to its original polynomial chaos basis. To regain the orthogonality, the construction of a new polynomial chaos basis is necessary, except in the case of the uniform distribution. This is done by using a three-term recurrence relation given by [21]

$$\begin{aligned} \pi_{k+1}(x) &= (x - \alpha_k)\pi_k(x) - \beta_k \pi_{k-1}(x), \\ \pi_{-1} &= 0, \quad \pi_0 = 1; \quad k = 0, 1, 2 \ldots, K, \end{aligned} \tag{4.13}$$

where $\pi_k(x)$ is a set of orthogonal polynomials defined as:

$$\pi_k(x) = x^k + \textit{lower degree terms}, \quad k = 0, 1, \ldots, K, \tag{4.14}$$

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. Different methods exist to calculate α_k and β_k. In this part, the Stieltjes procedure defined as

$$\alpha_k = \frac{\int_{\mathbb{R}} x \pi_k^2(x) d\lambda(x)}{\int_{\mathbb{R}} \pi_k^2(x) d\lambda(x)}, \quad k = 0, 1, 2, \ldots, K, \tag{4.15}$$

$$\beta_k = \frac{\int_{\mathbb{R}} \pi_k^2(x) d\lambda(x)}{\int_{\mathbb{R}} \pi_{k-1}^2(x) d\lambda(x)}, \quad k = 1, 2, \ldots, K, \tag{4.16}$$

is used, where β_0 is arbitrary and can be conveniently chosen as $\beta_0 = \int_{\mathbb{R}} d\lambda(x)$ and K is the highest order of the polynomials. If the measure consists of n discrete points, the integrals

in (4.15), (4.16) become summations. The Stieltjes procedure can ensure orthogonality with good precision when the polynomial chaos order K is not large. However, when K approaches a large n, the coefficients calculated with the Stieltjes procedure may lose their accuracy. Since only a polynomial chaos basis with low order is used here, the Stieltjes procedure is very accurate.

An Adaptive Procedure

Adaptivity must be used to achieve accuracy for a simulation of long duration or the presence of a discontinuity in an input random variable. The adaptive procedure implemented here is based on the relative local errors. For element k, its gPC expansion is defined as

$$y_k = \sum_{i=0}^{N_P} a_i^k \phi(\boldsymbol{\xi}_k); \tag{4.17}$$

The local mean μ_k and the local variance σ_k^2 of the kth element are, respectively, given by

$$\mu_k = a_0^k, \quad \sigma_k^2 = \sum_{i=1}^{N_P} (a_i^k)^2 E[\phi_{k,i}^2]; \tag{4.18}$$

The global mean μ_y and approximate global variance σ_y^2 are, respectively, given by

$$\mu_y = \sum_{k=1}^{m} a_0^k Pr(I_{B_k}), \tag{4.19}$$

$$\sigma_y^2 = \sum_{k=1}^{m} [(\sigma_k^2) + (a_0^k - \mu_y)^2] Pr(I_{B_k}). \tag{4.20}$$

In the adaptive procedure, the decay rate of the relative error of the PC in each element is first calculated from

$$\eta_k = \frac{\sum_{i=N_{(p-1)}+1}^{N_P} (a_i^k)^2 E[\phi_{k,i}^2]}{(\sigma_k^2)}. \tag{4.21}$$

Then, a random element is split into two equal probability parts provided the following criterion is satisfied:

$$\eta_k^\gamma Pr(I_{B_k}) \geq \theta_1, \quad 0 < \gamma < 1, \tag{4.22}$$

where θ_1 is a threshold and γ is a prescribed constant usually set to 0.5. When the random elements become smaller, the value of η_k that satisfies the criterion becomes large. (4.22) demonstrates that when the element becomes smaller, its probability becomes smaller so that the corresponding η_k can be relatively large. This relaxes the restrictions on achieving sufficient accuracy for small elements. To reduce the total number of elements while

increasing efficiency, only the random variables with higher sensitives are split to produce new elements. In this process, the sensitivity of each random variable is calculated via

$$r_i = \frac{(a_{i,P}^k)^2 E[\phi_{k,i,p}^2]}{\sum_{j=N_{(p-1)}+1}^{N_P} (a_i^k)^2 E[\phi_{k,j}^2]}, \tag{4.23}$$

where $(a_{i,p}^k)^2$ denotes the mode consisting only of random variable ξ_i with polynomial order P. All the sensitivities are compared to a threshold so that all random variables in the kth element satisfying

$$r_i \geq \theta_2 \max_{j=1,\ldots,N} r_j, \quad 0 < \theta_2 < 1, i = 1, \ldots, N, \tag{4.24}$$

are split into two equal probability elements, while all of the others stay unchanged. This is done to ensure accuracy without unduly sacrificing efficiency.

4.4.2 Implementation of the Multi-Element Generalized Polynomial Chaos Method in Power System Dynamic Simulations

The algorithm implementing the MEgPC method is described in Algorithm 2.

Algorithm 2 The MEgPC for a dynamic power systems model

```
Build a statistical dynamic power systems using gPC ;
for time step i = 1 to t do
    loop all the random elements;
        if η_k^γ Pr(I_{B_k}) ≥ θ_1 in element k then
            if r_i ≥ θ_2 max_{j=1,...,N} r_j then
                Split random variable into two equal ones;
                Map two new random variables;
                Construct the one-dimensional PC basis;
            end if
            Construct the N-dimensional PC basis ;
            Get statistics from the local elements via gPC;
        end if
        Update the global information using (4.19) and (4.20);
    end loop
end for
Plot the output graph with error bars;
```

Chapter 5

A Novel Polynomial-Chaos-based Kalman Filter

5.1 Introduction

Various nonlinear filters have been proposed in the literature, such as the extended Kalman filter (EKF), the unscented Kalman filter (UKF), the ensemble Kalman filter (EnKF), the Gauss-Hermite quadrature filter (GHQF) and the particle filter (PF), to cite a few [46, 51]. However, the application of these Kalman filters to solve real world problems has to overcome several challenges, such as possible strong system nonlinearities and large computational demands when applied to large-scale systems.

In general, the statistics/uncertainties of the states and the process and measurement noise are propagated through a carefully chosen point or several weighted sigma points or ensembles or particles in Kalman filters. this part provides an alternative way to propagate those statistics/uncertainties using the generalized polynomial chaos (gPC) method. Its basic philosophy is that the uncertainties can be propagated through a set of orthogonal polynomial chaos bases at selected collocation points [7]. Based on it, several Bayesian estimation tools have been proposed [52, 53], such as the gPC based EKF (gPC-EKF) and the ensemble Kalman filter (gPC-EnKF) [54–56]. However, the gPC-EKF may achieve poor performance in presence of strong system nonlinearities, which is not the case for the gPC-EnKF. But the latter suffers from the curse of dimensionality when the number of states is large [57–59].

To overcome the aforementioned difficulties, this part proposes a new polynomial-chaos-based Kalman filter (PCKF) that is able to track the dynamics of nonlinear dynamical systems subject to strong nonlinearities. To solve the curse of dimensionality, a dimension reduction strategy is proposed based on variance analysis. This allows us to construct more effective collocations points and to significantly improve the computational efficiency while maintaining a good accuracy.

5.2 Problem Statement

A nonlinear dynamic system can be typically expressed by the following two equations:

$$\boldsymbol{x}_k = \boldsymbol{f}(\boldsymbol{x}_{k-1}) + \boldsymbol{w}_k, \tag{5.1}$$

$$\boldsymbol{z}_k = \boldsymbol{h}(\boldsymbol{x}_k) + \boldsymbol{v}_k, \tag{5.2}$$

where $\boldsymbol{x}_k \in \mathbb{R}^{N\times 1}$ and $\boldsymbol{z}_k \in \mathbb{R}^{M\times 1}$ are the state vector and the measurement vector at time sample k, respectively; $\boldsymbol{f}$ and $\boldsymbol{h}$ are vector-valued nonlinear functions; $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are the system process and observation noise, respectively; they are assumed to be independent Gaussian white processes with zero mean and covariance matrices $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$, respectively.

To estimate $\boldsymbol{x}_k$, various types of gPC-based Kalman filters have been proposed in the literature. However, they suffer from the curse of dimensionality. this part proposes a new PCKF to address that. Before the introduction of the main results in Section I-V, a brief review of gPC is presented in Section III.

5.3 Review of Generalized Polynomial Chaos

The main idea of the gPC is to use a weighted sum of orthogonal polynomial chaos basis functions, which are constructed from the probability distribution of the input random variables, to represent the stochastic outputs. For instance, let y and $\boldsymbol{\xi} = [\xi_1, \xi_2, ...\xi_N]$ denote an output and a vector of random variables following a known probability distribution [4,7], respectively. If $\phi_i(\boldsymbol{\xi})$ and a_i denote a polynomial chaos basis function and the ith polynomial chaos coefficient, the stochastic output can be expressed as

$$y = \sum_{i=0}^{\infty} a_i \phi_i(\boldsymbol{\xi}). \tag{5.3}$$

A truncated expansion is typically used such that $y = \sum_{i=0}^{N_P} a_i \phi_i(\boldsymbol{\xi})$, where $N_P = \binom{N+P}{P} - 1$, N is the total number of the random variables and P is the maximum order of the polynomial chaos basis functions. Here, the N-variate gPC basis functions is constructed as the tensor product of the one-dimensional polynomial chaos basis associated with each input random variable of total degree less than or equal to P, that is: $\phi(\boldsymbol{\xi}) = \phi(\xi_1) \otimes \phi(\xi_2) \otimes \cdots \otimes \phi(\xi_N)$, where $\phi(\xi_i)$ denotes the one-dimensional polynomial chaos basis for the ith variable.

Orthogonality Relations

A set of one-dimensional polynomial chaos basis functions given by $\{\phi_i(\xi), i = 0, 1, 2, 3....\}$ with respect to some real positive measure satisfies the following relations: $\int \phi_r(\xi)\phi_s(\xi)d\lambda =$

$\gamma_r \delta_{rs}$ [21], where λ is a probability measure defined as the cumulative probability distribution function of ξ, for which, the associated orthogonal polynomials are unique; δ_{rs} is the Kronecker delta function, that is, $\delta_{rs} = 0$ if $s \neq r$ and $\delta_{rs} = 1$ if $s = r$; γ_r is a positive constant termed as the normalization constant for the rth order polynomial chaos basis.

Three-Term Recurrence Relation

The orthogonal polynomials satisfy a three-term recurrence relation given by [21]

$$\begin{aligned} \phi_{k+1}(\xi) &= (\xi - \alpha_k)\phi_k(\xi) - \beta_k \phi_{k-1}(\xi), \\ \phi_{-1} &= 0, \quad \phi_0 = 1; \quad k = 0, 1, 2 \ldots, K, \end{aligned} \tag{5.4}$$

where $\phi_k(\xi)$ is a set of orthogonal polynomials; α_k and β_k are the coefficients of the orthogonal polynomials of the kth order. For the standard normal distribution, its corresponding orthogonal polynomials are the Hermite polynomials and $\alpha_k = 0$ and $\beta_k = k$. After orthonormalization, the one-dimension orthonormal Hermite polynomials become

$$He_0 = 1, \quad He_1(\xi) = \xi, \quad He_2(\xi) = (\xi^2 - 1)/\sqrt{2} \quad \ldots \tag{5.5}$$

Statistical Moments

Using the aforementioned orthonormalized polynomial chaos bases, the mean, μ, and the variance, σ^2, of the output y can be determined by the polynomial chaos coefficients via

$$\mu = a_0, \ \sigma^2 = \sum_{i=1}^{N_P} a_i^2, \tag{5.6}$$

Similarly, for two orthonormal polynomial chaos expansions (PCE) $y_1 = \sum_{i=0}^{N_P} a_{1i}\phi_i(\boldsymbol{\xi})$ and $y_2 = \sum_{j=0}^{N_P} a_{2j}\phi_j(\boldsymbol{\xi})$, the covariance between them is given by [56–58]

$$\sigma^2_{y_1 y_2} = \sum_{i=1}^{N_P} a_{1i} a_{2i}. \tag{5.7}$$

Proof. Let's denote $\overline{y}_1$ and $\overline{y}_2$ as the means for y_1 and y_2, respectively. Since we have $\overline{y}_1 = a_{10}$,

$\overline{y}_2 = a_{20}$, the covariance $\sigma^2_{y_1y_2}$ is expressed as

$$\begin{aligned}\sigma^2_{y_1y_2} &= \int (\sum_{i=0}^{N_P} a_{1i}\phi_i(\boldsymbol{\xi}) - a_{10})(\sum_{j=0}^{N_P} a_{2j}\phi_j(\boldsymbol{\xi}) - a_{20})\lambda(\boldsymbol{\xi}) \\ &= \sum_{i=1}^{N_P}\sum_{j=1}^{N_P} a_{1i}a_{2j}E[\phi_i(\boldsymbol{\xi})\phi_j(\boldsymbol{\xi})]. \end{aligned} \tag{5.8}$$

Based on the orthogonality property, we reorganize (5.8) as

$$\begin{aligned}\sigma^2_{y_1y_2} &= \sum_{i=1}^{N_P}\sum_{j=i} a_{1i}a_{2j}E[\phi_i(\boldsymbol{\xi})\phi_j(\boldsymbol{\xi})] \\ &+ \sum_{i=1}^{N_P}\sum_{j\neq i} a_{1i}a_{2j}E[\phi_i(\boldsymbol{\xi})\phi_j(\boldsymbol{\xi})]. \end{aligned} \tag{5.9}$$

Here, $E[\phi_i(\boldsymbol{\xi})\phi_j(\boldsymbol{\xi})]$ in the second summation equals to 0. Since the polynomial chaos bases are orthonormal, (5.8) can be rewritten as

$$\sigma^2_{y_1y_2} = \sum_{i=1}^{N_P} a_{1i}a_{2i}E[\phi_i^2(\boldsymbol{\xi})] = \sum_{i=1}^{N_P} a_{1i}a_{2i}. \tag{5.10}$$

Thus, the proof is completed. □

Remark: In the proposed PCKF shown later, the polynomial chaos coefficients will be used to calculate the first two statistical moment information of the predicted states and measurements.

5.4 Polynomial-Chaos-Based Kalman Filter

The proposed PCKF approach consists of five components, namely the PCE dimension reduction, the construction of the gPC basis matrix, the system and observation process, and finally the correction process. In what follows, the proposed PCE dimension reduction strategy is presented first.

5.4.1 Dimension Reduction Strategy for the PCKF

PCE reduction

The PCE contains $\binom{N+P}{P}$ terms that make the traditional gPC-EKF and gPC-EnKF to suffer from the "curse of dimensionality" when the number of states N increases. As a result, its

application to practical systems is very limited. To solve this problem, a dimension reduction strategy based on variance analysis is proposed. Specifically, by rewriting the traditional truncated polynomial chaos expansion (6.22) as a function of N random variables, we get

$$y = a_0 + \sum_{i=1}^{N_P} a_i \phi_i(\boldsymbol{\xi}), \tag{5.11}$$

where a_0 reflects the mean and all the other terms have zero means and variance $a_i^2 E[\phi_i^2]$. This is proved as follows:

Proof. Let $y_i = a_i \phi_i(\boldsymbol{\xi})$. Its variance is expressed as

$$\sigma^2(y_i) = \int_{\Gamma^N} (a_i \phi_i(\boldsymbol{\xi}) - \mu_i)^2 d\lambda(\boldsymbol{\xi}), \tag{5.12}$$

where μ_i is the mean of y_i expressed as

$$\mu_i = \int_{\Gamma^N} (a_i \phi_i(\boldsymbol{\xi})) d\lambda(\boldsymbol{\xi}) = a_i \int_{\Gamma^N} \phi_i(\boldsymbol{\xi}) d\lambda(\boldsymbol{\xi}). \tag{5.13}$$

Since $\int_{\Gamma^N} \phi_i(\boldsymbol{\xi}) d\lambda(\boldsymbol{\xi}) = 0$, y_i has zeros mean, implying that (5.12) is given by

$$\sigma^2(y_i) = \int_{\Gamma^N} (a_i \phi_i(\boldsymbol{\xi}))^2 d\lambda(\boldsymbol{\xi}) = a_i^2 E[\phi_i^2]. \tag{5.14}$$

□

Note that a polynomial chaos expansion with $P = 2$ is a cost-effective choice in estimating the first two statistical moments [54], [56], [15, 17–19]. Therefore, (5.11) can be further reorganized as

$$y = a_0 + \sum_{i=1}^{N} a_i \phi(\xi_i) + \sum_{i=1}^{N} a_{i,i} \phi(\xi_i^2) \quad + \sum_{i=1}^{N} \sum_{1 \le j < i} a_{i,j} \phi(\xi_i, \xi_j), \tag{5.15}$$

where the first and the second summations represent the bases that are involved with the 1st order and the 2nd order statistical moment information, respectively; the third summation reflects the bases that are related to the mutual information of the two random variables. The former two summations contribute to the variance of y considering the information of individual variable while the third one does that by considering the coupling effects of two random variables [31, 33]. Note that the "curse of dimensionality" in the PCE is due to the inclusion of the third summation, which has a number of terms $\binom{N}{2}$.

In dynamic state estimation, the system is typically driven by the white process noise and the second moment information influenced by the coupling effect of the white noise is negligible.

For example, the GHQF, which uses Hermite polynomials to propagate the uncertainties, neglects the influences from the coupled Hermite Polynomials for the Gauss-Hermite Quadrature points [51]. Motivated by that, the truncated PCE can be simplified as follows:

$$y = a_0 + \sum_{i=1}^{N} a_i \phi(\xi_i) + \sum_{i=1}^{N} a_{i,i} \phi(\xi_i^2), \tag{5.16}$$

where only $2N+1$ polynomial chaos coefficients need to be approximated, which is much smaller than the original $\binom{N+2}{2}$ terms when N is large. To further improve the computational efficiency, the collocations points are re-generated next.

Collocation Points

Collocation Points (CPs) can be regarded as a finite sample of $\boldsymbol{\xi} = [\xi_1, \xi_2, ...\xi_N]$ for approximating the polynomial chaos coefficients. The elements of the CPs are generated by the roots of one higher-order one-dimensional polynomial for every random variable [4]. For example, for the Gaussian distribution with a second order PCE, its corresponding third order PCE is $\phi_3(\xi) = (\xi^3 - 3\xi)/\sqrt{6}$, yielding $\{-\sqrt{3}, 0, \sqrt{3}\}$. The widely used methods for constructing the combinations of CPs include the tensor-product rule [54], the sparse-grid method [**?**] and the pre-select method [17]. The tensor-product rule needs $(P+1)^N$ samples to simulate the forward solver whereas the sparse-grid method requires fewer samples by using nested grid samples. As indicated by its name, the pre-select method pre-selects $\binom{N+P}{P}$ proper samples, yielding the least number of samples with a good accuracy. It is worth pointing out that after the PCE dimension reduction is carried out, only $2N+1$ polynomial chaos coefficients are required by our proposed PCKF. Thus, only $2N+1$ proper combinations of CPs are needed. To this end, a $(N) \times (2N+1)$ matrix $\boldsymbol{\Xi}$ can be used as follows:

$$\boldsymbol{\Xi} = \begin{bmatrix} -\sqrt{3} & 0 & \dots & 0 & 0 & \sqrt{3} & 0 & \dots & 0 \\ 0 & -\sqrt{3} & \dots & \vdots & 0 & 0 & \sqrt{3} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -\sqrt{3} & 0 & 0 & 0 & \dots & \sqrt{3} \end{bmatrix}. \tag{5.17}$$

The ith column of $\boldsymbol{\Xi}$ is denoted as $\boldsymbol{\xi}_i$. It represents the ith combination of the CPs.

Remark: It is worth pointing out that the proposed combination strategy for dimension reduction and regeneration of the CPs allows us to mitigate the curse of dimensionality without any loss of estimation accuracy. This is the main contribution of our work as compared to the state-of-the-art.

5.4.2 Framework of the proposed PCKF

The other four steps of the proposed framework are as follows:

gPC Basis Matrix

Following [17], we put the collocation points of the orthonormal polynomial chaos functions into the $((2N+1)\times(2N+1))$ matrix $\boldsymbol{H}$ to get

$$\boldsymbol{H} = \begin{bmatrix} \phi_0(\boldsymbol{\xi}_1) & \dots & \phi_{2N}(\boldsymbol{\xi}_1) \\ \vdots & \ddots & \vdots \\ \phi_0(\boldsymbol{\xi}_{2N+1}) & \dots & \phi_{2N}(\boldsymbol{\xi}_{2N+1}) \end{bmatrix}. \tag{5.18}$$

Subsequently, the inverse of $\boldsymbol{H}$ is calculated as $\boldsymbol{H}_{inv} = \boldsymbol{H}^{-1}$, which is used to approximate the PC coefficients. In our approach, we will check the rank of $\boldsymbol{H}$ so that it is invertible. If it is badly scaled, an alternative approach of constructing an invertible $\boldsymbol{H}$ can be executed, such as the pre-selected strategy [17].

System Process Prediction

By taking into account the system process noise, the ith mapped CP is updated as follows:

$$\boldsymbol{x}^i_{k-1|k-1} = \hat{\boldsymbol{x}}_{k-1|k-1} + \sigma_{\boldsymbol{w}}\boldsymbol{\xi}_i, \tag{5.19}$$

where $\sigma_{\boldsymbol{w}}$ is the standard deviation of the process noise. Next, the polynomial chaos bases at each collocation point are propagated thorough the system process model (5.1), yielding a set of transformed samples expressed as

$$\boldsymbol{x}^i_{k|k-1} = \boldsymbol{f}(\boldsymbol{x}^i_{k-1|k-1}), \tag{5.20}$$

where $\boldsymbol{x}^i_{k|k-1} \in \mathbb{R}^{N\times 1}$ is the predicted state vector from the ith CP. These state vectors are organized into the following compact form

$$\hat{\boldsymbol{X}} = \begin{bmatrix} \boldsymbol{x}^1_{k|k-1} & \boldsymbol{x}^2_{k|k-1} & \dots & \boldsymbol{x}^{2N+1}_{k|k-1} \end{bmatrix}^T. \tag{5.21}$$

According to (5.16), the PC coefficients matrix of the system model can be put into a $(2N+1)\times(N)$ matrix $\boldsymbol{A}$, i.e.,

$$\boldsymbol{A} = \begin{bmatrix} a_{1,0} & \dots & a_{N,0} \\ \vdots & \ddots & \vdots \\ a_{1,2N} & \dots & a_{N,2N} \end{bmatrix}, \tag{5.22}$$

which is calculated by $\boldsymbol{A} = \boldsymbol{H}_{inv}\hat{\boldsymbol{X}}$. Note that the element $a_{i,j}$ of $\boldsymbol{A}$ represents the jth PC coefficient for the ith state variable. By defining $\boldsymbol{A_1}$ as the transpose of the first row of matrix $\boldsymbol{A}$, we get the following formulation from (5.6):

$$\hat{\boldsymbol{x}}_{k|k-1} = \begin{bmatrix} a_{1,0} & \dots & a_{N,0} \end{bmatrix}^T = \boldsymbol{A}_1; \tag{5.23}$$

Furthermore, the matrix to reflect the second moment information is defined as

$$\boldsymbol{A}_2 = \begin{bmatrix} a_{1,1} & \dots & a_{N,1} \\ \vdots & \ddots & \vdots \\ a_{1,2N} & \dots & a_{N,2N} \end{bmatrix}; \tag{5.24}$$

Then, by using (5.6) and (5.7), we get

$$\boldsymbol{P}^{xx}_{k|k-1} = \boldsymbol{A}_2^T \boldsymbol{A}_2 + \boldsymbol{Q}. \tag{5.25}$$

Observation Process Prediction

A similar procedure is applied to the observation model (5.2) to get

$$\boldsymbol{z}^i_{k|k-1} = \boldsymbol{h}(\boldsymbol{x}^i_{k|k-1}), \tag{5.26}$$

where $\boldsymbol{z}^i_{k|k-1} \in \mathbb{R}^{M\times 1}$ is the predicted measurement vector calculated by the ith CP, yielding the measurement prediction $\hat{\boldsymbol{Z}}$ matrix as

$$\hat{\boldsymbol{Z}} = \begin{bmatrix} \boldsymbol{z}^1_{k|k-1} & \boldsymbol{z}^2_{k|k-1} & \dots & \boldsymbol{z}^{2N+1}_{k|k-1} \end{bmatrix}^T. \tag{5.27}$$

Then, the PC coefficients matrix for the observation model, $\boldsymbol{W}$ can be calculated via

$$\boldsymbol{W} = \boldsymbol{H_{inv}} \hat{\boldsymbol{Z}}. \tag{5.28}$$

Let the transpose of the first row and the remaining parts of $\boldsymbol{W}$ as $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ reflect the first and second moment information, respectively. We then get

$$\hat{\boldsymbol{z}}_{k|k-1} = \boldsymbol{W}_1. \tag{5.29}$$

State Correction

Based on (5.6) and (5.7), the self and cross-covariance matrices, $\boldsymbol{P}^{zz}_{k|k-1}$ and $\boldsymbol{P}^{xz}_{k|k-1}$, are respectively calculated by

$$\boldsymbol{P}^{zz}_{k|k-1} = \boldsymbol{W}_2^T \boldsymbol{W}_2 + \boldsymbol{R}_k; \tag{5.30}$$

$$\boldsymbol{P}^{xz}_{k|k-1} = \boldsymbol{A}_2^T \boldsymbol{W}_2; \tag{5.31}$$

Finally, the state correction $\hat{\boldsymbol{x}}_{k|k}$ is performed using the Kalman filter, yielding

$$\boldsymbol{K}_k = \boldsymbol{P}^{xz}_{k|k-1} (\boldsymbol{P}^{zz}_{k|k-1})^{-1}, \tag{5.32}$$

$$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k(\boldsymbol{z} - \hat{\boldsymbol{z}}_{k|k-1}), \tag{5.33}$$

$$\boldsymbol{P}^{xx}_{k|k} = \boldsymbol{P}^{xx}_{k|k-1} - \boldsymbol{K}_k \boldsymbol{P}^{zz}_{k|k-1} \boldsymbol{K}_k^T. \tag{5.34}$$

Chapter 6

Response-Surface-Based Bayesian Inference for Power System Dynamic Parameter Estimation Using PMU Measurement

6.1 Introduction

Accurately estimating the key parameters of the generator dynamic models is of paramount importance to power system dynamic security analysis. Indeed, inaccurate parameter values may lead to inappropriate control actions, which in turn may result in power system failures. It was found that the well-known 1996 blackout in the Western Electricity Coordinating Council (WECC) system was caused by severe model discrepancy. This has prompted the North American Electric Reliability Corporation (NERC) to issue reports that require the validation and the calibration of the models of large generators and their controllers in North America every five years, aiming to have the dynamic model responses match the recorded measurements reasonably well [65–67]. To address this need, this part proposes a novel method for estimating generator dynamic model parameters using Bayesian formalism.

Because traditional offline stage-testing-based methods are very costly, time-consuming, and labor intensive for generator model validation and parameter estimation [68], a number of low-cost phasor-measurement-unit (PMU)-based methods have been proposed in recent years. Some of them are focused on estimating the generator moment of inertia [69–71], [72] while others are focused on estimating the aggregated inertia of the system [73]. However, the dynamic responses of the power systems are not governed by the moment of inertia only. Simply using the swing equations to describe the dynamic responses may lead to inaccurate estimation of the generator model parameters. In [68], [74], it has been shown

in the sensitivity-based methods that other generator model parameters also play key roles in influencing the power system dynamic performance. To estimate these key parameters when the system is subject to disturbances, methods based on the extended Kalman filter (EKF), the unscented Kalman filter (UKF) and the ensemble Kalman filter (EnKF) have been advocated in [67, 68, 74]. While overall these methods exhibit good performances, in practice, they still suffer from several weaknesses such as poor convergence rate and Gaussian assumption of the process and measurements error.

Unlike the weighted-least-squares (WLS)-based methods proposed, for example, in [75, 76], which provide only point estimated values for the generator model parameters without providing the confidence intervals, Bayesian-inference based methods provide maximum-a-posteriori (MAP) parameter estimated values along with the corresponding posterior probability distributions [72], [77]. Furthermore, they are able to cope with non-Gaussian errors [77]. For these reasons, Bayesian methods have gained a great deal of attention among power researchers. However, the posterior distribution in Bayesian inference is typically obtained by Monte-Carlo-based methods, which require prohibitive computing time when they are performed on detailed generator models.

To overcome the aforementioned weaknesses, this part proposes a response-surface-based Bayesian inference algorithm for power system dynamic parameter estimation, resulting in the following contributions:

- To eliminate the uncertainties of the line and transformer models and the loads within a centralized system model [72], a decentralized generator model is used [78], [79]. This significantly reduces the number of parameters, allowing effective Bayesian inference for power system dynamic parameter estimation.
- To deal with the non-Gaussian posterior distribution, the Metropolis-Hastings sampler is adopted. This is a Markov chain Monte Carlo (MCMC) method aimed at finding the global optima for the MAP estimator. This overcomes the shortcomings of the local optimization approach [72], in which the local optima might be highly biased when the initial guess is far from the true value or the posterior distribution is non-Gaussian.
- To greatly accelerate the MCMC algorithm [77], the response surfaces of the dynamical system represented by the state-of-the-art polynomial chaos expansion (PCE) are developed [4].
- To identify the key generator model parameters that satisfy system observability, a polynomial-chaos-based ANOVA algorithm is proposed.

Simulation results carried out on the New-England system reveal that our proposed method can accurately and simultaneously estimate several key generator dynamic model parameters with almost two order of magnitude improved in speed compared with the traditional MCMC-based Bayesian inference method for non-Gaussian posterior distribution.

this part is organized as follows: in Section II, the power system dynamic parameter estimation problem is formulated under the Bayesian framework. In Section III, the use of the polynomial chaos method to represent the response surface for Bayesian inference is presented. Section IV presents the proposed response surface based Bayesian inference via MCMC method. Section V presents the case study of the proposed method on the New-England test system. Conclusions and future work are given in Section VI.

6.2 Problem Formulation

This section formulates the Bayesian inference framework for the decentralized synchronous generator model and the measurement model. Following Sauer and Pai [44], the two-axis model with a IEEE-DC1A exciter and a TGOV1 turbine-governor is considered. It is represented by the following differential and algebraic equations:

$$T'_{d_0}\frac{dE'_q}{dt} = -E'_q - (X_d - X'_d)I_d + E_{fd}, \tag{6.1}$$

$$T'_{q_0}\frac{dE'_d}{dt} = -E'_d - (X_q - X'_q)I_q, \tag{6.2}$$

$$\frac{d\delta}{dt} = \omega - \omega_s, \tag{6.3}$$

$$\frac{2H}{\omega_s}\frac{d\omega}{dt} = T_M - P_e - D(\omega - \omega_s), \tag{6.4}$$

$$T_E\frac{dE_{fd}}{dt} = -(K_E + S_E(E_{fd}))E_{fd} + V_R, \tag{6.5}$$

$$T_F\frac{dV_F}{dt} = -V_F + \frac{K_F}{T_E}V_R - \frac{K_F}{T_E}(K_E + S_E(E_{fd}))E_{fd}, \tag{6.6}$$

$$T_A\frac{dV_R}{dt} = -V_R + K_A(V_{\text{ref}} - V_F - V), \tag{6.7}$$

$$T_{CH}\frac{dT_M}{dt} = -T_M + P_{SV}, \tag{6.8}$$

$$T_{SV}\frac{dP_{SV}}{dt} = -P_{SV} + P_C - \frac{1}{R_D}(\frac{\omega}{\omega_s} - 1), \tag{6.9}$$

$$V_d = V\sin(\delta - \theta), V_q = V\cos(\delta - \theta), \tag{6.10}$$

$$I_d = \frac{E'_q - V_q}{X'_d}, I_q = \frac{V_d - E'_d}{X'_q}, \tag{6.11}$$

$$P_e = V_dI_d + V_qI_q, Q_e = -V_dI_q + V_qI_d, \tag{6.12}$$

where T'_{d_0}, T'_{q_0}, T_E, T_F, T_A, T_{CH} and T_{SV} are the time constants; K_E, K_F and K_A are the controller gains; E'_d, E'_q, E_{fd}, V_F, V_R, T_M and P_{SV} are the d-axis and q-axis transient

voltages, field voltage, scaled output of the stabilizing transformer and scaled output of the amplifier, synchronous machine mechanical torque and steam valve position, respectively; X_d, X'_d, X_q and X'_q are the generator parameters; H, D and R_D are the inertia constant in seconds, damping ratio and droop respectively; V_{ref} and P_C are the known voltage reference and power references of the exciter and speed governor, respectively.

Following Zhao and Mili [79], once the local PMU measurements are recorded and the parameters of the ith generator model to be estimated are selected while satisfying state observability, we take the voltage phasors $V_i\angle\theta_i$ as the known model inputs and the current phasors $I_i\angle\phi_i$ as the model outputs. Based on (6.10) and (6.11), we have $V_{di} = V_i\sin(\delta_i - \theta_i)$, $V_{qi} = V_i\cos(\delta_i - \theta_i)$, $I_{di} = (E'_{qi} - V_{qi})/X'_{di}$ and $I_{qi} = (V_{di} - E'_{di})/X'_{qi}$. However, paper [80] shows that the terminal real and reactive powers enable a better observability of the generator states than the current phasors. From (6.12), we further take the active power and reactive power for the ith decentralized generator as model outputs expressed as

$$P_{ei} = V_{di}I_{di} + V_{qi}I_{qi} + e_{Pi}, \tag{6.13}$$

$$Q_{ei} = -V_{di}I_{qi} + V_{qi}I_{di} + e_{Qi}, \tag{6.14}$$

where e_{Pi} and e_{Qi} are the measurement noise. Up to now, we have formulated the decentralized generator model. Once we capture the local voltage phasor $V_i\angle\theta_i$ for the ith generator in the selected time period, given its corresponding generator parameters, we can obtain the trajectories of its active and reactive power by calculating the ith generator model's output using the aforementioned differential and algebraic equations given by (6.1)-(6.14).

Discussion: This decentralized model has been widely used in Kalman filter based methods in both parameter estimation and state estimation [67,68,79]. In [67,68], Huang *et al.* name it an "event playback". This procedure can be shown in Fig. 6.1. For the ith generator, when an event occurs in the system, the external system is represented as an infinite bus with varying bus voltage magnitude, V_i, and phase angle, $\angle\theta_i$. Then, the PMU-measured V_i and $\angle\theta_i$ is called the *play-in* function that can be injected into the aforementioned differential-algebraic equations to simulated the model output P and Q. The simulated P and Q trajectories are further used to estimate the model parameters based on PMU-measured P and Q trajectories.

Remark: In this part, the decentralized generator model in parameter estimation will be the first time formulated into Bayesian inference framework.

Tarantola [81] formulated the model for Bayesian inference as

$$\boldsymbol{d} = \boldsymbol{f}(\boldsymbol{m}) + \boldsymbol{e}, \tag{6.15}$$

where $\boldsymbol{d}$ contains the observations, which consists of d_1 that contains the active power measurements, P_{ei}, and d_2 that contains the reactive power measurements, Q_{ei}; $\boldsymbol{m} \in \mathbb{R}^N$ are the parameters to be estimated; N is the number of parameters to be estimated that depends on the specific applications; $\boldsymbol{f}(\cdot)$ is the vector-valued forward function that represents the

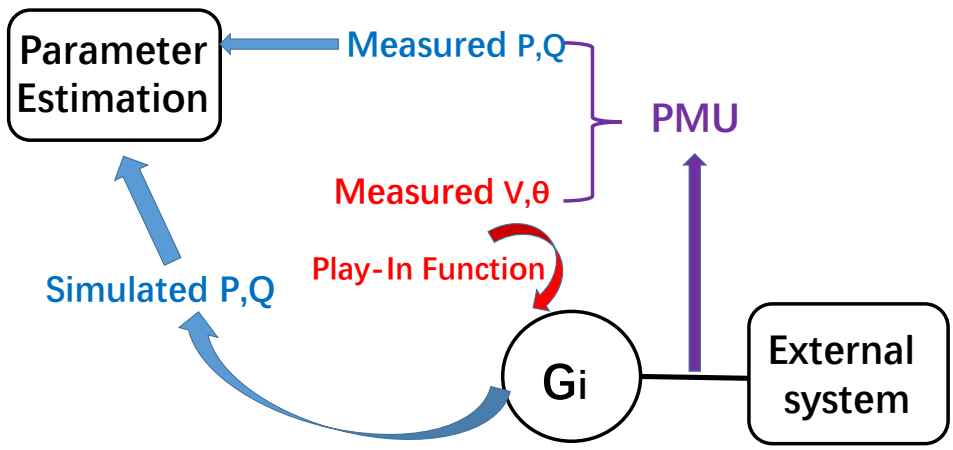


Figure 6.1: Parameter estimation with event playback.

aforementioned differential and algebraic equations, which map the model parameters $\boldsymbol{m}$ to the observations $\boldsymbol{d}$; $\boldsymbol{e} \in \mathbb{R}^2$ stands for the measurement error vector whose components are assumed to be mutually independent random variables with the joint probability density functions $\pi_{\boldsymbol{e}}$ defined as $\pi_{\boldsymbol{e}} = \prod_{i=1}^{2} \pi_{e_i}(e_i)$. In the Bayesian inference, each parameter m_i is also viewed as a random variable with a given prior probability distribution, whose probability density function (pdf) is denoted as $\pi_i(m_i)$. The corresponding joint prior density function for a vector $\boldsymbol{m}$ is given by

$$\pi_{\text{prior}}(\boldsymbol{m}) = \prod_{i=1}^{N} \pi_i(m_i). \tag{6.16}$$

Note that here $\boldsymbol{e}$ and $\boldsymbol{m}$ are also assumed to be mutually independent.

Given the observation $\boldsymbol{d}$, the posterior pdf $\pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d})$ for the parameters $\boldsymbol{m}$ is derived as [81]

$$\pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d}) \propto \pi_{\text{like}}(\boldsymbol{d}|\boldsymbol{m})\pi_{\text{prior}}(\boldsymbol{m}). \tag{6.17}$$

Here $\pi_{\text{like}}(\boldsymbol{d}|\boldsymbol{m})$ denotes the likelihood function, expressed as

$$\pi_{\text{like}}(\boldsymbol{d}|\boldsymbol{m}) = \prod_{i=1}^{2} \pi_{e_i}(d_i - f_i(\boldsymbol{m})). \tag{6.18}$$

Given a set of parameters contained in $\boldsymbol{m}$, we obtain the trajectories of P_{ei} and Q_{ei} from the forward solver $\boldsymbol{f}(\cdot)$. By comparing them to the PMU metered values for the simulated time period t_{end}, the likelihoods for the corresponding trajectories are evaluated. Let us denote

$\pi_{e_i}^t$, d_i^t and f_i^t as the likelihood, the observation and the realization at time t, respectively. The likelihood function for the trajectories in the log-form is then expressed as

$$\log \pi_{e_i}(d_i - f_i(\boldsymbol{m})) = \sum_{t=0}^{t_{\text{end}}} \log \pi_{e_i}^t (d_i^t - f_i^t(\boldsymbol{m})). \tag{6.19}$$

Thanks to the high speed sampling rate of the PMUs, which is typically equal to 60 samples/s, for a short time period such as 3 seconds, we get 180 samples for P_{ei} and for Q_{ei}, which provide good tracking information of the dynamic responses of a power system following a disturbance. Now the relationship given by 6.17) can be put into the following form:

$$\log(\pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d})) \propto \sum_{i=1}^{2} \log \pi_{e_i}(d_i - f_i(\boldsymbol{m})) + \sum_{i=1}^{N} \log(\pi_i(m_i)), \tag{6.20}$$

yielding the MAP estimator defined as

$$\hat{\boldsymbol{m}}_{\text{MAP}} = \arg\min_{\boldsymbol{m}} \quad \{-\log(\pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d}))\}. \tag{6.21}$$

Note that due to the non-linearity of $\boldsymbol{f}(\cdot)$, even if the prior assumption is Gaussian, $\pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d})$ may be non-Gaussian. This motivates us to use the MCMC method instead of simply making the Gaussian assumptions for the posterior pdfs of the parameters $\boldsymbol{m}$. The detailed procedure to calculate the MCMC via the Metropolis-Hastings sampler will be described in Section IV.

6.3 Response-Surface-Based Bayesian Inference

The MCMC method is very time-consuming when the forward solver is complex. To accelerate it, we propose to replace the forward solver with a PCE-based response surface, as explained next.

6.3.1 Review of the Generalized Polynomial Chaos Expansion

Introduced by Wiener and further developed by Xiu and Karniadakis [7], the generalized polynomial chaos expansion has been shown to be a cost-effective tool in modeling response surfaces [82], [83], [17]. In the gPC 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 $\boldsymbol{\xi} = [\xi_1, \xi_2, ...\xi_N]$ be a vector of random variables following a standard probability distribution (e.g. the Gaussian or the beta distribution), to which, as shown in [7], a unique orthogonal polynomial is associated.

Let $\Phi_i(\xi_1, \xi_2, ...\xi_N)$ denote this procedure's corresponding polynomial chaos basis and let a_i denote the ith polynomial chaos coefficient. Formally, we have

$$z = \sum_{i=0}^{\infty} a_i \Phi_i(\boldsymbol{\xi}). \tag{6.22}$$

In practice, a truncated expansion is used such that

$$z = \sum_{i=0}^{N_P} a_i \Phi_i(\boldsymbol{\xi}), \tag{6.23}$$

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. It is found that a relatively low maximum polynomial chaos order, typically 2, is found to provide output results with enough accuracy [83], [17], [88]. From the polynomial chaos coefficients, the mean, μ, and the variance, σ^2, of the output z can be determined as follows:

$$\mu = a_0, \quad \sigma^2 = \sum_{i=1}^{N_P} a_i^2 E[\Phi_i^2], \tag{6.24}$$

where $E[.]$ is the expectation operator.

Collocation Points

collocation points can be regarded as a finite sample of $\boldsymbol{\xi} = [\xi_1, \xi_2, \cdots, \xi_N]$ 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 [4, 17]. 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. For example, the $N_P + 1$ unknown coefficients can be estimated by the WLS method introduced below. Besides this tensor-product method, the widely used methods for constructing the combinations of CPs include the sparse-grid method [4] and the pre-select method [17]. Unlike the tensor-product rule that needs $(P+1)^N$ samples to simulate the forward solver, the sparse-grid method requires fewer samples by using nested grid samples. As indicated by its name, the pre-select method pre-selects $\binom{N+P}{P}$ proper samples, yielding the least number of samples but sacrifacing some accuracy. In this part, we adopt the sparse-grid method. This has been shown to be a cost-effective way to approximate PCE coefficients in [72].

6.3.2 Building PCE-based Response Surface for Dynamic Power Systems

In the Bayesian inference, the parameters $\boldsymbol{m}$ are viewed as random variables and hence, are given prior PDFs. By mapping the parameters $\boldsymbol{m}$ into $\boldsymbol{\xi}$, we can build a PCE as the response surface of the dynamic power system model. The detailed gPC procedure is as follows:

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

$$P_i = F_i^{-1}(T(\xi_i)), \tag{6.25}$$

where F_i^{-1} is the inverse cumulative probability distribution function of m_i and T is the cumulative probability distribution function of ξ_i.

(b) Construct the polynomial chaos basis, then express the output z in the gPC expansion form.

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

$$H_{pc} = \begin{pmatrix} \Phi_0(\boldsymbol{\xi_1}) & \Phi_1(\boldsymbol{\xi_1}) & \dots & \Phi_{N_P}(\boldsymbol{\xi_1}) \\ \Phi_0(\boldsymbol{\xi_2}) & \Phi_1(\boldsymbol{\xi_2}) & \dots & \Phi_{N_P}(\boldsymbol{\xi_2}) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_0(\boldsymbol{\xi_M}) & \Phi_1(\boldsymbol{\xi_M}) & \dots & \Phi_{N_P}(\boldsymbol{\xi_M}) \end{pmatrix}; \tag{6.26}$$

(d) Compute the power system dynamic model output for the selected collocation points to get the $(M \times 1)$ output Z matrix given by

$$Z = \begin{pmatrix} z(t, \boldsymbol{\xi_1}) & z(t, \boldsymbol{\xi_2}) & \dots & z(t, \boldsymbol{\xi_M}) \end{pmatrix}^T; \tag{6.27}$$

(e) Estimate the unknown coefficients A based on the collocation points that are selected and the model output from:

$$Z = H_{pc}A. \tag{6.28}$$

A is the $(N_P \times 1)$ coefficient vector expressed as

$$A = \begin{pmatrix} a_0(t) & a_1(t) & \dots & a_{N_P}(t) \end{pmatrix}^T; \tag{6.29}$$

(f) Let $\hat{A}$ denote the estimated coefficient vector and let us define a residual vector r as $r = Z - H_{pc}\hat{A}$. Let us minimize the 2-norm of the residual vector to estimate $\hat{A}$, that is,

$$J(\hat{A}) = \arg\min_{\hat{A}} \quad r^T r, \tag{6.30}$$

The necessary and sufficient condition of optimal is given by

$$\frac{dJ(\hat{A})}{d\hat{A}} = \frac{d[(Z - H_{pc}\hat{A})^T(Z - H_{pc}\hat{A})]}{d\hat{A}} \tag{6.31}$$

$$= -2H_{pc}^T(Z - H_{pc}\hat{A}) = 0. \tag{6.32}$$

It yields $\hat{A} = (H_{pc}^T H_{pc})^{-1} H_{pc}^T Z$. It is worth pointing out that the aforementioned three methods to generate CPs can provide accurate approximations of the PCE coefficients. The suggested sparse-grid method can not only provide good numerical stability of the regression problem but requires less computing time than the tensor-product method.

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

Remark: In practice, the industrial generator model is more complicated than the generator model we described above. The industrial generator model has more states and parameters. Without using the PCE based response surface , it would be very time-consuming to conduct the dynamic simulations. However, when the number of parameter increases, the number of terms involved in PCE will also increase. This will relatively reduce the efficiency of the PCE-based response surface. But it remains much more cost-effective than the full dynamic solver.

6.3.3 Incorporating Polynomial Chaos Expansion into the Bayesian Inference Framework

In the gPC-based Bayesian inference [82,83], we use the approximated gPC solution in (7.9) to replace the exact forward solver solution $\boldsymbol{f(m)}$ in (6.20) as

$$\log(\pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d})) \propto \sum_{i=1}^{2} \log \pi_{e_i}(d_i - z(\boldsymbol{\xi})) + \sum_{i=1}^{N} \log(\pi_i(m_i)). \tag{6.33}$$

Once a sample point $\boldsymbol{m}$ is proposed in MCMC, we can first map $\boldsymbol{m}$ to $\boldsymbol{\xi}$ via (6.25) and then evaluate $\boldsymbol{\xi}$ with the PCE-based response surface without resorting to actual simulations of the forward solver. Therefore, we achieve very high accuracy in sampling the posterior distribution at a much less computational cost.

6.4 Application of the Proposed Response-Surface-Based Bayesian Inference Algorithm

This section illustrate the application of the proposed response-surface-based Bayesian inference algorithm via the Metropolis-Hastings (M-H) algorithm to achieve a decentralized

power system dynamic parameter estimation. Starting from the initial guess, which can be the manufactured data, the M-H method employs a given PDF, $q(\boldsymbol{m_k}, \cdot)$ at each sample point $\boldsymbol{m}_k$ to generate a proposed sample point $\boldsymbol{m_{k+1}}$. Once generated, the sample point is either accepted or rejected by the M-H method. This procedure is then applied to the next sample point, yielding a chain of sample points from the posterior pdf $\pi_{\text{post}}(\boldsymbol{m})$ [81]. To further stabilize the numerical computation of the M-H algorithm, the posterior pdf is transformed into the log-form as suggested in [77]. The M-H algorithm used in the classical Bayesian inference is described in Algorithm 3. In this algorithm, the most time-consuming

Algorithm 3 The Bayesian Inference using the M-H Algorithm

1: Choose the initial guess of the parameters $\boldsymbol{m_0}$ from the manufactured data as the Bayesian prior $\boldsymbol{m}_{\text{prior}}$;
2: Compute $\log(\pi_{\text{post}}(\boldsymbol{m_0}|\boldsymbol{d}))$ for the exact decentralized dynamic model from (6.20);
3: **for** $k = 0, \ldots, N_{\text{samples}} - 1$ **do**
4: Generate new sample $\boldsymbol{m_{k+1}}$ from the proposal function $q(\boldsymbol{m_k}, \cdot)$;
5: **Compute $\log(\boldsymbol{\pi_{\text{post}}}(\boldsymbol{m_{k+1}}|\boldsymbol{d}))$ from the exact decentralized dynamic model by (6.20);**
6: Calculate the correction factor $c = \frac{q(\boldsymbol{m_{k+1}}, \boldsymbol{m_k})}{q(\boldsymbol{m_k}, \boldsymbol{m_{k+1}})}$;
7: Compute $\alpha(\boldsymbol{m_k}, \boldsymbol{m_{k+1}})$ defined as

$$\alpha(\boldsymbol{m_k}, \boldsymbol{m_{k+1}}) = \log(min\{1, \frac{\pi_{\text{post}}(\boldsymbol{m_{k+1}}|\boldsymbol{d})}{\pi_{\text{post}}(\boldsymbol{m_k}|\boldsymbol{d})} \cdot c\});$$

8: Draw $u \sim \mathcal{U}([0, 1))$;
9: **if** $\log(u) < \alpha(\boldsymbol{m_k}, \boldsymbol{m_{k+1}})$ **then**
10: Accept: Set $\boldsymbol{m}_{k+1} = \boldsymbol{m_{k+1}}$;
11: **else**
12: Reject: Set $\boldsymbol{m}_{k+1} = \boldsymbol{m}_k$;
13: **end if**
14: **end for**
15: Plot the PDF of the N_{samples} of $\boldsymbol{m}$ obtained via the above procedure as the $\pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d})$ and find the MAP points.

step happens in step 5, this is accelerated by the PCE-based response surface, as shown in the modified M-H algorithm given in Algorithm 4. In this modified version, the sample size needed for PCE surrogate construction is much smaller than the samples required by the original M-H algorithm. This step can also be executed offline. In the following, we compare the two Bayesian-inference-based parameter estimation methods.

Algorithm 4 The Response-surface-based Bayesian Inference using the M-H Algorithm

Choose the initial guess of the parameters $\boldsymbol{m_0}$ from the manufactured data as the Bayesian prior $\boldsymbol{m}_{\text{prior}}$;
Build the PCE surrogates as the response surface of the decentralized dynamic model;
Compute $\log(\pi_{\text{post}}(\boldsymbol{m_0}|\boldsymbol{d}))$ from the PCE surrogate via (7.15);
for $k = 0, \ldots, N_{\text{samples}} - 1$ **do**
    Generate new sample $\boldsymbol{m_{k+1}}$ from $q(\boldsymbol{m_k}, \cdot)$;
    Compute $\log(\boldsymbol{\pi_{\text{post}}}(\boldsymbol{m_{k+1}}|\boldsymbol{d}))$ from the PCE surrogate via (7.15);
    Calculate the correction factor $c = \frac{q(\boldsymbol{m_{k+1}}, \boldsymbol{m_k})}{q(\boldsymbol{m_k}, \boldsymbol{m_{k+1}})}$;
    Compute $\alpha(\boldsymbol{m_k}, \boldsymbol{m_{k+1}})$ defined as

$$\alpha(\boldsymbol{m_k}, \boldsymbol{m_{k+1}}) = \log(\min\{1, \frac{\pi_{\text{post}}(\boldsymbol{m_{k+1}}|\boldsymbol{d})}{\pi_{\text{post}}(\boldsymbol{m_k}|\boldsymbol{d})} \cdot c\});$$

    Draw $u \sim \mathcal{U}([0, 1))$;
    if $\log(u) < \alpha(\boldsymbol{m_k}, \boldsymbol{m_{k+1}})$ **then**
        Accept: Set $\boldsymbol{m}_{k+1} = \boldsymbol{m_{k+1}}$;
    else
        Reject: Set $\boldsymbol{m}_{k+1} = \boldsymbol{m}_k$;
    end if
end for
Plot the PDF, $\pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d})$, and find the MAP points.

Chapter 7

Real-Time Dynamic Parameter Estimation using PCE-Based Bayesian Inference with Hybrid MCMC

7.1 Introduction

It is well-known that 1996 blackout in the Western Electricity Coordinating Council (WECC) system was caused by severe model discrepancy [65]. This shows that the parameter estimation for dynamic generator models is of paramount importance for power system dynamic security analysis. Ignoring it will lead to inappropriate control actions, which in turn may result in power system failures.

However, the traditional offline stage-testing-based methods to conduct dynamic parameter estimation are very costly, time-consuming, and labor intensive [68]. This motivates more and more researchers to seek the online, low-cost phasor-measurement-unit (PMU)-based methods. Some of them focused on estimating the generator moment of inertia [69, 71]. While some of them are estimating several generator parameters or dynamic load parameters via Kalman filter [**?**,67,68,74]. Nevertheless, the former provides only point estimated values for the inertia parameters without providing the confidence intervals. The latter still suffers from several weaknesses of Kalman filters, such as the Gaussian assumption of the process and measurements noise and the convergence rate.

Unlike the above methods, Bayesian-inference based methods provide maximum-a-posteriori (MAP) parameter estimated values along with the corresponding posterior probability distributions [72], [77]. Under Bayesian Inference framework, the Non-Gaussian posterior distribution are typically obtained via MCMC based algorithms. Among them, the Metropolis-Hastings sampler may be the most widely used [81]. However, this algorithm places huge computational burden, especially when detailed generator models are adopted or when the

dimension of the parameters is large . To overcome it, Petra proposed to keep the approximated Gaussian assumption and take the local optimization approach [72]. Although the computing time is reduced, the local optima might be highly biased when the initial guess is far from the true value or the posterior distribution is highly non-Gaussian.

To overcome the aforementioned weaknesses, this part proposes a polynomial-chaos-expansion (PCE)-based Bayesian inference algorithm for power system dynamic parameter estimation using Hybrid MCMC algorithm. Here, we formulate the Bayesian framework for the parameter estimation of the decentralized generator model [79]. This eliminates the uncertainties of the line and transformer models and the loads within a centralized system model [16, 79]. In dealing with the non-Gaussian posterior distribution, the MCMC mechanism is adopted. To improve the computing efficiency of the MCMC procedure, we proposed two strategies:

- The time-consuming dynamical solver are represented by the state-of-the-art PCE-based surrogate, also known as the response surface or the reduced-order modeling [4]. This surrogate allows us to analyze the time-consuming dynamic solver at parameter values with negligible cost. Under PCE framework, Analysis of Variance (ANOVA) can quickly screen out the key parameters that we can estimate within the system observability.
- To reduce the large sample size required in the M-H algorithm, a two-stage hybrid MCMC algorithm is proposed. The first stage is based on the Gradient-based Langevin MCMC (LMCMC) algorithm. This algorithm enables a better convergence rate when bad priors are given. Its gradient information can be obtained easily from the PCE coefficients. The second stage is based on the Adaptive MCMC algorithm. This algorithm enables an accurate estimation when strong correlations between parameters exist.

Based on these strategies, the proposed method can achieve an accurate real-time dynamic parameter estimation. The simulation results carried out on the New-England system reveal that our method can accurately and simultaneously estimate several key generator dynamic model parameters with almost three-magnitude improved in computing speed compared with the traditional MCMC-based Bayesian inference method with full dynamic model involved.

7.2 Problem Formulation

This section formulates the Bayesian inference framework for the decentralized dynamic parameter estimation.

Decentralized Generator Model

The decentralized synchronous generator model, its the measurement model and the notations are all following the work of Zhao and Mili [79], where the two-axis model with a IEEE-DC1A exciter and a TGOV1 turbine-governor is considered [44]. Here, when a disturbance happens in the system, the local PMU ith generator will record its measurement, namely $V_i\angle\theta_i$ and $I_i\angle\phi_i$. Then we have $V_{di} = V_i\sin(\delta_i - \theta_i)$, $V_{qi} = V_i\cos(\delta_i - \theta_i)$, $I_{di} = (E'_{qi} - V_{qi})/X'_{di}$ and $I_{qi} = (V_{di} - E'_{di})/X'_{qi}$. This enables us to calculate the active power and reactive power for the ith generator as model outputs expressed as

$$P_{ei} = V_{di}I_{di} + V_{qi}I_{qi} + e_{Pi}, \tag{7.1}$$

$$Q_{ei} = -V_{di}I_{qi} + V_{qi}I_{di} + e_{Qi}, \tag{7.2}$$

where e_{Pi} and e_{Qi} are the measurement noise. By this way, once we capture the local voltage phasor $V_i\angle\theta_i$ for the ith generator in the selected time period, given its corresponding generator parameters, we can obtain the trajectories of its active and reactive power by calculating the ith generator model's output using the differential and algebraic equations given in [79]. This is the decentralized generator model used in this part.

Bayesian Inference

Tarantola [81] formulated the Bayesian inference model as

$$\boldsymbol{d} = \boldsymbol{f}(\boldsymbol{m}) + \boldsymbol{e}, \tag{7.3}$$

where $\boldsymbol{d}$ contains the observations, which consists of d_1 that contains the active power measurements, P_{ei}, and d_2 that contains the reactive power measurements, Q_{ei}; $\boldsymbol{m} \in \mathbb{R}^N$ are the parameters to be estimated; N is the number of parameters to be estimated that depends on the specific applications; $\boldsymbol{f(\cdot)}$ is the vector-valued forward function that represents the aforementioned decentralized generator model, which map the model parameters $\boldsymbol{m}$ to the observations $\boldsymbol{d}$; $\boldsymbol{e} \in \mathbb{R}^2$ stands for the measurement error vector whose components are assumed to be mutually independent random variables with the joint probability density functions $\pi_{\boldsymbol{e}}$ defined as $\pi_{\boldsymbol{e}} = \prod_{i=1}^{2} \pi_{e_i}(e_i)$. In the Bayesian inference, each parameter m_i is also viewed as a random variable with a given prior probability distribution, whose probability density function (pdf) is denoted as $\pi_i(m_i)$. The corresponding joint prior density function for a vector $\boldsymbol{m}$ is given by $\pi_{\text{prior}}(\boldsymbol{m}) = \prod_{i=1}^{N} \pi_i(m_i)$. Note that here $\boldsymbol{e}$ and $\boldsymbol{m}$ are also assumed to be mutually independent. Given the observation $\boldsymbol{d}$, the posterior pdf $\pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d})$ for the parameters $\boldsymbol{m}$ is derived as [81]

$$\pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d}) \propto \pi_{\text{like}}(\boldsymbol{d}|\boldsymbol{m})\pi_{\text{prior}}(\boldsymbol{m}). \tag{7.4}$$

Here $\pi_{\text{like}}(\boldsymbol{d}|\boldsymbol{m})$ denotes the likelihood function, expressed as

$$\pi_{\text{like}}(\boldsymbol{d}|\boldsymbol{m}) = \prod_{i=1}^{2} \pi_{e_i}(d_i - f_i(\boldsymbol{m})). \tag{7.5}$$

Given a set of parameters contained in $\boldsymbol{m}$, we obtain the trajectories of P_{ei} and Q_{ei} from the forward solver $\boldsymbol{f}(\cdot)$. By comparing them to the PMU metered values for the simulated time period t_{end}, the likelihoods for the corresponding trajectories are evaluated. Let us denote $\pi_{e_i}^t$, d_i^t and f_i^t as the likelihood, the observation and the realization at time t, respectively. The likelihood function for the trajectories in the log-form is then expressed as

$$\log \pi_{e_i}(d_i - f_i(\boldsymbol{m})) = \sum_{t=0}^{t_{\text{end}}} \log \pi_{e_i}^t(d_i^t - f_i^t(\boldsymbol{m})). \tag{7.6}$$

Thanks to the high speed sampling rate of the PMUs, which is typically equal to 60 samples/s, for a short time period such as 3 seconds, we get 180 samples for P_{ei} and for Q_{ei}, which provide good tracking information of the dynamic responses of a power system following a disturbance. Now the relationship given by (7.4) can be put into the following form:

$$\log(\pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d})) \propto \sum_{i=1}^{2} \sum_{t=0}^{t_{\text{end}}} \log \pi_{e_i}^t(d_i^t - f_i^t(\boldsymbol{m})) + \sum_{i=1}^{N} \log(\pi_i(m_i)), \tag{7.7}$$

yielding the MAP estimator defined as

$$\hat{\boldsymbol{m}}_{\text{MAP}} = \arg\min_{\boldsymbol{m}} \quad \{-\log(\pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d}))\}. \tag{7.8}$$

Note that due to the non-linearity of $\boldsymbol{f}(\cdot)$, even if the prior assumption is Gaussian, $\pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d})$ may be non-Gaussian. This motivates us to use the MCMC method instead of simply making the Gaussian assumptions for the posterior pdfs of the parameters $\boldsymbol{m}$. The Hybrid MCMC that we proposed will be described in Section IV.

7.3 PCE-Based Bayesian Inference

Here, we introduce to way to build PCE-based surrogate to represent original forward dynamic solvers.

7.3.1 Review of the Generalized Polynomial Chaos Expansion

Introduced by Wiener and further developed by Xiu and Karniadakis [4], the generalized polynomial chaos expansion has been shown to be a cost-effective tool in modeling response surfaces [82], [83], [17]. In the gPC 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 $\boldsymbol{\xi} = [\xi_1, \xi_2, ...\xi_N]$ be a vector of random variables following a standard probability distribution (e.g. the Gaussian or the beta distribution), to which, as shown in [4], a unique orthogonal polynomial is associated.

Let $\Phi_i(\xi_1, \xi_2, ...\xi_N)$ denote this procedure's corresponding polynomial chaos basis and let a_i denote the ith polynomial chaos coefficient. Formally, we have

$$z = \sum_{i=0}^{N_P} a_i \Phi_i(\boldsymbol{\xi}), \tag{7.9}$$

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. It is found that a relatively low maximum polynomial chaos order, typically 2, is found to provide output results with enough accuracy [83], [17], [15]. From the polynomial chaos coefficients, the mean, μ, and the variance, σ^2, of the output z can be determined as follows

$$\mu = a_0, \quad \sigma^2 = \sum_{i=1}^{N_P} a_i^2 E[\Phi_i^2]. \tag{7.10}$$

where $E[.]$ is the expectation operator.

7.3.2 PCE-Based ANOVA

Not all the parameters can be analyzed by the proposed method due to the observability limitation [80]. Some parameters have negligible influences on the model output, P_{ei} and Q_{ei}, that make the accurate prediction impossible. To address this issue, we propose to use the ANOVA index to screen out the key model parameters. The ANOVA expansion represents a function y with N random variable in the form of as

$$\begin{aligned} y(\xi_1, \ldots, \xi_N) = y_{j_0} + \sum_{1 \le j_1 \le N} y_{j_1}(\xi_{j_1}) \\ + \sum_{1 \le j_1 < j_2 \le N} y_{j_1,j_2}(\xi_{j_1}, \xi_{j_2}) + \ldots + y_{1,2,\ldots,N}. \end{aligned} \tag{7.11}$$

The functions $y_{j_k}(\xi_{j_k})$, $1 \le j_k \le N$, $y_{j_k,j_l}(\xi_{j_k}, \xi_{j_l})$, $1 \le j_k \le j_l \le N$, etc., are called the first-order, second-order,... ANOVA components, respectively. Let $T_{j_1,\ldots,j_s}$ denote the fraction of the variance $\sigma^2(y)$ that is contributed by $y_{j_1,\ldots,j_s}$, which is defined as

$$T_{j_1,\ldots,j_s} = \frac{\sigma^2(y_{j_1,\ldots,j_s})}{\sigma^2(y)}. \tag{7.12}$$

Here T_{j_1} denotes the variance contributed by the single variable ξ_{j_1}, T_{j_1,j_2} denotes the variance contributed by the coupling effect of the two random inputs, (ξ_{j_1}, ξ_{j_2}), and so on for the higher-order terms.

Generally, ANOVA is a Monte-Carlo based method that are thought to be time consuming. However, Sudret [88] proposes the ANOVA index under the PCE framework as follows:

$$T_{j_1} = \frac{\sigma^2(y_{j_1})}{\sigma^2(y)} = \frac{a_i^2 E[\phi(\xi_i)^2] + a_{i,j}^2 E[\phi(\xi_i, \xi_j)^2]_{i=j}}{\sigma^2}, \tag{7.13}$$

$$T_{j_1,j_2} = \frac{\sigma^2(y_{j_1,j_2})}{\sigma^2(y)} = \frac{a_{i,j}^2 E[\phi(\xi_i,\xi_j)^2]_{i\neq j}}{\sigma^2}. \tag{7.14}$$

This means once we obtained the PCE coefficients, we can obtain ANOVA index with almost no computing cost. T_{j_1} is typically chosen to measure the importance of the random variables. If T_{j_1} is too small, such as less than 0.1%, we call it an ineffective dimension that can not be estimated. T_{j_1,j_2} is typically far smaller than T_{j_1}.

7.3.3 Incorporating Polynomial Chaos Expansion into the Bayesian Inference Framework

In the PCE-based Bayesian inference [82,83], we use the approximated gPC solution in (7.9) to replace the exact forward solver solution $\boldsymbol{y(m)}$ in (7.7) as

$$\log(\pi_{\text{post}}(\boldsymbol{\xi}|\boldsymbol{d})) \propto \sum_{i=1}^{2}\sum_{t=0}^{t_{\text{end}}} \log \pi_{e_i}^t(d_i^t - z_i^t(\boldsymbol{\xi})) + \sum_{i=1}^{N} \log(\pi_i(\xi_i)). \tag{7.15}$$

Once a sample point $\boldsymbol{\xi}$ is proposed in MCMC, we can evaluate $\boldsymbol{\xi}$ with the PCE-based response surface without resorting to actual simulations of the forward solver. Therefore, we achieve very high accuracy in sampling the posterior distribution at a much less computational cost. Note that, after obtaining the posterior distribution of $\boldsymbol{\xi}$, we need to use (6.25) to map $\boldsymbol{\xi}$ back into $\boldsymbol{m}$ to obtain the posterior distribution of $\boldsymbol{m}$.

7.4 Application of the Proposed Method

This section shows the proposed PCE-based Bayesian inference via the Hybrid MCMC algorithm for the dynamic parameter estimation. Originated from the classical M-H algorithm, we further developed a hybrid MCMC algorithm by incorporating the advantages of the Langevin MCMC algorithm and the adaptive MCMC algorithm . We have also extend this algorithm into the PCE framework. This algorithm can greatly accelerate the M-H algorithm when the prior knowledge is poor and be able to overcome the strong correlation between parameters.

7.4.1 Review of MCMC Methods

Metropolis-Hastings (M-H) algorithm

As the most classical MCMC algorithm, starting from the initial guess, $\boldsymbol{m_0}$, the M-H method employs a proposal distribution, $q(\boldsymbol{m_k},\cdot)$, at each sample point $\boldsymbol{m}_k$ to generate a proposed sample point $\boldsymbol{m_{k+1}}$. Once generated, the sample point is either accepted or rejected by the

M-H algorithm. This procedure is then applied to the next sample point, yielding a chain of sample points that provide the posterior pdf $\pi_{\text{post}}(\boldsymbol{m})$ [81]. This procedure is described in Algorithm 1.

Algorithm 5 Metropolis-Hastings Algorithm to sample pdf π

Choose the initial parameters $\boldsymbol{m_0}$;
Compute $\pi(\boldsymbol{m_0})$
for $k = 0, \ldots, N_{\text{samples}} - 1$ **do**
    Generate new sample $\boldsymbol{m_{k+1}}$ from $q(\boldsymbol{m_k}, \cdot)$,
    Compute $\pi_{post}(\boldsymbol{m_{k+1}})$
    Calculate the correction factor $c = \frac{q(\boldsymbol{m_{k+1}}, \boldsymbol{m_k})}{q(\boldsymbol{m_k}, \boldsymbol{m_{k+1}})}$;
    Compute $\alpha(\boldsymbol{m_k}, \boldsymbol{m_{k+1}}) = \min\{1, \frac{\pi_{\text{post}}(\boldsymbol{m_{k+1}}|\boldsymbol{d})}{\pi_{\text{post}}(\boldsymbol{m_k}|\boldsymbol{d})} \cdot c\}$
    Draw $u \sim \mathcal{U}([0, 1))$;
    if $u < \alpha(\boldsymbol{m_k}, \boldsymbol{m_{k+1}})$ **then**
        Accept: Set $\boldsymbol{m}_{k+1} = \boldsymbol{m_{k+1}}$;
    else
        Reject: Set $\boldsymbol{m}_{k+1} = \boldsymbol{m}_k$;
    end if
end for

Langevin MCMC algorithm (LMCMC)

Unlike the M-H algorithm that simply uses the proposal function driven by random variables to approach the posterior distribution asymptotically, the LMCMC algorithm uses a combination of two mechanisms to generate the samples of a random walk [84, 85]. It first proposes new samples using Langevin dynamics mechanism, which uses the evaluations of the gradient of the target probability density function, then accept or reject new samples using the M-H mechanism. This process can be described as

$$\boldsymbol{m_{k+1}} = \boldsymbol{m_k} + \frac{\varepsilon^2}{2} \bigtriangledown \log(\pi_{\text{post}}(\boldsymbol{m_k}|\boldsymbol{d})) + \varepsilon \boldsymbol{\rho}. \tag{7.16}$$

Here $\varepsilon > 0$ is a fixed step size, $\bigtriangledown \log(\pi_{\text{post}}(\boldsymbol{m_k}|\boldsymbol{d}))$ represents the gradient information for the Langevin dynamics and $\boldsymbol{\rho}$ is an independent draw from a multivariate probability distribution that brings the randomness for M-H mechanism. This Langevin dynamics enables the LMCMC a better convergence rate than the M-H algorithm. This benefits tends to be more obvious when the dimension increases [85]. This is specially the case for the decentralized dynamic parameter estimation problem, where several generator dynamic parameters can influence the dynamic performances of the systems.

However, the LMCMC algorithm also suffers several drawbacks. First, analyzing the gradient information takes extra computing time. Second, the selection of the step size ε will influence

the searching path for the global optimum. Once the LMCMC sampler is approaching the local optimums, it may not jumper out from the local optimum when ε is too small. In contrast, when ε is too large, the accept rate will be very low. To obtain unbiased results, the sample size must be increased significantly. Last, both M-H algorithm and LMCMC algorithm cannot overcome the strong correlation between parameters

Adaptive MCMC

The basic idea of the AM strategy is to create a Gaussian proposal distribution function with a covariance matrix calibrated using the previous samples in the MCMC chain. The key point of the AM adaption is how the covariance of the proposal distribution function depends on the history of the chain. After an initial non-adaption period, the Gaussian proposal function is centered at the current position of the Markov Chain, $\boldsymbol{m_k}$, and reset its covariance C_k. This procedure can be expressed as

$$C_k = \begin{cases} C_0 & k \leq k_0 \\ s_N \mathrm{Cov}(\boldsymbol{m_0}, \ldots, , \boldsymbol{m_{k-1}}) + s_N \epsilon I_N & k > k_0 \end{cases} \tag{7.17}$$

Here C_0 denotes initial covariance chosen according to a prior knowledge, which may be quite poor; k_0 denotes a time index that defines this initial non-adaptive period. Obviously, the larger k_0 is, the longer it takes for the adaption to take place; s_N denotes the scaling parameter set to be $2.4^2/N$ [86]; $\epsilon > 0$ is a constant that we may choose very small to ensure the non-singularity of C_k; I_N denotes N dimensional identity matrix. This adaption mechanism enables the AM algorithm to overcomes the local optimums caused by the strong correlation between parameters and to choose the appropriate variance values for the proposal function. It is suggested in [87] that the adaption should not be done at each time step, but only at given time intervals, N_{interval}.

The benefits of this AM strategy are clear. Once the strong correlation between parameters exist, the accept rate of the M-H algorithm and LMCMC algorithm tends to be very low. Therefore, the local optimum caused by strong correlations will be formed. This may lead to highly-biased results. This can be overcome by AM strategy. However, it may be difficulty to get the adaption process started if the initial guess for the proposal distribution is far from a correct one. The long burn-in period will destroy the adaption period and make this process cannot get started.

7.4.2 Proposed Hybrid MCMC algorithm

In this part, we combine the LMCMC algorithm and the AM strategy to make the use of the advantages of both methods. Note that we are not merging AM strategy into LMCMC framework, but propose a two-stage hybrid MCMC algorithm that use LMCMC algorithm and AM strategy at different time period separately. The first stage is the LMCMC stage,

which take use of the fast converge rate of the LMCMC. Under the PCE framework, it can converge to a reasonable range with a small sample size, N_{LMCMC}, in a short time period. Later on, the MCMC sampler is switched to adopt the AM strategy. The gradient information in this period is not analyzed anymore. The adaption procedure is executed at given time intervals. By this way, the LMCMC can help AM strategy to overcome its shortcoming in dealing with bad priors and the AM strategy can help LMCMC algorithm handle parameter correlation, reduce computing time, and provide global optimum results.

PCE-based LMCMC

Here, let's first derive the LMCMC under PCE framework. From (7.15), we can see that the original complicated dynamic solver at $\boldsymbol{m}$ space is simplified into the polynomial forms at $\boldsymbol{\xi}$ space. This can greatly simplify the derivation of the gradients of the posterior likelihood. Let us assume the $\boldsymbol{e} \sim \mathcal{MVN}(\mathbf{0}, \Sigma_e)$ where Σ_e is a diagonal matrix $diag\{\sigma_{e1}^2, \sigma_{e2}^2\}$. Under Gaussian assumption for the prior $\boldsymbol{m}$, Normalized Hermite polynomials [4] are chosen for $\boldsymbol{\xi}$, which means $\boldsymbol{\xi} \sim \mathcal{MVN}(\mathbf{0}, \Sigma_\xi)$ and Σ_ξ is a $N \times N$ identity matrix. Thus, (7.15) is expressed as

$$\log(\pi_{\text{post}}(\boldsymbol{\xi}|\boldsymbol{d})) \propto \sum_{i=1}^{2} \sum_{t=0}^{t_{\text{end}}} -\frac{(d_i^t - z_i^t(\boldsymbol{\xi}))^2}{2{\sigma_{e_i}}^2} + \sum_{i=1}^{N} -\frac{(\xi_i)^2}{2}. \tag{7.18}$$

Then we derive the gradients as

$$\bigtriangledown \log(\pi_{\text{post}}(\boldsymbol{\xi}|\boldsymbol{d})) = \sum_{i=1}^{2} \sum_{t=0}^{t_{\text{end}}} \frac{(d_i^t - z_i^t(\boldsymbol{\xi}))}{{\sigma_{e_i}}^2} \bigtriangledown z_i^t(\boldsymbol{\xi}) + \boldsymbol{g}(\boldsymbol{\xi}). \tag{7.19}$$

Here $\boldsymbol{g}(\boldsymbol{\xi})$ represents the gradients from prior information.

$$\boldsymbol{g}(\boldsymbol{\xi}) = \begin{bmatrix} -\xi_1 & -\xi_2 & \dots & -\xi_N \end{bmatrix}^T; \tag{7.20}$$

Now, let us derive the gradient of $z_i^t(\boldsymbol{\xi})$. Following [15], let us reorganize (7.9) with $P = 2$ at time point t as

$$z_i^t = a_0^t + (\sum_{i=1}^{N} a_i^t \xi_i + \sum_{i=1}^{N} a_{i,i}^t \frac{{\xi_i}^2 - 1}{\sqrt{2}} \quad + \sum_{j=1}^{N} \sum_{1 \le i < j} a_{i,j}^t \xi_i \xi_j. \tag{7.21}$$

From this simple polynomial form, we can obtain $\bigtriangledown z_i^t$ via

$$\bigtriangledown z_i^t = \begin{bmatrix} \frac{\partial z_i^t}{\partial \xi_1} \\ \frac{\partial z_i^t}{\partial \xi_2} \\ \vdots \\ \frac{\partial z_i^t}{\partial \xi_N} \end{bmatrix} = \begin{bmatrix} a_1^t \\ a_2^t \\ \vdots \\ a_N^t \end{bmatrix} + \sqrt{2} \begin{bmatrix} a_{1,1}^t \xi_1 \\ a_{2,2}^t \xi_2 \\ \vdots \\ a_{N,N}^t \xi_N \end{bmatrix}$$
$$+ \begin{bmatrix} 0 & a_{1,2}^t & \dots & a_{1,N-1}^t & a_{1,N}^t \\ a_{1,2}^t & 0 & \dots & a_{2,N-1}^t & a_{2,N}^t \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{1,N-1}^t & a_{2,N-1}^t & \dots & 0 & a_{N-1,N}^t \\ a_{1,N}^t & a_{2,N}^t & \dots & a_{N-1,N}^t & 0 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_{N-1} \\ \xi_N \end{bmatrix}; \tag{7.22}$$

This form remains the same for both active P and reactive power Q with different PCE coefficients. Based on (7.16), we can propose new sample ξ_{k+1} from current sample ξ_k via PCE- based LMCMC as

$$\boldsymbol{\xi_{k+1}} = \boldsymbol{\xi_k} + \frac{\varepsilon^2}{2} \bigtriangledown \log(\pi_{\text{post}}(\boldsymbol{\xi_k}|\boldsymbol{d})) + \varepsilon \boldsymbol{\rho}. \tag{7.23}$$

Note that the derivation in this part is under Gaussian assumption for $\boldsymbol{e}$. However, the Bayesian inference is not restricted to Gaussian noise. When the pdf of noise follows other distribution, such as the Laplace distribution, the gradients can be derived through the similar procedure.

Application in the Decentralized Parameter Estimation using LMCMC

Now, we can formulate the decentralized parameter estimation with the PCE-based LMCMC. The detailed procedure is summarized in Algorithm 1. In the latter, the most time-consuming step is Step 7. Using a full dynamic solver brings prohibitive computational burden. This is greatly accelerated by the PCE-based surrogate. Compared with the MH method that needs a large sample size to converge when the prior is highly biased, the LMCMC can converge faster with a much smaller sample size. All these can make the MCMC-based algorithm fast enough for online applications.

Application in Decentralized Dynamic Parameter Estimation using Hybrid MCMC

Here, we can formulate the decentralized parameter estimation with Hybrid MCMC under the PCE framework. The detailed procedure is summarized in Algorithm 2. Here, the proposal function is selected as Gaussian distribution. Then the correction factor in step 6 of Algorithm 1 is simplified into 1. The first LMCMC stage is assigned with N_{LMCMC}

Algorithm 6 Polynomial-Chaos-Expansion-based Bayesian Inference using the Langevin MCMC

Choose the initial guess of the parameters $\boldsymbol{m_0}$ from the manufactured data as the Bayesian prior $\boldsymbol{m}_{\text{prior}}$;
Mapping $\boldsymbol{m}$ into $\boldsymbol{\xi}$ via (6.25), build the PCE surrogates as the response surface of the decentralized dynamic model;
Compute $\log(\pi_{\text{post}}(\boldsymbol{\xi_0}|\boldsymbol{d}))$ from the PCE surrogate via (7.15);
for $k = 0, \ldots, N_{\text{samples}} - 1$ **do**
    Compute $\nabla \log(\pi_{\text{post}}(\boldsymbol{\xi_k}|\boldsymbol{d}))$ from (7.19);
    Generate a new sample $\boldsymbol{\xi_{k+1}}$ from (7.23);
    Compute $\boldsymbol{\log\left(\pi_{\text{post}}(\xi_{k+1}|d)\right)}$ from the PCE surrogate via (7.15);
    Calculate the correction factor $c = \frac{q(\boldsymbol{\xi_{k+1}}, \boldsymbol{\xi_k})}{q(\boldsymbol{\xi_k}, \boldsymbol{\xi_{k+1}})}$;
    Compute $\alpha(\boldsymbol{\xi_k}, \boldsymbol{\xi_{k+1}})$ defined as

$$\alpha(\boldsymbol{\xi_k}, \boldsymbol{\xi_{k+1}}) = \log(\min\{1, \frac{\pi_{\text{post}}(\boldsymbol{\xi_{k+1}}|\boldsymbol{d})}{\pi_{\text{post}}(\boldsymbol{\xi_k}|\boldsymbol{d})} \cdot c\});$$

    Draw $u \sim \mathcal{U}([0, 1))$;
    if $\log(u) < \alpha(\boldsymbol{\xi_k}, \boldsymbol{\xi_{k+1}})$ **then**
        Accept: Set $\boldsymbol{\xi}_{k+1} = \boldsymbol{\xi_{k+1}}$;
    else
        Reject: Set $\boldsymbol{\xi}_{k+1} = \boldsymbol{\xi}_k$;
    end if
end for
Mapping samples of $\boldsymbol{\xi}$ back to $\boldsymbol{m}$, then plot the PDF of $\pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d})$, and find the MAP points.

Algorithm 7 PCE-based Bayesian Inference using the Hybrid MCMC for Decentralized Dynamic Parameter Estimation

1: Choose the initial guess of the parameters $\boldsymbol{m_0}$ from the manufactured data as the Bayesian prior $\boldsymbol{m}_{\text{prior}}$;
2: Mapping $\boldsymbol{m}$ into $\boldsymbol{\xi}$ via (6.25), build the PCE surrogates as the response surface of the decentralized dynamic model;
3: Computing ANOVA index via (7.13). Find the effective dimensions and ignore the ineffective dimensions;
4: Compute $\log(\pi_{\text{post}}(\boldsymbol{\xi_0}|\boldsymbol{d}))$ from via (7.15);
5: **for** $k = 0, \ldots, N_{\text{LMCMC}} - 1$ **do**
6: Compute $\bigtriangledown \log(\pi_{\text{post}}(\boldsymbol{\xi_k}|\boldsymbol{d}))$ from (7.19);
7: Generate new sample $\boldsymbol{\xi_{k+1}}$ from (7.23);
8: Compute $\boldsymbol{\log\left(\pi_{\text{post}}(\xi_{k+1}|d)\right)}$ via (7.15);
9: Compute $\alpha(\boldsymbol{\xi_k}, \boldsymbol{\xi_{k+1}})$ defined as

$$\alpha(\boldsymbol{\xi_k}, \boldsymbol{\xi_{k+1}}) = \log(\min\{1, \frac{\pi_{\text{post}}(\boldsymbol{\xi_{k+1}}|\boldsymbol{d})}{\pi_{\text{post}}(\boldsymbol{\xi_k}|\boldsymbol{d})}\});$$

10: Draw $u \sim \mathcal{U}([0, 1))$;
11: **if** $\log(u) < \alpha(\boldsymbol{\xi_k}, \boldsymbol{\xi_{k+1}})$ **then**
12: Accept: Set $\boldsymbol{\xi}_{k+1} = \boldsymbol{\xi_{k+1}}$;
13: **else**
14: Reject: Set $\boldsymbol{\xi}_{k+1} = \boldsymbol{\xi}_k$;
15: **end if**
16: **end for**
17: **for** $k = N_{\text{LMCMC}}, \ldots, N_{\text{samples}} - 1$ **do**
18: Update covariance C_k for proposal distribution via (7.17) based the samples generated in the AM stage with the interval set as N_{interval};
19: Generate new sample $\boldsymbol{\xi_{k+1}}$ from $q(\boldsymbol{\xi_k}, \cdot)$ with its covariance C_k;
20: Compute $\boldsymbol{\log\left(\pi_{\text{post}}(\xi_{k+1}|d)\right)}$ via (7.15);
21: Repeat M-H mechanism shown in Step 9 to Step 15;
22: **end for**
23: Mapping samples of $\boldsymbol{\xi}$ back to $\boldsymbol{m}$, then plot the PDF of $\pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d})$, and find the MAP points.

samples and the total sample size for the two stage is N_{samples}. The AM stage update covariance matrix at given intervals, N_{interval}.

In this algorithm, the most time-consuming period happens in step 7 and 20. Using a full dynamic solver brings prohibitive computational burden. This is greatly accelerated by the PCE-based surrogate. Furthermore, compared with the M-H algorithm that needs a large sample size to converge, the hybrid MCMC algorithm we proposed can converge faster by using the LMCMC method at the first stage and be able to overcome the parameter correlation by using the AM strategy at the second stage. Therefore, the hybrid MCMC algorithm can greatly reduce the sample size required by M-H algorithm while providing better estimation results. All these make this method applicable for real-time implementation.

Chapter 8

Case Studies

This chapter demonstrates the case studies conducted for Chapter 3 to Chapter 7. Section 8.1 demonstrates the case studies for probabilistic power flow and variance analysis with the proposed polynomial-chaos-ANOVA method that described in Chapter 3. Section 8.2 demonstrates the case studies for statistical power system dynamic simulation using generalized polynomial chaos method and the multi-element polynomial chaos method that described in Chapter 4. Section 8.3 demonstrates the case studies conducted for the power system dynamic state estimation with the proposed polynomial-chaos-based Kalman filter that described in Chapter 5. Section 8.4 demonstrates the cases studies conducted for the decentralized power system dynamic parameter estimation with the proposed response-surface based Bayesian inference method that described in Chapter 6. Following section 8.4, section 8.5 will further improve the proposed method using a response-surface based Bayesian inference with a hybrid MCMC strategy that described in Chapter 7.

8.1 Case Studies for Probabilistic Power Flow and Variance Analysis Problem

All the case studies were conducted in the MATPOWER software and the power flow is solved by the default Newton Method. In terms of gPC method, no specific solver is required.

8.1.1 Simulation Results on the IEEE 118-bus System

This case study is conducted using the IEEE 118-bus system provided with correlated power generation outputs from six renewable energy resources as shown in Table 8.1 and 8.2 [36]. 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 5% of their means.

Table 8.1: Information Renewable Energy Units

Renewable	Capacity	Location	Probability distribution
Solar 1	150 MW	Bus 13	Beta(α=2.06,β=2.5)
Wind 1	200 MW	Bus 43	Beta(α=1.98,β=3.52)
Wind 2	50 MW	Bus 44	Beta(α=1.98,β=3.52)
Solar 2	10 MW	Bus 82	Beta(α=2.06,β=2.5)
Solar 3	20 MW	Bus 83	Beta(α=2.06,β=2.5)
Wind 3	30 MW	Bus 115	Weibull(k=2.06, λ=7.41)

Table 8.2: Correlation matrix of the renewable energy units

	S 1	**W 1**	**W 2**	**S 2**	**S 3**	**W 3**
S 1	1	0.497	0.201	0.551	0.450	0.522
W 1	0.497	1	0.345	0.263	0.301	0.450
W 2	0.201	0.345	1	0.641	0.125	0.238
S 2	0.551	0.263	0.641	1	0.261	0.316
S 3	0.449	0.301	0.125	0.261	1	0.760
W 3	0.522	0.450	0.238	0.316	0.760	1

Table 8.3: The scenarios for gPC-ANOVA calculation efficiency

Scenario	Threshold (ϵ)	2nd-order basis (M)	CPU time
Case (a)	0.999	3	1.2031 s
Case(b)	0.999	20	3.3594 s

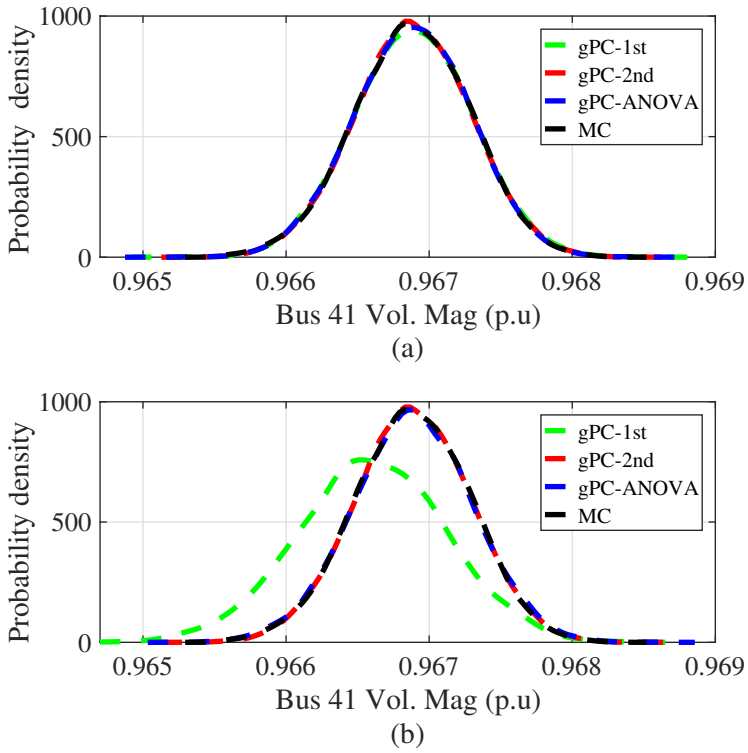


Figure 8.1: Plots for voltage magnitude at Bus 41: (a) probability distribution function of Bus 41 using the gPC-ANOVA 1st-order with accurate approximation, gPC-2nd-order, gPC-ANOVA, and MC methods; (b) probability distribution function of Bus 41 using the gPC-ANOVA 1st order with inaccurate approximation, gPC-2nd-order, gPC-ANOVA, and MC methods.

Accuracy

As an example, we provide the simulation results for the bus voltage magnitude at Bus 41. They are depicted in Fig. 3.1 and the calculation times and accuracy are provided in Table 8.3 and Table 8.4. Fig. 1 shows that the probability density function of bus voltage magnitude at Bus 41 of the gPC method (P=2) and the gPC-ANOVA method ($\epsilon = 0.999$) match that of the MC method with 20,000 samples. The green lines are the 1st-order gPC calculation that the gPC-ANOVA method makes use of to estimate to effective 2nd-order bases. As respectively shown in Fig. 1 (a) and (b), the 1st-order gPC calculation provides an accurate result for Case (a), but an inaccurate result for Case (b). By contrast, the gPC-ANOVA method provides accurate results for both cases.

Adaptivity

By comparing the results displayed in Table 8.3 with those displayed in Table 8.4, it is observed that in both Case (a) and (b), the gPC-ANOVA method is much faster than the gPC method (P=2). When the the 1st-order gPC calculation provides a good estimation, the gPC-ANOVA method adaptively chooses a smaller number of the 2nd-order basis functions, resulting in an improved computation efficiency. When the 1st-order gPC calculation provides an inaccurate estimation, the gPC-ANOVA method adaptively chooses a larger number of the 2nd-order basis functions, which decreases the computational efficiency.

Computational Efficiency

We execute 300 times the algorithms that implement the gPC method (P=2), the gPC-ANOVA method ($\epsilon = 0.99$) and the gPC-ANOVA method ($\epsilon = 0.999$). Their average computing times are listed in Table 8.4. We observe that the computing time of the gPC algorithm has been greatly improved when the gPC method is combined with the ANOVA method, that is, when using the gPC-ANOVA method proposed in this work. We also note that with a higher threshold ϵ, the computing time of the gPC-ANOVA algorithm increases. This is because with larger values of ϵ, more 2nd-order bases are selected, implying that more deterministic power flow calculations have to be carried out.

Table 8.4: Average CPU Times of the gPC 2nd-order, gPC-ANOVA, and MC Methods

Method	gPC(P=2)	gPC-A(0.99)	gPC-A(0.999)	MC
CPU time	4.4955 s	1.2178 s	1.46538 s	2028.5 s

8.1.2 Simulation Results on the European 1354-bus System

This case study is conducted on the European high-voltage transmission system, which consists of 1,354 buses, 260 generators, and 1,991 branches [34], [35]. Note that the slack bus is at bus 4231 in region 2. The loads are classified according to their degree of correlations; the highly correlated loads are grouped into one region, yielding 3 distinct regions. The loads are assumed to follow a Gaussian distributions with mean values equal to the original bus loads and standard deviations equal to 5%, 6% and 5% of their means for each of the three regions. In region one, there are seven wind farms with a capacity of 200 MW, 160 MW, 150MW, 140 MW, 60 MW, 120 MW and 100 MW, and two solar generation units with a capacity of 60 MW and 120 MW. In region two, there are five wind farms with a capacity of 260 MW, 80 MW, 130 MW, 160 MW, 120 MW and three solar generation units with a capacity of 200 MW, 120 MW and 60 MW. In region three, there are three wind farms with a capacity of 200 MW, 120 MW and 100 MW, and three solar generation units

with a capacity of 50 MW, 140 MW and 120 MW. The power generated by the renewable energy units (both solar and wind) are assumed to obey a beta distribution and to be highly correlated when they belong to the same region.

Accuracy

Let us randomly select separately as output variables the voltage magnitudes and phase angles on Bus 346, Bus 1435, Bus 4060 and Bus 6619. Their PDFs obtained by the gPC-ANOVA method ($\epsilon = 0.98$) are depicted in Fig. 3.2. The simulation results are compared with the results of the gPC method (P=2) and the MC method over 20,000 samples. Fig. 3.2 shows that the PDFs of the gPC method and the gPC-ANOVA method match very well the PDF of the MC method. They can provide accurate results even for very asymmetric curves, as the ones displayed in Fig. 3.2.(a). It is interesting to notice that Bus 4060 has a very small variation in the voltage magnitude but large variations in the voltage angle, see Fig. 8.2. (c)-(g). The reason is that Bus 4060 is very close to the swing bus, Bus 4231 and its angle has to vary to help compensate the system losses following the changes of the swing bus. As a result, it yields smaller confidence interval as compared to other 3 buses.

Trade-offs Between Computing Efficiency and Accuracy

For this application, we choose as an example the voltage magnitude at Bus 4060 to be the output and execute 200 times the gPC method (P=2), the gPC-ANOVA method with thresholds equal to $\epsilon = 0.95$, $\epsilon = 0.96$, $\epsilon = 0.97$, $\epsilon = 0.98$ and $\epsilon = 0.99$, in that sequence. In the simulations, we pre-generate 200 different groups of the collocation points. For each case, the algorithms that implement these methods are executed using one of the 200 groups of the collocation points that have been determined. The average calculation times, the average mean values, and the average standard deviation values taken by each of these methods are listed in Table 8.5. The results are compared with the results of the MC method with 20,000 samples, which are also taken as the ground truth for comparison. Table 8.5 shows that the gPC method greatly reduces the calculation time of the MC method. But the decrease in the computing time is not as good as that shown in Table 8.4. The gPC-ANOVA method further improves the computing time of the gPC method. With higher threshold values, its computational efficiency decreases, but remains better than that of the gPC method. This is because with higher threshold settings, more effective bases will be adaptively selected. This is demonstrated in Fig. 3.3. Therefore, we can say that with a higher threshold, more deterministic calculations need to be conducted, resulting in additional computing times. Furthermore, the mean values and the standard deviation values for all the methods reach high accuracy. In conclusion, the gPC-ANOVA method greatly improves the calculation efficiency of the gPC method without losing accuracy.

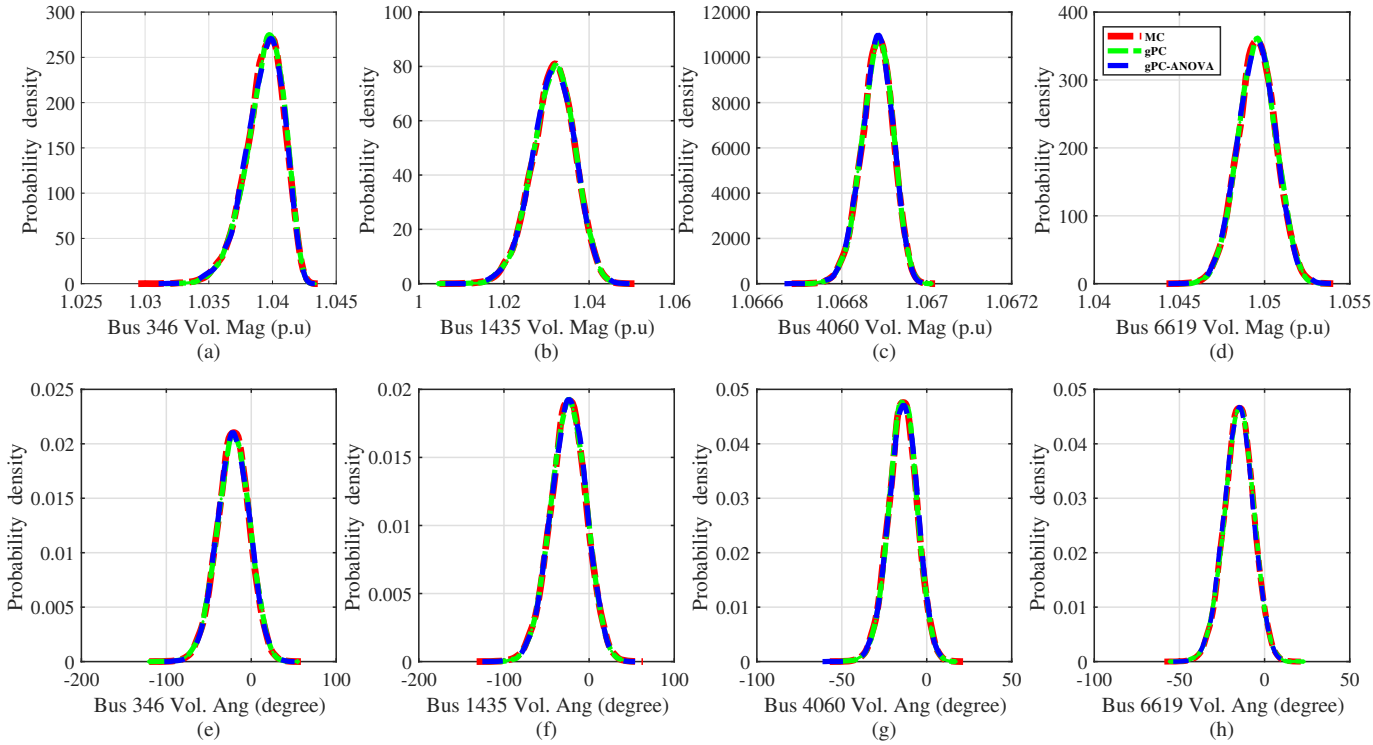


Figure 8.2: Plots of Bus magnitudes at (a) Bus 346; (b) Bus 1436; (c) Bus 4060; (d) Bus 6619 and Plots of Bus Angles at (e) Bus 346; (f) Bus 1436; (g) Bus 4060; (h) Bus 6619 with the MC, the gPC and the gPC-ANOVA methods.

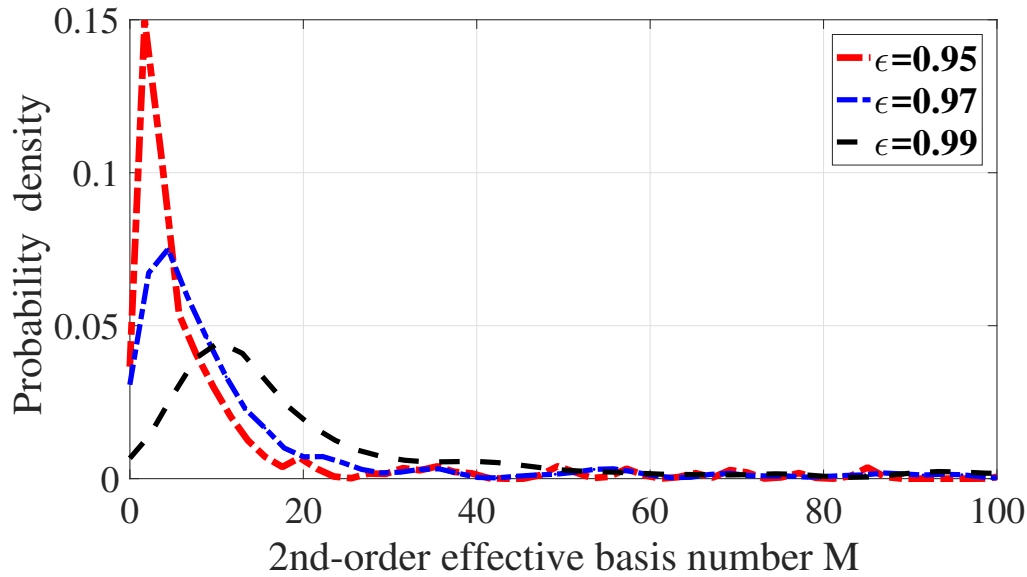
.

Table 8.5: Comparison between the gPC 2 nd-order, gPC-ANOVA, and MC Methods for Bus 4060 Vol. Mag. (p.u)

Method	CPU time (s)	μ (p.u)	standard deviation (p.u)
MC	2639.3	1.0669	3.6860e-05
gPC	46.7072	1.0669	3.6786e-05
gPC-A(0.99)	9.5034	1.0669	3.6826e-05
gPC-A(0.98)	7.9587	1.0669	3.6883e-05
gPC-A(0.97)	7.3052	1.0669	3.6836e-05
gPC-A(0.96)	6.9399	1.0669	3.6852e-05
gPC-A(0.95)	6.6661	1.0669	3.6947e-05

Variance Analysis

Fig 3.4 displays the plots of the variance fraction T_{j_1} and T_{j_1,j_2} of the voltage magnitude on Bus 346. It is observed that the T_{j_1} accounts for most of the variance while T_{j_1,j_2} contribution



(a)

Figure 8.3: Plots for the Probability density functions of the 2-nd order effective basis for the Bus 4060 Vol. Mag with different threshold values.

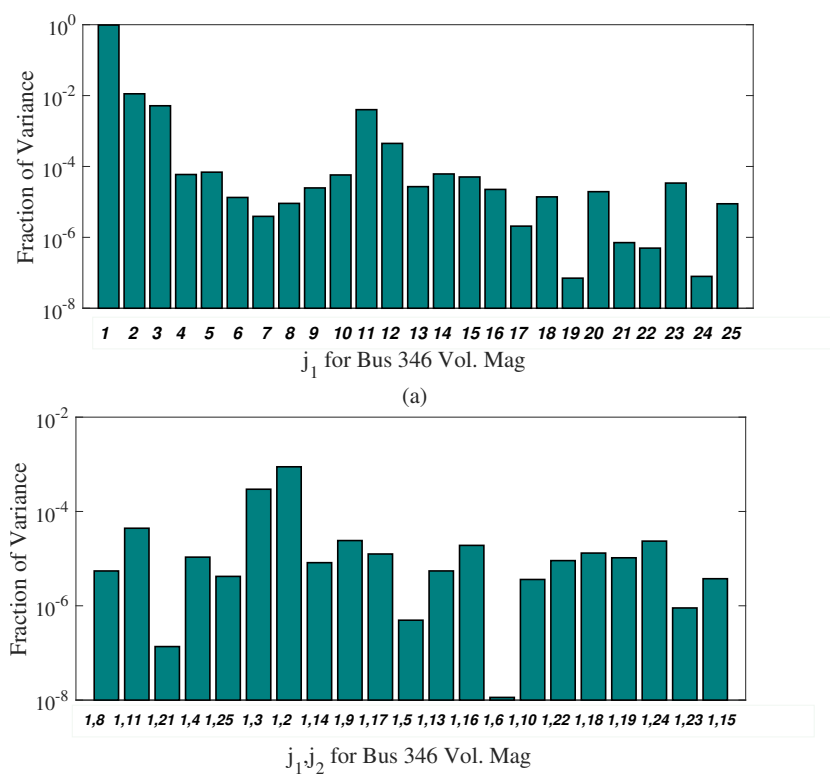


Figure 8.4: Histograms of the fraction of variance for (a) one-dimensional input variables; (2) effective dimensions with two random variables.

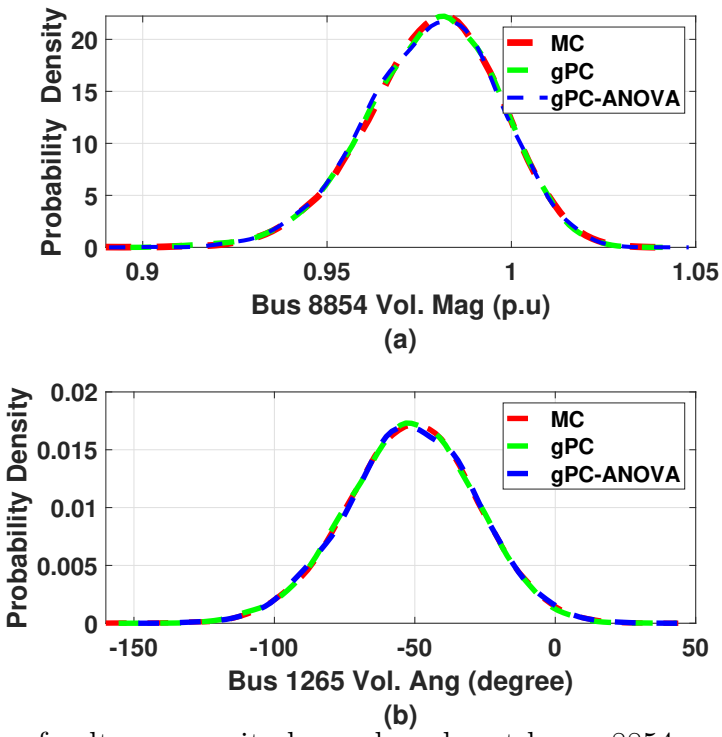


Figure 8.5: PDFs of voltage magnitudes and angles at buses 8854 and 1265 in (a) and (b), respectively.

is small. This demonstrates the good performance of the gPC-ANOVA method. Regarding the gPC method, it suffers from the "curse of dimensionality" because the two dimensional bases that are used only provide very small improvements to the output. By contrast, the gPC-ANOVA method selects the effective bases to alleviate the "curse of dimensionality".

Adaptive ANOVA Procedure

Fig. 3.5 displays the 2nd-order effective dimension sets $\{U_1, U_2, \ldots, U_M\}$ selected by the adaptive ANOVA procedure for different outputs. Let us take the voltage magnitude on Bus 346 as an example. Without adaptivity, the 2nd-order ANOVA decomposition will generate 300 sub-problems. However, using our adaptive algorithm with a threshold of 0.98, we only need to solve 11 sub-problems and with a higher threshold of 0.995, we only need to solve 23 sub-problems. The same tests are conducted while using the voltage magnitudes on the Bus 1435 and Bus 4060 that are in different regions than Bus 346. The simulation results show that while the effective dimension sets in different regions may be different, they always have a much smaller number compared with the 300 2nd-order sets that the gPC method needs.

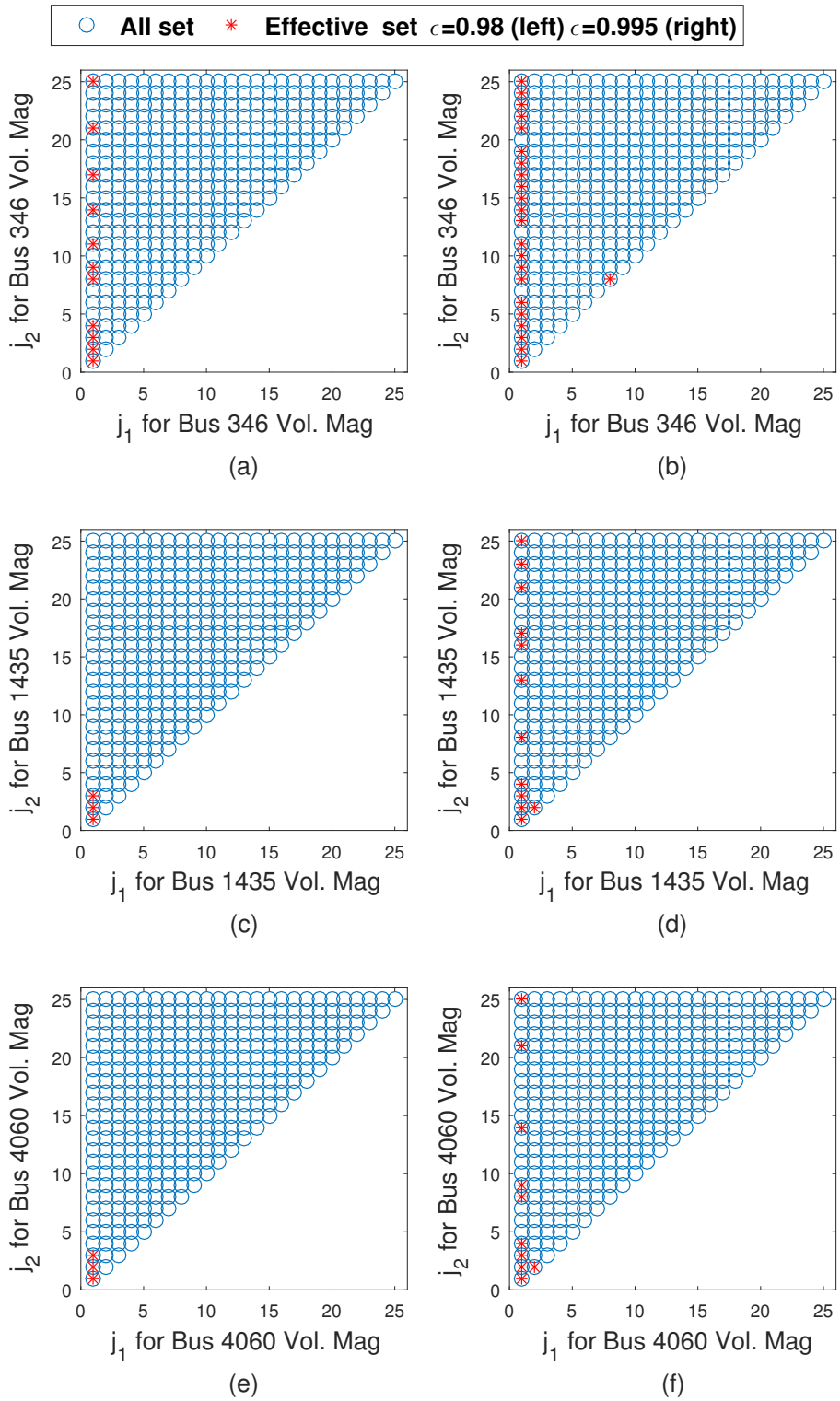


Figure 8.6: Scatter plots of effective dimension set for voltage magnitudes at (a) Bus 346 with $\epsilon = 0.98$; (b) Bus 346 with $\epsilon = 0.995$; (c) Bus 1435 with $\epsilon = 0.98$; (d) Bus 1435 with $\epsilon = 0.995$; (e) Bus 4060 with $\epsilon = 0.98$; (f)Bus 4060 with $\epsilon = 0.995$.

Extreme Case Analysis

The buses whose voltage magnitudes and angles have high probability of violating operational limits should be paid more attentions than others, e.g., plus and minus 5% threshold of the voltage magnitude under normal operations. To this end, the buses with largest standard deviation of voltage magnitude and angles are found to be Bus No. 8854 and No. 1265, respectively. Their estimated PDFs by the MC method, the gPC method and the gPC-ANOVA method are shown in Fig. 8.5. It can be observed that all three methods are able to obtain good results, demonstrating their capabilities of dealing with cases that yield possible operational limits violations. As a result, the proposed gPC-ANOVA method is further validated to achieve much higher computational efficiency while maintaining the same level of accuracy as compared to the MC and the gPC methods.

8.1.3 Conclusions

In this chapter, a hierarchical adaptive polynomial chaos-ANOVA method is proposed to solve the PPF and variance analysis problem. This method greatly alleviates the famous "curse of dimensionality" from which the gPC method suffers. We proved the equivalence between the polynomial chaos expansion and the ANOVA decomposition and showed that the variance analysis can be executed with no additional computing time. Furthermore, by using the Stieltjes procedure, we showed that an arbitrary probability distribution can be assumed for the input random variables. The accuracy, efficiency, and adaptivity of the proposed method have been demonstrated on several case studies carried out on the IEEE 118 system and the 1354-bus European system. Future work will focus on further improving the computing time of the proposed method in large-scale systems.

8.2 Case Studies for Propagating Uncertainties in Power System Dynamic simulations

In this section, two case studies are analyzed. The first is conducted on the standard three-machine-nine-bus Western Electricity Coordination Council (WSCC) power system [44] with the classical generator model. This case study compares the gPC method, the MEgPC method, and the traditional MC method in terms of calculation efficiency and accuracy for stochastic power system dynamic studies. It also shows the condition at which the gPC fails whereas the MEgPC works. Furthermore, the characteristics for the decomposition of random space in MEgPC is also discussed.

The second case study is conducted on the New England 10 machine 39-bus system with a detailed two-axis generator model. All system parameters come from [45]. This test demonstrates the performance of MEgPC in propagating uncertainty in power system dynamic

simulations. Both stable and unstable cases are considered.

These two test cases show that the PC-based approach can accurately quantify the uncertainties in power systems, which is essential in power system dynamic security analysis. Furthermore, the calculation efficiency is significantly improved over that of traditional MC simulations.

8.2.1 Simulation Results on the WECC 3-Machine 9-Bus System

One Random Variable Case

Table 8.6: CPU Times of the gPC, MEgPC, and MC Methods

Method	gPC(P=4)	gPC(P=70)	MEgPC	MC
CPU time	1.0s	21.9s	18.7s	2028.5s

This case study is conducted on the WECC power system using the classical generator model [44]. Here it is assumed that the power injection in Generator 3 follows the Gaussian distribution with the mean being the original bus power injection and the standard deviation being 10% [6] to account for the uncertainty induced by a renewable energy unit. A three-phase fault is applied at Bus 8 at 1s. After 5 cycles, it is cleared by opening Line 8-9. The output of interest is the rotor angle δ_{21} of Machine 2 with respect to that of Machine 1. The gPC method is tested using 4th and 70th order PC, the MEgPC method using $3rd$ order PC and $\theta_1 = 10^{-2}$, and the MC method with 20,000 samples [17]. The sample mean and variance of δ_{21} as a function of time obtained using the aforementioned methods are plotted separately in Fig. 4.1(a), (b) and their corresponding calculation times are given in Table. 8.6. Fig. 4.1(a) shows that the four sample mean curves of δ_{21} provided by the the $4th$-order and the $70th$-order gPC methods, the MEgPC method, and the MC method are very close to each other. Both the gPC methods and the MEgPC method provide the same results as that of the MC method. Fig. 4.1(b) shows that the two sample variance curves of δ_{21} provided by the $4th$-order and the $70th$-order gPC methods are very close to that of the MC method for $t < 5$s but significantly deviate after that time. By contrast, the sample variance curves provided by the MEgPC and MC methods remain close to each other at all times. Fig. 4.1(c) displays the sample mean and variance curves provided by the MEgPC method. Here, the sample mean is plotted in black while the standard deviations are plotted as error bars in blue. These curves help grid operators to make decisions aimed at improving the dynamic security of the power system of which they are in charge.

Regarding the computing times, the gPC method is much faster than the MEgPC method for low order (e.g., 4th order) PC but becomes slower for higher orders (e.g., 70th order). However, the gPC and MEgPC methods are much faster than the MC method in all cases.

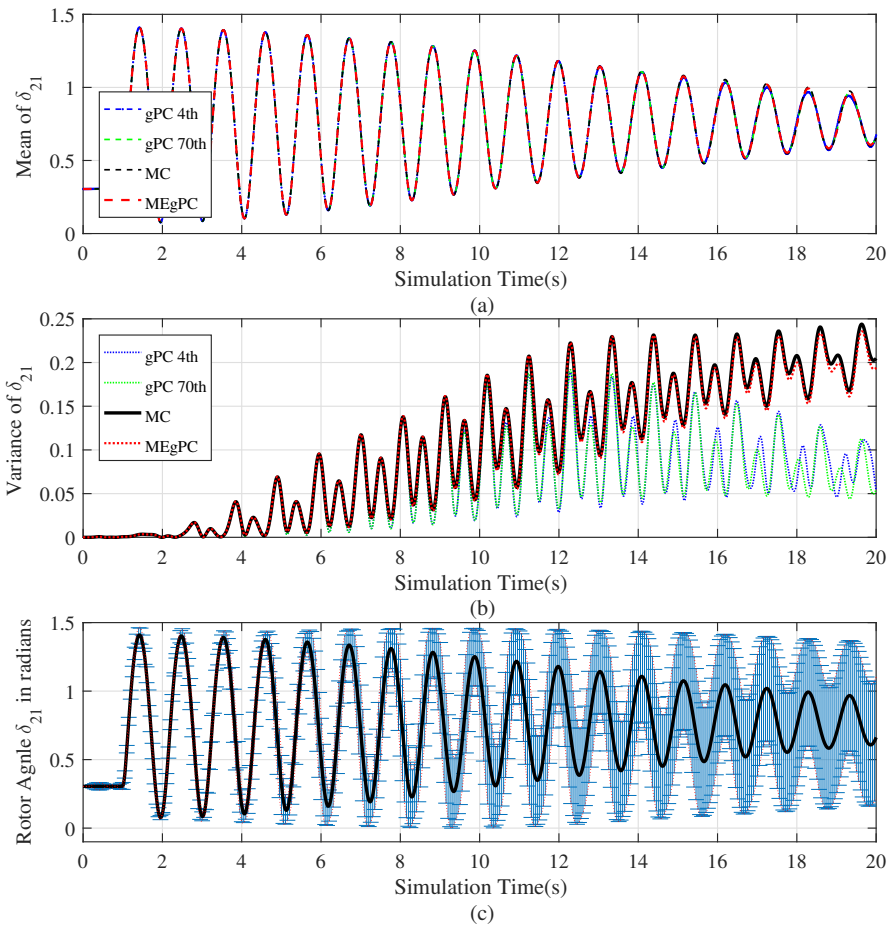


Figure 8.7: Plots for δ_{21}: (a) sample mean of δ_{21} using the gPC, MEgPC, and MC methods; (b) sample variance of δ_{21} for all three methods; and (c) mean and error bar for δ_{21} using MEgPC.

Rules of the MEgPC Method

How the domain of the random input variables in the MEgPC method is decomposed using different thresholds and PC order is demonstrated here. The element number, the calculation time, and the reliability of the MEgPC method are given in Table. 8.7. Table 8.7 leads to the following three conclusions: 1) with smaller threshold values, more elements are generated, resulting in an increase in computational time; 2) with equal threshold values, fewer local elements are generated for the higher order PC, resulting in smaller computational times since the higher order PC has relatively lower local errors; 3) with larger threshold values and fewer local elements, higher PC order does not provide accurate results. Thus, PC with orders larger than 8 are not recommended when using the MEgPC method. Fig. 4.2 shows that most of the splitting of the domain of the random input variable occurs over small intervals with high probabilities. The MEgPC method improves its accuracy by focusing in the intervals with high probability.

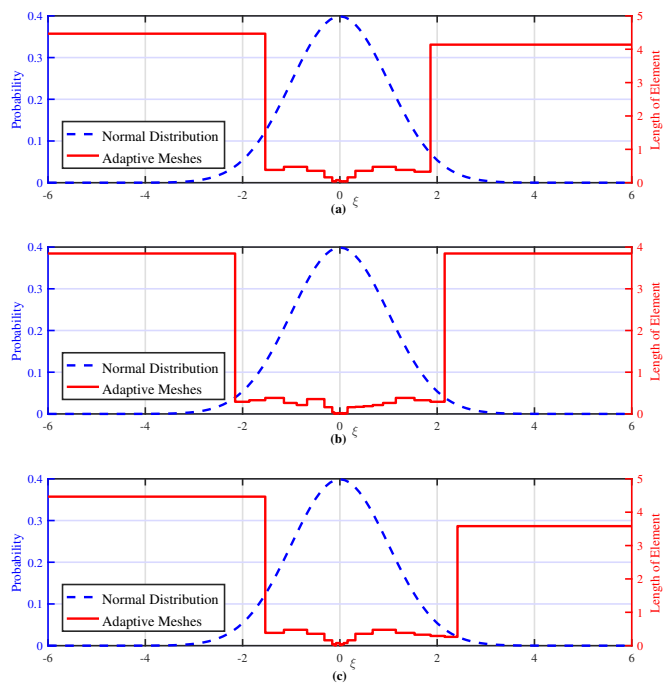


Figure 8.8: Elements and corresponding distribution for δ_{21}: (a) $P = 3$, $\theta = 10^{-2}$, (b) $P = 3$, $\theta = 10^{-3}$, and (c) $P = 4$, $\theta = 10^{-3}$.

Impact of the Initial Conditions

To demonstrate the capability of our proposed approach to handle uncertainties associated with the initial conditions, we assume that the latter associated with the rotor machine

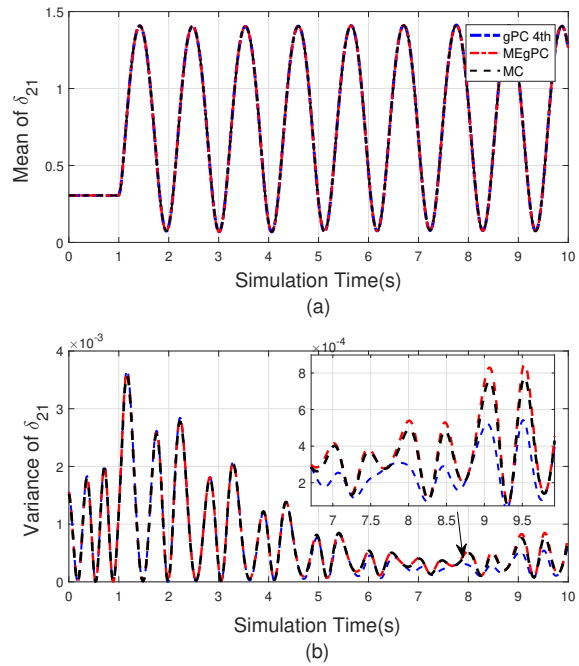


Figure 8.9: Plots for δ_{21} under uncertain initial condition: (a) sample mean of δ_{21} using the gPC, MEgPC, and MC methods; (b) sample variance of δ_{21} for all three methods;

Table 8.7: Space Decomposition for multi-element gPC.

Threshold	**MEgPC**	**p=3**	**p=4**	**p=5**	**p=6**
$\theta_1 = 10^{-2}$	**Elements**	18	6	4	2
	Run Time	18.7s	7.5s	5.8s	2.8s
	Success	Yes	Yes	No	No
$\theta_1 = 10^{-3}$	**Elements**	42	22	7	4
	Run Time	43.4s	28.9s	12.5s	6.3s
	Success	Yes	Yes	Yes	No
$\theta_1 = 10^{-4}$	**Elements**	238	98	33	18
	Run Time	256.6s	127.6s	52.8s	31.8s
	Success	Yes	Yes	Yes	Yes
$\theta_1 = 10^{-5}$	**Elements**	673	375	182	75
	Run Time	720.7s	523.8s	322.9s	134.6s
	Success	Yes	Yes	Yes	Yes

angles δ_1 are subject to uncertainties modelled by a Gaussian distribution with the mean being the original value and the standard deviation being 10%. The simulation results of the sample means and the sample variances of δ_{21} obtained by the gPC method, the MEgPC method and the MC method with 20,000 samples are plotted in Fig. 4.3, respectively. It is observed from Fig. 4. 3(a) that the $4th$-order gPC methods, the MEgPC method, and the MC method achieve a comparable performance. However, for the sample variance curve displayed in Fig. 2(b), the $4th$-order gPC method achieves accurate results during the first 7 seconds but produces increased errors later on. By contrast, our proposed MEgPC is able to maintain very good accuracy at all times.

8.2.2 Case studies on the 10-Machine, 39 Bus System

Propagating Uncertainty in Complex Power Systems

This case study is conducted on the New England system using a two-axis generator model with the IEEE-DC1A exciter and the TGOV1 turbine governor [44] for both stable and unstable conditions. 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 5% of their means. Generator 8 is a solar energy unit, whose real power generation is assumed to follow a Beta distribution over an interval of [0, 400 MW] with shape parameters equal to 2.06 and 2.5 [17]. Two scenarios are considered. In the first scenario, a three-phase fault is applied at Bus 15 at 0.5 s and is cleared after 3 cycles by opening Lines 15-16 and 14-15. The rotor angle of Generator 5 with respect to that of Generator 1, δ_{51}, and its angular frequency, ω_5, are selected to analyze the stability of the power system. The simulation results are depicted in Figs. 4.4 and 4.5. In the second scenario, the same set-up is considered except that the fault is permanent and the protection devices fail to operate, which causes the system to be unstable. The simulation results are depicted in Figs. 4.6 and 4.7.

Figs. 4.4 and 4.5 show that the sample means and variances of the gPC method (P=6) and the MEgPC method match that of the MC method with 20,000 samples up to about 10 s for the mean and about 5 s for the variance. However, after that only the MEgPC method matches the MC results, while the gPC method begins to deviate significantly from time to time. The plots of δ_{51} and ω_5 curves reveal that the disturbance is well damped as the system returns to a stable equilibrium point.

Figs. 4.6 and 4.7 show that the gPC method is accurate for the unstable case during about the first 3 s. After that, it loses its accuracy. The MEgPC method maintains a good accuracy throughout the simulations. From Fig. 4.6, it is observed that the sample means of δ_{51} and ω_5 calculated using the MEgPC method diverge, while their standard deviations increase, indicating that the system loses its stability.

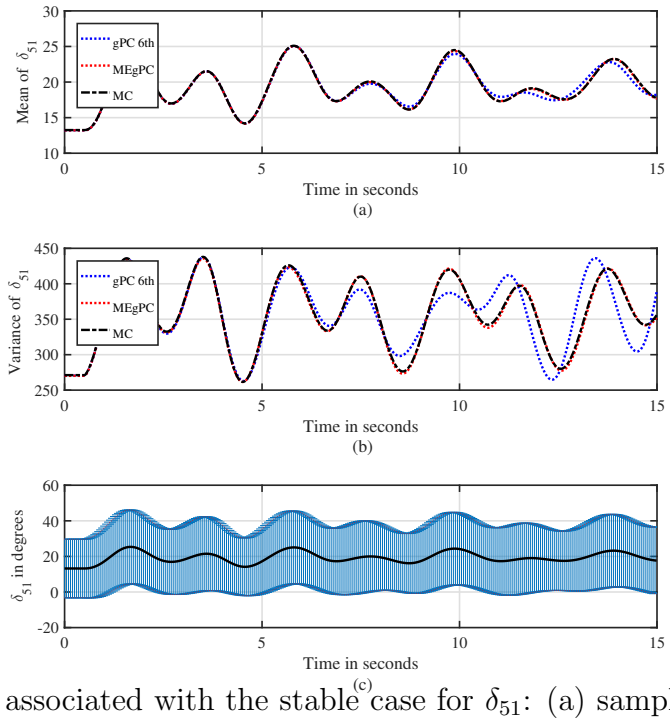


Figure 8.10: Plots associated with the stable case for δ_{51}: (a) sample mean of δ_{51} using the gPC, MEgPC, and MC methods; (b) sample variance of δ_{51}; and (c) mean and error bar for δ_{51} using MEgPC.

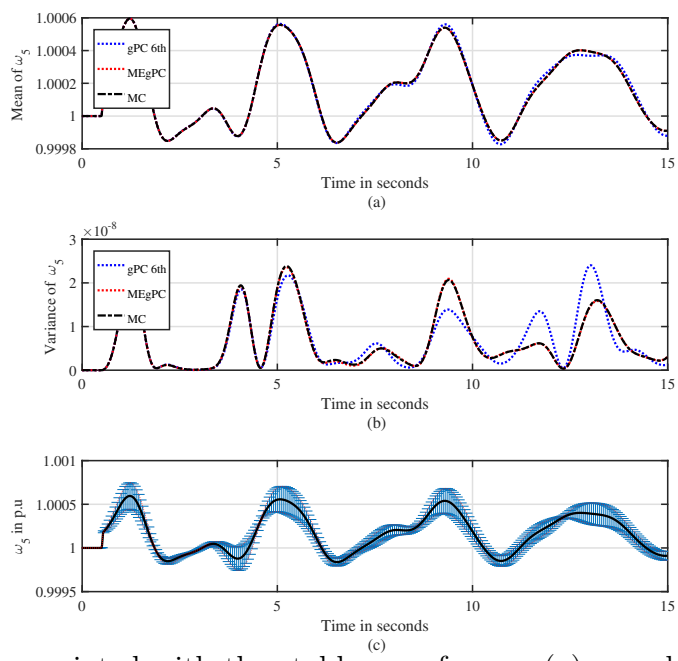


Figure 8.11: Plots associated with the stable case for ω_5: (a) sample mean of ω_5 using the gPC, MEgPC, and MC methods; (b) sample variance of ω_5; and (c) mean and error bar for ω_5 using MEgPC.

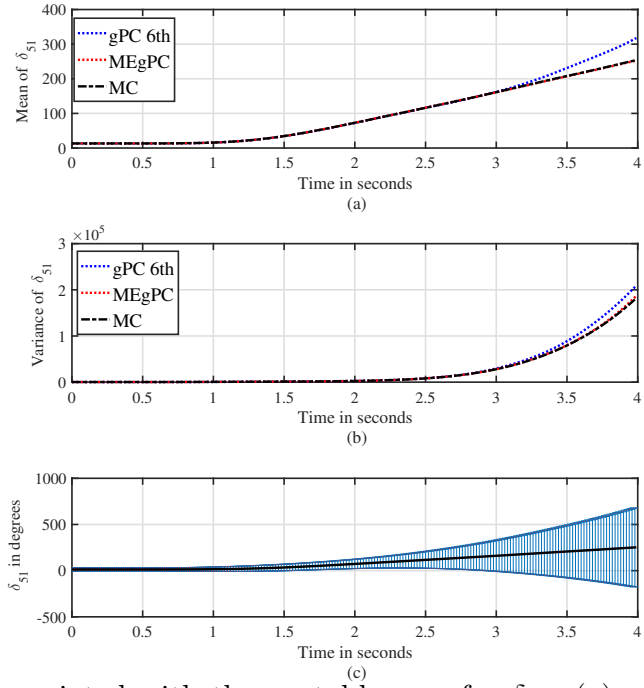


Figure 8.12: Plots associated with the unstable case for δ_{51}: (a) sample mean of δ_{51} using the gPC, MEgPC, and MC methods; (b) sample variance of δ_{51}; and (c) mean and error bar for δ_{51} using MEgPC.

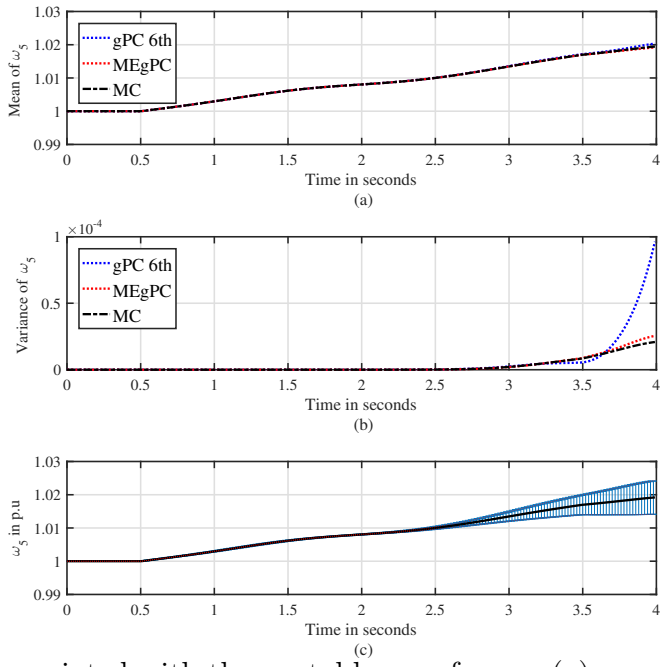


Figure 8.13: Plots associated with the unstable case for ω_5: (a) sample mean of ω_5 using the gPC, MEgPC, and MC methods; (b) sample variance of ω_5; and (c) mean and error bar for ω_5 using MEgPC.

Rules of the MEgPC Method

Table 8.8 displays the threshold and PC order used for four cases of the stable scenario with MEgPC decomposition. For example, consider δ_{51}, the decomposition of the domain for the two random variables is displayed in Fig. 4.8. The interval of every element is represented with a rectangle with a different color. The CPU times for each case as well as for the gPC and MC methods are listed in Table 8.9.

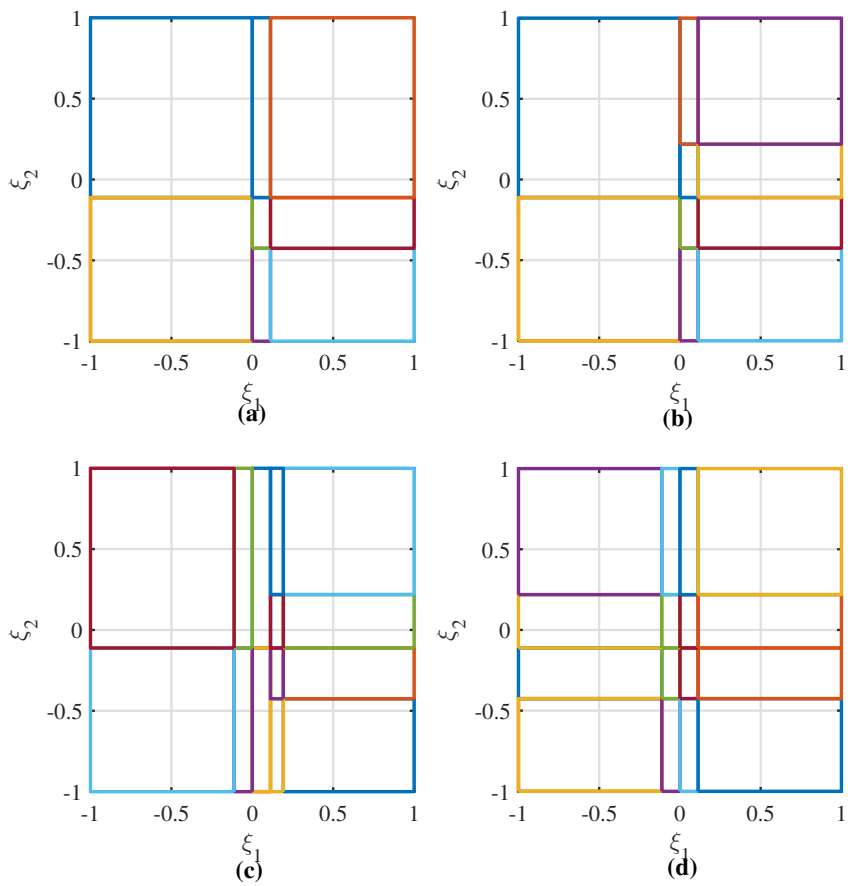


Figure 8.14: Elements for δ_{51} with (a) $P = 3$, $\theta_1 = 10^{-2}$, $\theta_2 = 10^{-2}$; (b) $P = 4$, $\theta_1 = 10^{-2}$, $\theta_2 = 10^{-2}$; (c) $P = 4$, $\theta_1 = 10^{-3}$, $\theta_2 = 10^{-2}$; and (d) $P = 4$, $\theta_1 = 10^{-3}$, $\theta_2 = 10^{-4}$

Figs. 4.8(a), (b), and (c) show that with a small θ_1 and high PC order, the domain will be split to more intervals, which means more elements are generated. Fig. 4.8(c) and (d) show that θ_2 aids in choosing the most sensitive dimensions for decomposing the domain of the input random variables. If θ_2 is too small, all the dimensions are decomposed into an equal number of intervals, which increases the computing time. Table 8.9 shows that the gPC method is still faster than the MEgPC and the MC methods. However, considering both the

Table 8.8: Stable scenario conditions for MEgPC.

Scenario	Case(a)	Case(b)	Case(c)	Case(d)
PC order	3	4	4	4
θ_1	10^{-2}	10^{-2}	10^{-3}	10^{-3}
θ_2	10^{-2}	10^{-2}	10^{-2}	10^{-4}

computational time and the accuracy, the MEgPC method is the most cost-effective one to use.

Table 8.9: CPU times for the gPC(p=6), and MC Methods

Method	gPC	Case(a)	Case(b)	Case(c)	Case(d)	MC
Time(s)	4.6	17.7	28.5	40.8	46.1	6267

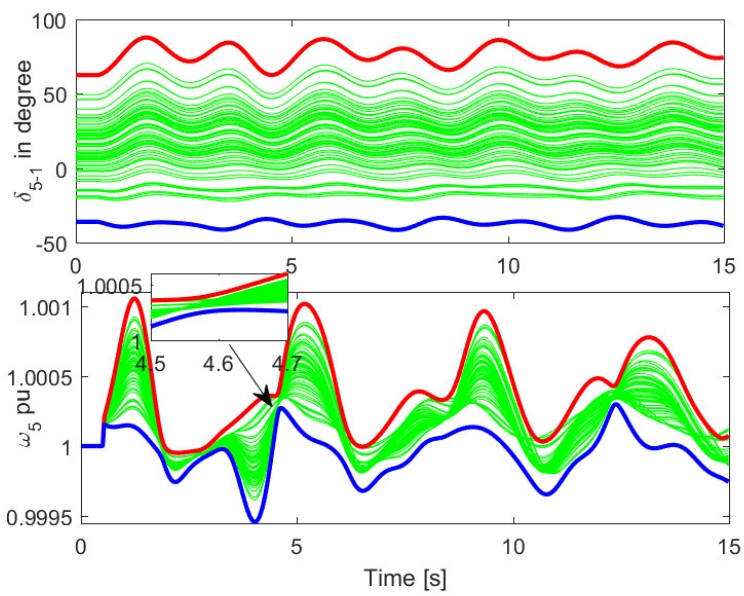


Figure 8.15: Plots for the upper and lower bounds of the state variables δ_{51} and ω_5 in stable case, where the upper (lower) bounds for MEgPC method are in red (blue, respectively); the Monte Carlo samples are all in green.

Upper and Lower Bounds of the Obtained Results

The upper and lower bounds of the obtained results may be useful for power system operations and controls. Following the work of [38, 40], we use the 3σ rule to calculate the corresponding upper and lower bounds of our proposed approaches. Specifically, the upper

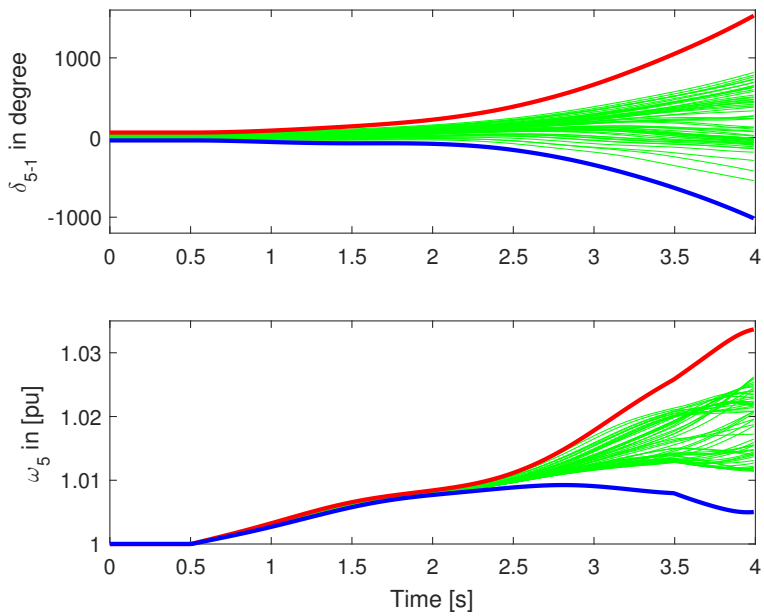


Figure 8.16: Plots for the bounds of the state variables δ_{51} and ω_5 in unstable case, where the upper (lower) bounds for MEgPC method are in red (blue, respectively); the Monte Carlo samples are all in green.

and lower bounds for the state variables δ_{51} and ω_5 are calculated for both stable and unstable cases as displayed in Figs. 4.9 and 4.10, respectively. Note that the upper bound is shown as a red curve, the lower bound is displayed as a blue curve, and the MC results as green curves. It is found that the results obtained by our MEgPC method are within the bounds for both the stable and unstable cases.

8.2.3 Conclusions and Future Work

In this part, both the gPC and the MEgPC methods are applied to quantify the uncertainties in statistical power system dynamic analysis. When compared with the traditional MC method, these two polynomial-chaos-based methods greatly reduce the computational time. When compared with the perturbation method, they perform well in both stable and unstable conditions. Furthermore, use is made of the Stieltjes procedure to generate the orthogonal polynomial chaos functions associated with arbitrary probability distributions of the input random variables. Simulation results reveal that the gPC method is the best method for short-term dynamical simulations from the standpoint of computational time, while the MEgPC method is recommended for long-term dynamical simulations. Future work will focus on the development of an algorithm to increase the efficiency of the MEgPC method for large-scale power systems with renewable energy generation using detailed dynamical models. Furthermore, the way to conduct statistical power system dynamic security

assessment considering uncertainty will be further developed.

8.3 Case Studies in Polynomial-Chaos-Based Kalman Filter

To demonstrate the effectiveness of the proposed PCKF, several case studies are carried out using two standard IEEE benchmark systems, namely the 3-machine, 9-bus system and the 10-machine, 39-bus system. The system data can be found in [36]. The system dynamics are governed by the following two differential equations under the assumption that the classical synchronous generator model is used:

$$\frac{d\delta_i}{dt} = \omega_i - \omega_s, \tag{8.1}$$

$$\frac{2H}{\omega_s}\frac{d\omega_i}{dt} = T_M - \frac{|E|}{X'_d}(|V_j|sin(\delta_i - \theta_j)) - D(\omega_i - \omega_s). \tag{8.2}$$

where the subscript i denotes the ith generator and j denotes the jth system bus; δ denotes the rotor angle in radians; ω and ω_s are respectively the actual and the synchronous rotor speed in radians/s; T_M, H and D are the mechanical power input, the inertia constant and the damping factor, respectively; X'_d is the transient reactance; $|E|\angle\delta$ represents the voltage behind the transient reactance, and $|V_j|\angle\theta_j$ is the bus voltage phasor. It is assumed that $|V_j|$ and $|\angle\theta_j|$ can be directly measured by the phasor measurement units (PMUs). Using the PMU measurements, the observation model can be expressed as

$$P_e = \frac{|E|}{X'_d}(|V_j|sin(\delta_i - \theta_j)),\ Q_e = -\frac{V_j^2}{X'_d} + \frac{|E|}{X'_d}(|V_j|cos(\delta_i - \theta_j)) \tag{8.3}$$

where P_e and Q_e represent the generator real and reactive power, respectively [44]. For more details, the reader is referred to [60, 61].

For the 9-bus system, a three-phase fault is applied at Bus 8 at 1s. After 5 cycles, it is cleared by opening the Line 8-9. While for the IEEE 39-bus system, it is assumed that a large system disturbance occurs at t=0.5s by opening the transmission line between Buses 15 and 16. All the true states and measurements are obtained through time-domain simulations; Gaussian noise with zero means and standard deviations 0.01 is added to the true values to simulate the realistic measurements. The traditional gPC-EnkF with tensor-product rules, sparse-grid method and pre-select CPs samples are used for comparisons. To test the statistical performances of these Kalman filters, 100 Monte Carlo simulations are performed and the probability density functions (PDF) of the calculation times and the root-mean square errors (RMSE) are compared. The simulation results are shown in Fig. 8.17 and Table 8.10. It can be found that for small-scale system, the improved computational efficiency of the proposed approach over other ones are not significant, see Fig.8.17-(a)-(b). The tensor rule based

gPC-EnKF method shows the worst performance. However, when the scale of the system is increased, our proposed PCKF achieves much higher computational efficiency than all the other methods proposed in the literature while maintaining the same level of accuracy as shown in Fig. 8.17-(c)-(d). To further elaborate the results, the averaged RMSEs and the CPU times of the PCKF and the EnKF with 100 samples for the two cases are displayed in TABLE I. It is found that the proposed PCKF improves the computational efficiency significantly while achieving similar estimation accuracy as other methods. Note that 3^{20} CPs are required for the Tensor-based gPC-EnKF method in the 10-machine system, which is almost impossible to calculate. Thus, its results are neglected in Table I. The state estimation results with different types of Kalman filter for these two test systems can be found in Figs. 8.18- 8.19.

Table 8.10: Average CPU Times and RMSEs

Method	PCKF	Pre-select	Sparse-grid	Tensor	EnKF
Group A: 3-Machine, 9-Bus Test System					
Time (s)	0.1195	0.1670	0.2203	0.9371	0.2987
RMSE	1.763e-4	1.742e-4	1.754e-4	1.86e-4	9.481e-4
Group B: 10-Machine, 39-Bus Test System					
Time (s)	0.5364	1.4841	3.8698	—-	0.8428
RMSE	1.714e-4	1.710e-4	1.708e-4	—-	5.158e-4

8.3.1 Conclusion

This paper presents a new PCKF that is able to deal with nonlinear dynamical systems with strong nonlinearity. Based on the gPC theory, the uncertainties from the model and the measurements are propagated through the nonlinear functions via PC bases. To cope with the curse of dimensionality, we propose a dimension reduction strategy used together with the collocation points re-generation. This allows us to significantly improve the computational efficiency without any loss of estimation accuracy. Simulations carried out on various IEEE benchmark test systems show that our proposed method outperforms the other gPC-Kalman filters proposed in the literature. Although the proposed PCKF is able to deal with strong system nonlinearity with large number of state variable, it is not robust to outliers, a vulnerability shared by most Kalman filters. As a future work, we will resort to robust statistics to develop a robust PCKF.

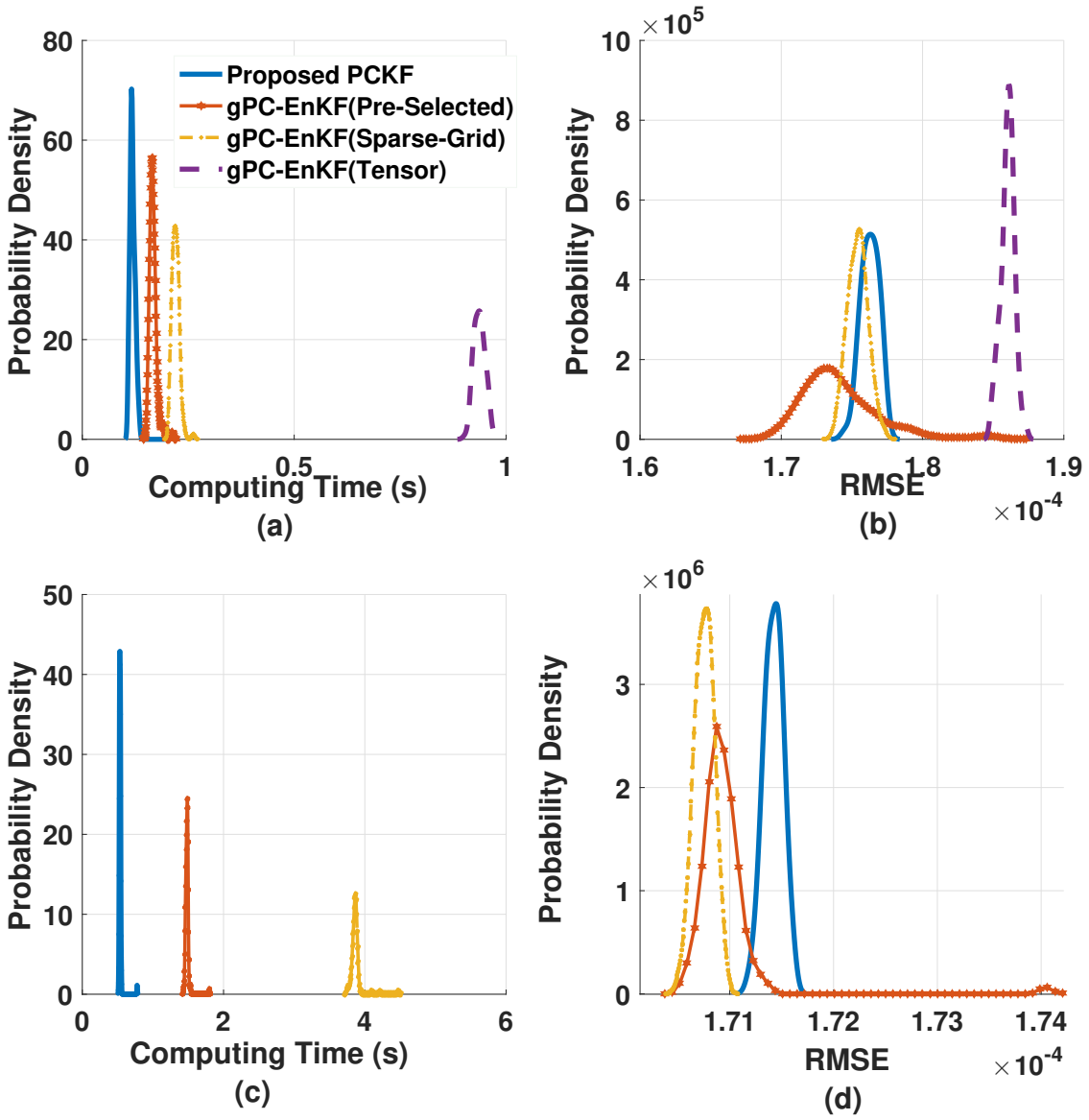


Figure 8.17: Results of four different Kalman filters: (a)-(b) PDFs of the computing times and RMSEs for the 3-machine system (c)-(d) PDFs of computing times and RMSEs for the 10-machine system

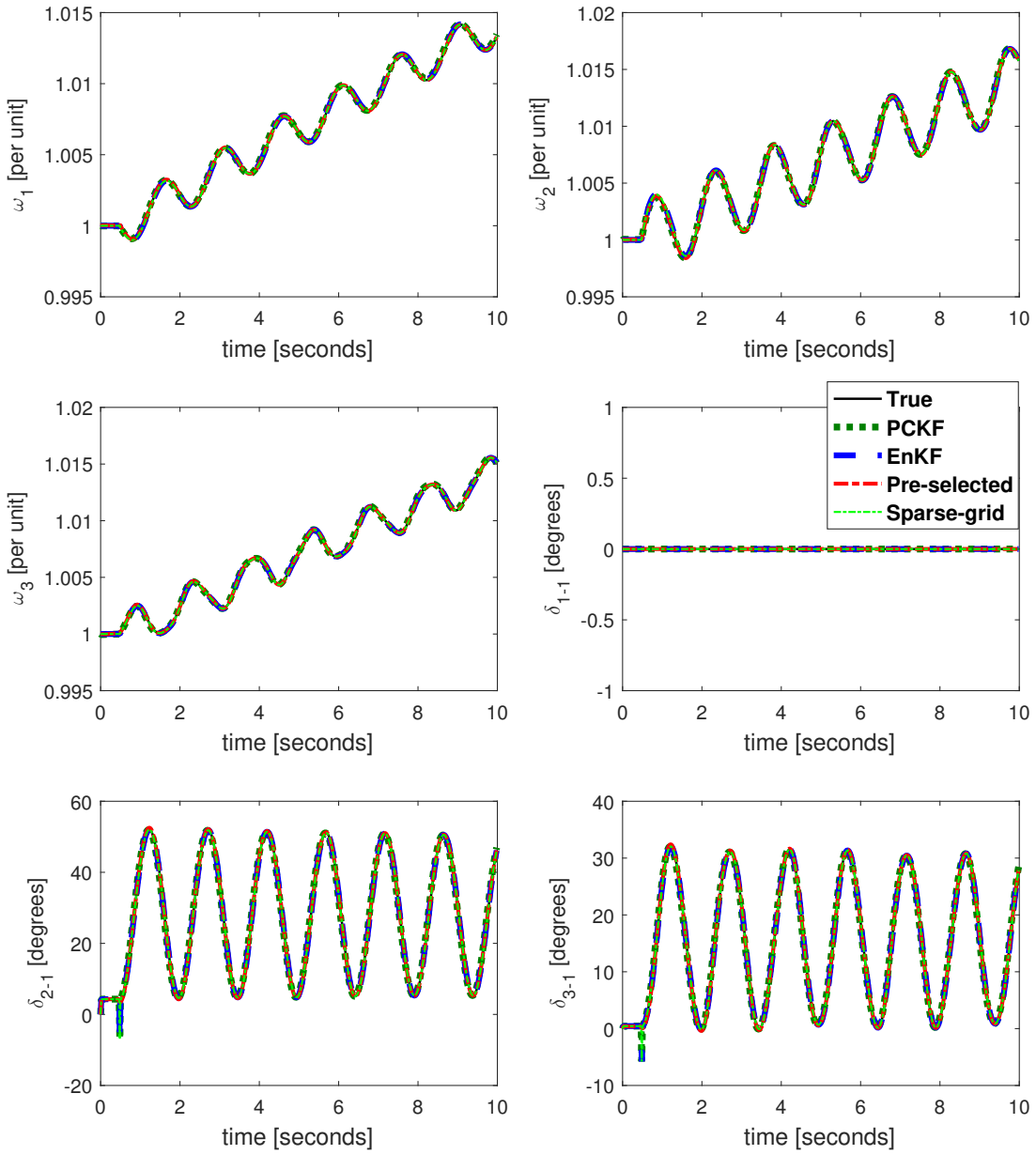


Figure 8.18: States estimation results of the 3-Machine, 9-Bus test system with different types of Kalman filter

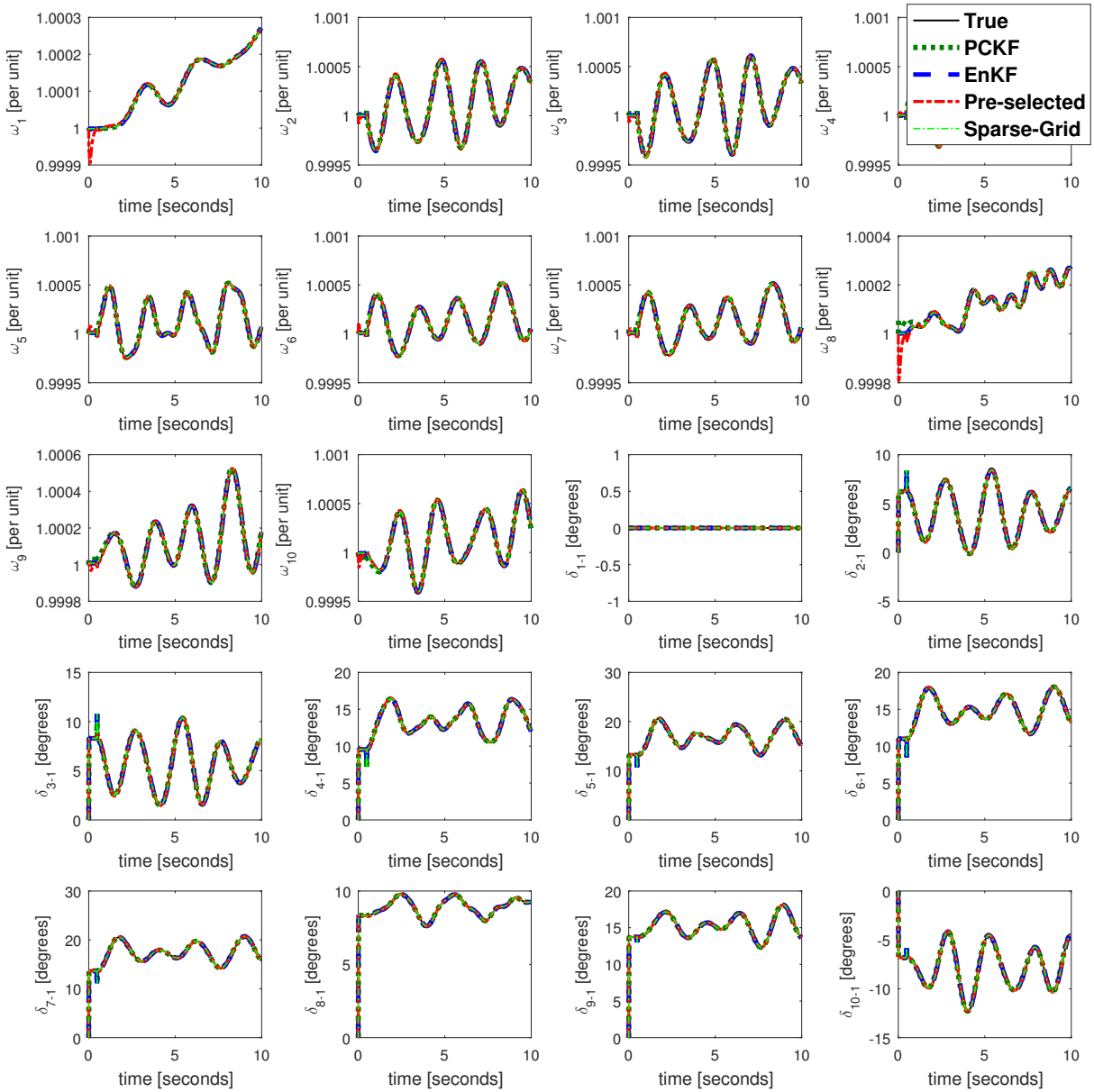


Figure 8.19: States estimation results of the 10-Machine, 39-Bus test system with different types of Kalman filter

8.4 Case Studies for the Response-Surface based Bayesian Inference in Power System Dynamic Simulations using PMU measurement

In this section, we test our method on the New England test system [36] using MATLAB® R2016b. Case studies are conducted on the New England system using the aforementioned two-axis generator model with the IEEE-DC1A exciter and the TGOV1 turbine governor as described in [44]. Parameter estimation studies are conducted to evaluate the accuracy, the calculation efficiency, the impact of the Bayesian prior knowledge, the high dimension performance, and the applicability of the methods in different realistic scenarios.

8.4.1 Studies on the Prior Information

The posterior distribution of the Bayesian inference is determined by the likelihood function and the prior PDF. For the practice implementation, the prior PDF of the parameters can be chosen using the data provided by the manufacturer, which is considered to be reasonable in general. However, the errors are still inevitable. This study analyzes the estimation accuracy under different prior information. Suppose the parameters to be estimated are the moment of inertia H and the amplifier gain K_A in the exciter. The errors from the original data provided by the manufacturer for every generators are assumed to have about 5% deviations from of their true values in Group A and 10% for the Group B. These two parameters are assumed to follow the Gaussian distributions with their means equal to the manufacturer's data and the standard deviations being set to 10% of the means to account for the parameter uncertainties. The random vector q is assumed to follow a multivariate Gaussian probability distribution with zero means and standard deviations equal to 10% of the original values; $\boldsymbol{e}$ are assumed to be identical, independent Gaussian noise with 0.01 standard deviation. The transmission line between Bus 19 and Bus 33 is removed after 0.5 seconds. The time interval is selected as 3 seconds. The parameter estimation is conducted separately for every generator in a decentralized manner. The sample size for the M-H sampler is set to $100,000$. We choose the Hermite polynomials as surrogate simulators for the Bayesian Inference. Our proposed method completes the calculation for every single generator in just 6s while the traditional method takes around 29 minutes, thus achieving two orders of magnitude in speed. The simulation results from the proposed response-surface-based method are given in Table 8.11. It shows that the MAP estimates for both groups match the true values quite well. We also observe that the prior PDF of the parameters has a negligible influence on the accuracy of the MAP estimates. This is because, for every generator, the sum of the log likelihoods given by (7.15) that are associated with just a few uncertain parameters and are provided by 360 observations in total are negligible. Evidently, the fast sampling rate of the PMUs has greatly decreased the estimation error brought by an inappropriate prior knowledge. Given a good prior information, the estimation algorithm converges faster.

Table 8.11: Comparison of the MAPs with the Actual Values for the ith Machines under Different Prior Information

Group A: Good Prior Information

i	H^{prior}	H^{true}	H^{MAP}	K_A^{prior}	K_A^{true}	K_A^{MAP}
1	40	42	42.24	38	40	40.01
2	29	30.3	30.61	38	40	40.00
3	34	35.8	36.05	38	40	39.97
4	27.5	28.6	28.90	38	40	40.03
5	24.5	26	26.09	38	40	39.98
6	36	34.8	35.07	42	40	40.02
7	28	26.4	26.71	42	40	40.00
8	25.5	24.3	24.51	42	40	39.98
9	36	34.3	34.68	42	40	40.01
10	525	500	502.1	42	40	40.00

Group B: Bad Prior Information

i	H^{prior}	H^{true}	H^{MAP}	K_A^{prior}	K_A^{true}	K_A^{MAP}
1	38	42	42.26	36	40	40.00
2	27.5	30.3	30.38	36	40	40.03
3	32	35.8	35.94	36	40	40.01
4	26.5	28.6	28.96	36	40	40.02
5	23.5	26	26.07	36	40	40.04
6	37.5	34.8	35.07	44	40	40.01
7	29	26.4	26.72	44	40	40.00
8	27	24.3	24.44	44	40	40.00
9	38	34.3	34.73	44	40	39.98
10	550	500	502.5	44	40	39.80

8.4.2 Validation under Different Events

Several papers in the literature discuss the importance of validating the proposed methods under different events [67], [68], [78] since the uncertain parameters can lead to very different model output for different cases. To verify this result, we create seven different events by removing seven different transmission lines. All the other settings are the same with Group A described in the previous case study. The MAP estimates and their corresponding error rates in percentages obtained by our proposed response surface method are displayed in Table 8.12. From the simulation results, we observe that our proposed method yields good accuracy in calculating the MAP parameter estimates for every generator in most of the cases under consideration. However, for Event 7, considering the estimation results for the moment of the inertia H, there are two cases with an error rate higher than 2%. This is because H mainly influences the active power output P while K_A mainly influences the reactive power Q. For Event 7, the active power output has a much smaller variation compared with the variation in the reactive power output. Under this circumstance, with different H values proposed in the M-H sampler, the trajectories of P do not change much, and the corresponding likelihood functions vary little as a result. This leads to potential inaccuracy in the MAP estimates. Luckily, this inaccuracy can be self-detected from the posterior distribution of the Bayesian inference. Unlike the WLS-based estimator that only provides point estimates for the generator parameters, the Bayesian method provides their probability distributions. For example, let us plot the PDFs of the H_2 for Event 3 and Event 7 and the PDF for the H_7 for Event 7, respectively, in Fig. 8.21. Comparing Fig. 8.21.(a) and 8.21.(b), we find that, for H_2, the confidence interval provided by the PDFs under different events can be very different. For Event 3, the PDF mainly covers the range of $30.2 - 30.6$, which is much smaller than that provided by Event 7. This means Event 3 provides a more accurate parameter estimation than that of Event 7, even though the MAP estimators under these two events are very similar. Comparing Fig. 8.21.(b) and 8.21.(c), we find that, under the same Event 7, the PDFs for both H_2 and H_7 estimate larger confidence intervals. Even though the MAP estimator of H_2 is accurate, it can not always be guaranteed. This is demonstrated in 8.21.(c) that the MAP estimate of H_7 is not accurate. This means Event 7 is not an appropriate event to estimate the parameter H. When obtaining estimation results with large confidence intervals, we should also test other events to provide more reliable results. These scenarios have also been discussed in [67]. The Bayesian posterior distribution gives us a good guidance in distinguishing these inappropriate events.

Here, let us further validate our proposed method using Group B data. All the other settings are the same with previous tests. The simulation results are shown in Table. 8.13. It is found that for parameter K_A, the estimation results are always accurate. The error rates are similar to the results shown in Table. 8.12. For parameter H, the error rate is relatively larger, but still provide a relatively accurate results for most of the cases. Again, we found that event 7 cannot provide an accurate estimation. This matches our conclusions before.

Table 8.12: Comparison of the MAPs for Different Events for the Generators

Event	1	2	3	4	5	6	7
Lines	15-16	19-33	29-38	39-9	10-13	26-28	3-18
H_1^{MAP}	41.70	42.24	42.41	42.65	42.71	42.08	41.43
error(%)	0.714	0.571	0.976	1.548	1.690	0.190	1.357
H_2^{MAP}	30.59	30.61	30.44	30.44	30.51	30.27	30.45
error(%)	0.957	1.023	0.462	0.462	0.693	0.099	0.495
H_3^{MAP}	36.05	36.05	35.95	36.21	35.99	35.63	35.28
error(%)	0.698	0.698	0.391	1.145	0.531	0.475	1.452
H_4^{MAP}	29.06	28.90	28.90	28.65	28.66	28.82	28.48
error(%)	1.54	1.05	1.05	0.175	0.210	0.769	0.420
H_5^{MAP}	25.9	26.29	26.14	26.17	25.96	26.43	25.90
error(%)	0.385	1	0.538	0.654	0.154	1.654	0.385
H_6^{MAP}	34.90	35.07	35.06	35.13	34.88	34.90	34.65
error(%)	0.287	0.776	0.747	0.948	0.230	0.287	0.431
H_7^{MAP}	26.60	26.72	26.67	26.46	26.31	26.34	27.15
error(%)	0.758	1.21	1.02	0.227	0.341	0.227	2.84
H_8^{MAP}	24.50	24.51	24.54	24.50	24.57	24.30	24.32
error(%)	0.823	0.857	0.988	0.823	1.11	0.00	0.082
H_9^{MAP}	34.51	34.68	34.69	34.53	34.17	34.70	34.63
error(%)	0.612	1.11	1.14	0.670	0.379	1.117	0.962
H_{10}^{MAP}	503.90	502.1	502.02	503.50	503.21	499.89	486.77
error(%)	0.78	0.42	0.404	0.7	0.642	0.022	2.646
$K_{A_1}^{\mathrm{MAP}}$	40.01	40.01	39.98	39.97	40.01	40.00	40.02
error(%)	0.025	0.025	0.050	0.075	0.025	0.000	0.050
$K_{A_2}^{\mathrm{MAP}}$	40.02	40.00	40.01	40.02	39.97	39.99	39.99
error(%)	0.050	0.000	0.025	0.050	0.075	0.025	0.025
$K_{A_3}^{\mathrm{MAP}}$	39.98	39.97	39.99	39.97	40.02	39.99	39.99
error(%)	0.050	0.075	0.025	0.075	0.050	0.025	0.025
$K_{A_4}^{\mathrm{MAP}}$	40.00	40.03	39.99	39.98	40.03	39.98	40.00
error(%)	0.000	0.075	0.025	0.050	0.075	0.050	0.000
$K_{A_5}^{\mathrm{MAP}}$	39.96	39.98	40.01	40.05	40.02	40.02	40.01
error(%)	0.100	0.050	0.025	0.125	0.050	0.050	0.025
$K_{A_6}^{\mathrm{MAP}}$	40.02	40.02	40.04	39.99	39.98	39.99	40.00
error(%)	0.050	0.050	0.1	0.025	0.050	0.025	0.000
$K_{A_7}^{\mathrm{MAP}}$	39.98	40.00	39.99	40.01	40.00	39.99	40.01
error(%)	0.050	0.000	0.025	0.025	0.000	0.025	0.025
$K_{A_8}^{\mathrm{MAP}}$	40.02	39.98	39.97	39.99	40.00	40.02	39.98
error(%)	0.050	0.050	0.075	0.025	0.000	0.050	0.050
$K_{A_9}^{\mathrm{MAP}}$	40.03	40.01	40.03	40.02	40.03	40.01	40.01
error(%)	0.075	0.025	0.075	0.050	0.075	0.025	0.025
$K_{A_{10}}^{\mathrm{MAP}}$	39.99	40.00	40.02	40.01	40.00	40.00	39.98
error(%)	0.025	0.000	0.050	0.025	0.000	0.000	0.050

Discussion: It is worth pointing out that the error rate for parameter H increases due to the fact that the deviation of prior knowledge is larger than that in Group A. The response surface constructed over the prior distributions of input parameters usually provide an accurate estimation when the samples are closer to the mean values of the prior distributions and gradually loses some accuracy when the samples are close to the tail regions of the prior distributions. Under the Bayesian inference framework that typically assumes the knowledge of a reasonable prior, the estimation accuracy can generally be guaranteed. Furthermore, there are existing techniques to ensure that the initial guesses for bad parameters will not deviate too far from their true values as shown in [67].

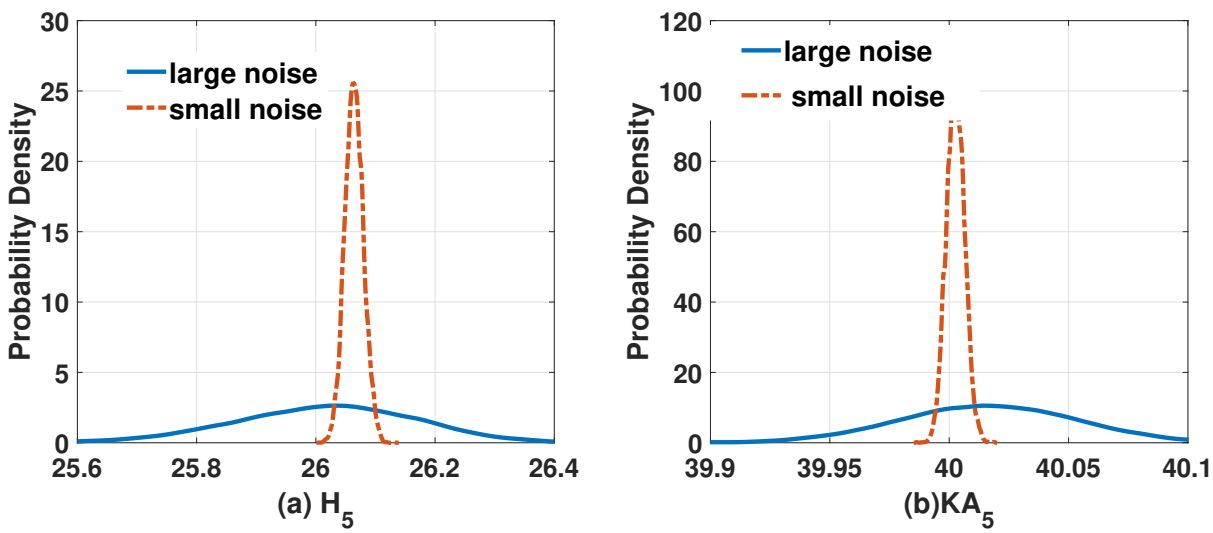


Figure 8.20: Plots for the PDFs of the parameters under different noise level: (a) for H_5 and (b) for K_{A_5}.

8.4.3 Impacts from Different Noise

In this section, we are discussing the impacts of the estimation results under different noise level. Let us take generator 5 under event 1 as an example. The manufacturer's data for H_5 and KA_5 is $\{30, 36\}$ while the true values are $\{26, 40\}$. Here, let us conduct a comparison study that assume $\boldsymbol{e}$ to be identical, independent Gaussian noise with 0.01 standard deviation and 0.001 standard deviation separately. The estimation results are shown in Fig. 8.20. It is shown that under different noise level, the estimation results are all accurate. However, when the noise increases, the confidence interval increases as well. This leads to an estimation result with relatively lower confidence.

Table 8.13: Validation under Different Events for all the Generators using Group-B Data

Event	1	2	3	4	5	6	7
Lines	15-16	19-33	29-38	39-9	10-13	26-28	3-18
H_1^{MAP}	41.70	42.41	42.2	42.0	41.75	42.3	42.5
error(%)	0.714	0.976	0.476	0.000	0.595	0.714	1.190
H_2^{MAP}	30.39	30.41	30.36	30.24	30.50	30.33	30.7
error(%)	0.297	0.363	0.198	0.198	0.660	0.099	1.320
H_3^{MAP}	35.75	36.2	35.93	36.01	35.94	35.82	34.7
error(%)	0.140	1.117	0.363	0.587	0.391	0.056	3.073
H_4^{MAP}	28.60	28.70	28.70	28.68	28.2	28.80	27.9
error(%)	0.000	0.350	0.350	0.280	1.399	0.699	2.448
H_5^{MAP}	26.07	26.08	26.11	25.91	26.05	26.08	25.60
error(%)	0.269	0.308	0.423	0.346	0.192	0.308	1.538
H_6^{MAP}	34.79	34.98	35.04	34.86	34.96	34.81	35.5
error(%)	0.029	0.517	0.690	0.172	0.460	0.029	2.011
H_7^{MAP}	26.80	26.44	26.71	26.90	26.82	26.2	27.22
error(%)	1.515	0.152	1.174	1.894	1.591	0.758	3.106
H_8^{MAP}	24.60	24.45	24.54	24.58	24.81	24.74	24.50
error(%)	1.235	0.617	0.988	1.152	2.099	1.811	0.823
H_9^{MAP}	34.35	34.7	34.68	34.0	34.83	34.71	34.90
error(%)	0.146	1.166	1.108	0.875	1.545	1.195	1.749
H_{10}^{MAP}	502.50	501.5	502.8	505.50	501.5	504.02	496.5
error(%)	0.500	0.300	0.560	1.100	0.300	0.804	0.700
$K_{A_1}^{\mathrm{MAP}}$	39.98	39.97	39.99	39.97	40.02	39.99	39.99
error(%)	0.05	0.075	0.025	0.075	0.05	0.025	0.025
$K_{A_2}^{\mathrm{MAP}}$	40.02	40.01	40.01	40.03	39.98	39.99	40.01
error(%)	0.050	0.025	0.025	0.075	0.05	0.025	0.025
$K_{A_3}^{\mathrm{MAP}}$	40.02	39.97	39.99	39.97	40.02	39.99	40.03
error(%)	0.050	0.075	0.025	0.075	0.050	0.025	0.075
$K_{A_4}^{\mathrm{MAP}}$	40.00	40.03	39.99	39.98	40.03	39.98	40.00
error(%)	0.000	0.075	0.025	0.050	0.075	0.050	0.000
$K_{A_5}^{\mathrm{MAP}}$	39.97	39.99	39.99	40.02	40.00	39.98	40.01
error(%)	0.075	0.025	0.025	0.05	0.00	0.050	0.025
$K_{A_6}^{\mathrm{MAP}}$	40.02	40.01	40.03	40.02	40.03	40.01	40.01
error(%)	0.050	0.025	0.0.075	0.050	0.075	0.025	0.025
$K_{A_7}^{\mathrm{MAP}}$	39.99	40.02	39.99	40.01	40.01	39.99	40.01
error(%)	0.025	0.050	0.025	0.025	0.025	0.025	0.025
$K_{A_8}^{\mathrm{MAP}}$	40.00	39.98	40.00	40.02	39.99	40.02	39.97
error(%)	0.00	0.050	0.00	0.050	0.025	0.050	0.075
$K_{A_9}^{\mathrm{MAP}}$	39.97	40.02	39.99	39.38	40.03	40.01	40.02
error(%)	0.075	0.050	0.025	0.050	0.075	0.025	0.05
$K_{A_{10}}^{\mathrm{MAP}}$	40.01	40.02	39.97	39.99	40.02	40.01	39.97
error(%)	0.025	0.050	0.075	0.025	0.050	0.025	0.075

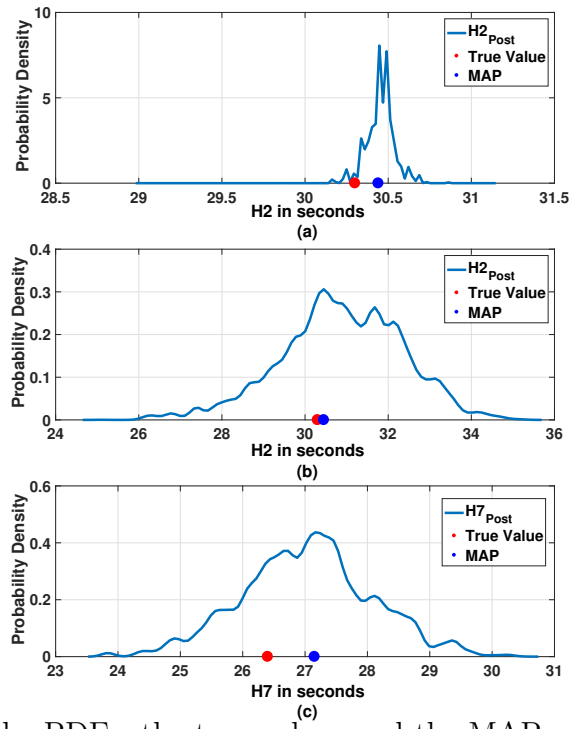


Figure 8.21: Plots for the PDFs, the true values and the MAPs: (a) for $H1$ under Event 3; (b) for $H1$ under event 7; (c) for $H7$ under Event 7.

8.4.4 Tolerance Test

In this section, we are going to test how much deviation on parameters that the proposed method can tolerate. Let us select aforementioned Generator 5 under Event 2 as an example. Their true values for H and K_A are $\{26, 40\}$ The prior values are assumed to have different levels of deviation from the true values, namely 5%, 10%, 15%, 20%, 30% and 40% separately. We are also going to explore the relationship between different standard deviations assumed for the prior pdfs and the estimation accuracy. Let us conduct the tests that assume the prior pdf to be Gaussian distribution with standard deviation as 10%, 30% and 60% of their mean values separately. Therefore, response surfaces with different standard deviations are built for them separately. Let us call the standard deviation used for constructing the response surface as $STDRS$. All the other settings remains unchanged with above experiments in Section V-A. The simulation results are shown in Table 8.14. All the simulations are finished in 5 s. From Table 8.14, we can obtain the following conclusion:

(1) When the true values are at the tail regions of the assumed prior pdfs, the proposed approach will gradually lose its accuracy.

(2) Increasing $STDRS$ can increase estimation accuracy when the prior values have extreme large deviations, but will lose some accuracy when the prior values have very small deviations.

Table 8.14: Tolerance Test under Different Deviations

STDRS(10%)						
Deviation	H^{prior}	K_A^{prior}	H^{MAP}	**error(%)**	K_A^{MAP}	**error(%)**
40%	15.6	64	18.99	26.9	39.17	2.1
30%	18.2	52	22.48	13.5	40.12	0.3
20%	20.8	48	25.11	3.42	40.06	0.15
15%	22.1	46	25.57	1.65	39.99	0.025
10%	23.4	44	26.06	0.23	40.01	0.025
5%	24.7	42	26.08	0.031	40.01	0.025
STDRS(30%)						
Deviation	H^{prior}	K_A^{prior}	H^{MAP}	**error(%)**	K_A^{MAP}	**error(%)**
40%	15.6	64	21.12	18.7	39.3	1.75
30%	18.2	52	24.41	6.11	40.01	0.025
20%	20.8	48	26.28	1.07	40.12	0.3
15%	22.1	46	26.44	1.69	40.12	0.3
10%	23.4	44	26.25	0.96	40.08	0.2
5%	24.7	42	26.11	0.42	40.05	0.125
STDRS(60%)						
Deviation	H^{prior}	K_A^{prior}	H^{MAP}	**error(%)**	K_A^{MAP}	**error(%)**
40%	15.6	64	25.40	2.31	41.15	2.88
30%	18.2	52	26.80	3.07	40.75	1.88
20%	20.8	48	26.28	1.10	40.4	1
15%	22.1	46	25.89	0.42	40.25	0.625
10%	23.4	44	25.74	1	40.1	0.25
5%	24.7	42	25.76	0.923	39.95	0.125

8.4.5 Studies in Higher Dimension

In this part, we are demonstrating the performance of the proposed method with six key parameters that are known to influence the dynamic response of the system and can not be directly measured easily. These are the moment of inertia H, three gains in the exciter, namely, K_A, K_E, K_F, the damping ratio D, and the droop R_D. Let us take Generator 5 as an example. These six parameters are assumed to follow Gaussian distributions with means being the original data provided by the manufacturer, that is, $\{26, 40, 82.5, 1, 1, 0.067\}$ and standard deviations being 10% of the means to account for the parameter uncertainties. The disturbances are caused by Event 2. The simulation results are shown in Fig. 8.22. It is found that the proposed method yields good accuracy of the MAP estimates, which are equal

to $\{26.07, 39.9, 82.7, 1.014, 0.999, 0.0634\}$. Even though there are several peaks regarded as local optima, the M-H sampler still provides the global optima, namely the MAP estimates. However, the calculation efficiency is decreased. To guarantee the accuracy, the sample number is increased to 1,000,000. With our proposed response-surface-based method, the calculation can still be completed within 2 minutes. By contrast, traditional method will take about 5 hours. This further demonstrates the high efficiency of our proposed method.

8.4.6 Validation under Fault

In this case, we demonstrate the proposed method's performance under the disturbances caused by a fault. We also demonstrate the effects of using different probability distributions for the prior pdfs. A three-phase fault that lasts 5 cycles is applied at Bus 19 at 0.5 s. Let us take Generator 5 as an example. These six parameters are assumed to follow uniform distributions with means being the original data provided by the manufacturer, and standard deviations being 20% of the means to account for the parameter uncertainties. The prior values, the estimated values and the true values are listed in Table 8.15. All the other setting remaining the same as in Section V-E. Following the tradition, let us remove the first 30% samples in the burn-in period [83], the 1-D and 2-D posterior marginals of parameters for K_A, K_E, H, D, R_D and K_F obtained by the proposed method are provided by Fig. 8.23. The comparison study results with the traditional Bayesian inference method that uses full dynamic solver as shown in Algorithm 1 are also provided in Table 8.15. From the 2-D posterior marginals of parameters, we can found there are strong correlation between parameters, such as $\{K_A, K_E\}$ and $\{D, R_D\}$. In order to generate enough correlated samples in the proposal functions, we must increase the sample size. Besides the dimension issue, this is one main reason we need as large as 10^6 samples to obtain the unbiased results. The simulation results again show that our proposed method can provide accurate estimation results under a three-phase fault, but with a much smaller computing time compared with the results obtained by Algorithm 3. Furthermore, the experiment shows that the proposed method can work well under the assumption for non-Gaussian prior pdfs.

8.4.7 Effective Dimension Analysis

Not all the parameters can be analyzed by the proposed method due to the observability limitation [80]. Some parameters have negligible influences on the output that makes the accurate prediction impossible. To address this issue, we propose to use the ANOVA index to screen out the key model parameters. The ANOVA expansion represents a function f with N random variable in the form of as

$$\begin{aligned} f(\xi_1, \ldots, \xi_N) = f_0 + \sum_{1 \le j_1 \le N} f_{j_1}(\xi_{j_1}) \\ + \sum_{1 \le j_1 < j_2 \le N} f_{j_1, j_2}(\xi_{j_1}, \xi_{j_2}) + \ldots + f_{1,2,\ldots,N}. \end{aligned} \tag{8.4}$$

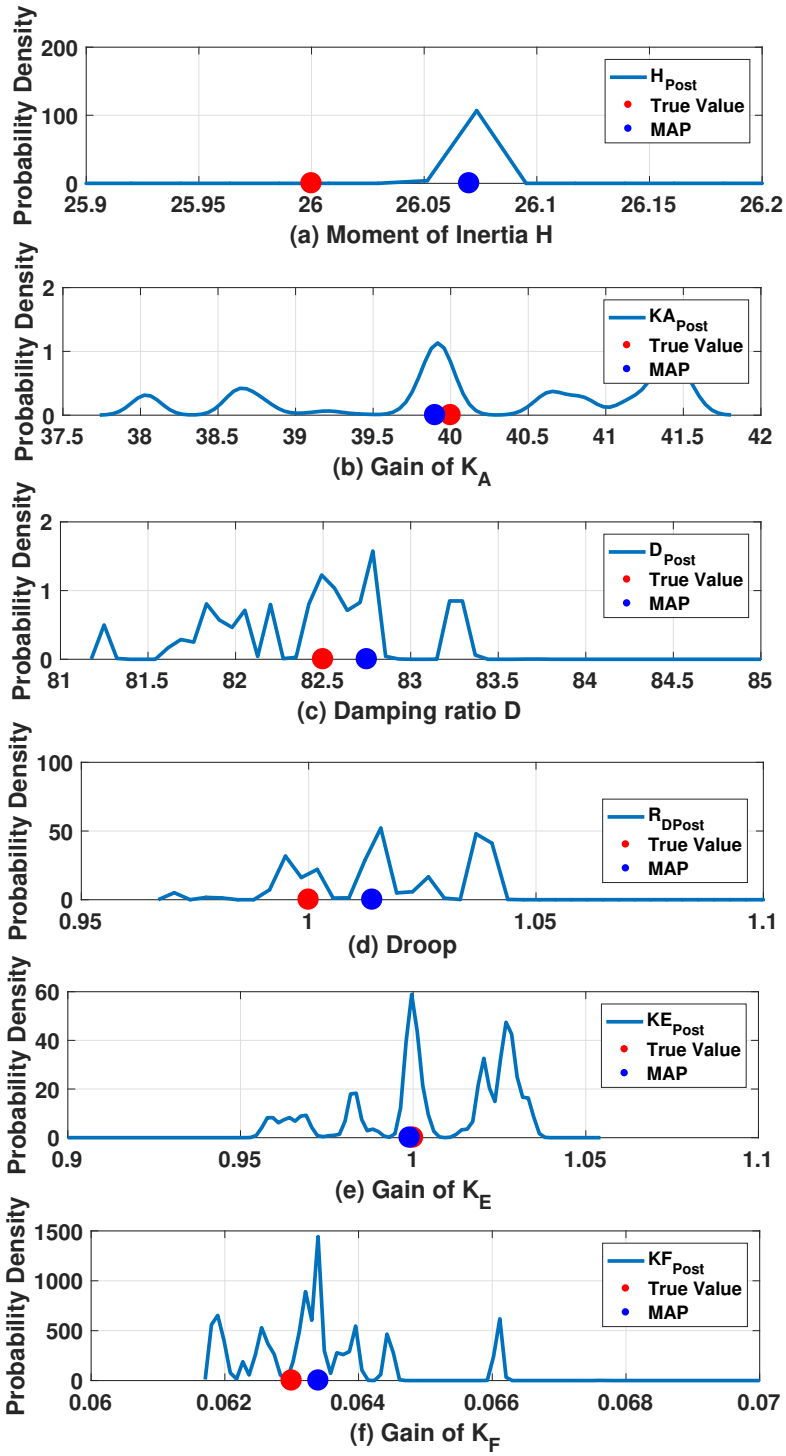


Figure 8.22: Plots for the pdfs, the true values and the MAPs: (a) for moment of inertia H; (b) for Gain KA; (c) for Damping ratio D;(d) for Droop R_D; (e) for Gain KE; (f) for Gain K_F.

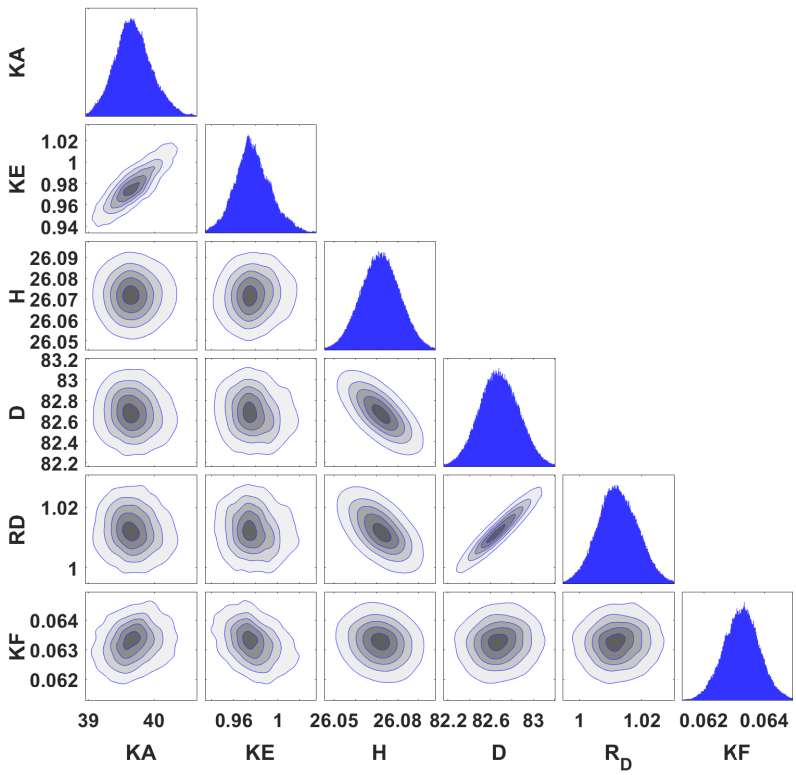


Figure 8.23: 1-D and 2-D posterior marginals of parameters for K_A, K_E, H, D, R_D and K_F obtained by the proposed method

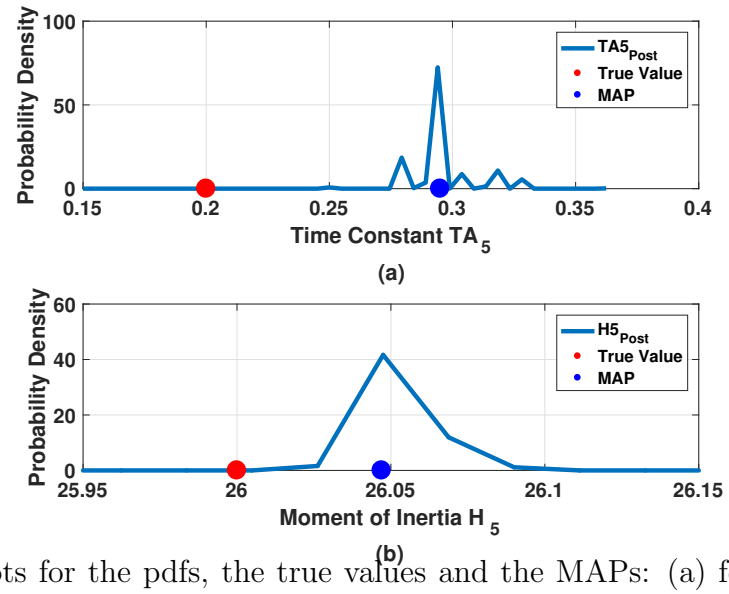


Figure 8.24: Plots for the pdfs, the true values and the MAPs: (a) for Time constant T_A; (b) for moment of inertia H.

Table 8.15: Comparison of the MAPs obtained by Different MCMC methods for Generator 5

	Prior	**True**	**Algorithm 1**	**Algorithm 2**
$H(s)$	30	26	26.05	26.07
K_A	45	40	39.8	39.7
K_E	0.9	1	0.98	0.97
K_F	0.058	0.063	0.062	0.063
D	80	82.5	82.3	82.6
R_D	1.1	1	0.98	0.97
Samples	—-	—-	10^6	10^6
Time	—-	—-	4.5 h	98 s

The functions $f_{j_k}(\xi_{j_k})$, $1 \leq j_k \leq N$, $f_{j_k,j_l}(\xi_{j_k}, \xi_{j_l})$, $1 \leq j_k \leq j_l \leq N$, etc., are called the first-order, second-order,... ANOVA components, respectively. Let $T_{j_1,\ldots,j_s}$ denote the fraction of the variance $\sigma^2(f)$ that is contributed by $f_{j_1,\ldots,j_s}$, which is defined as

$$T_{j_1,\ldots,j_s} = \frac{\sigma^2(f_{j_1,\ldots,j_s})}{\sigma^2(f)}. \tag{8.5}$$

Here T_{j_1} denotes the variance contributed by the single variable ξ_{j_1}, T_{j_1,j_2} denotes the variance contributed by the coupling effect of the two random inputs, (ξ_{j_1}, ξ_{j_2}), and so on for the higher-order terms. Sudret [88] proposes the ANOVA index under the PCE framework as follows:

$$T_{j_1} = \frac{\sigma^2(f_{j_1})}{\sigma^2(f)} = \frac{a_i^2 E[\phi(\xi_i)^2] + a_{i,j}^2 E[\phi(\xi_i, \xi_j)^2]_{i=j}}{\sigma^2}, \tag{8.6}$$

$$T_{j_1,j_2} = \frac{\sigma^2(f_{j_1,j_2})}{\sigma^2(f)} = \frac{a_{i,j}^2 E[\phi(\xi_i, \xi_j)^2]_{i\neq j}}{\sigma^2}. \tag{8.7}$$

Let us take the estimation of the time constant T_{A5} and the moment of inertia H_5 for Generator 5 under Event 2 as an example. The manufacturer's data for T_{A5} and H_5 are $\{0.3, 28\}$ while the true values are $\{0.2, 26\}$. As observed from the simulation results displayed in Fig. 8.24, the MAP estimate of H_5 is accurate while the MAP estimate of T_{A5} has almost no shift from the original data. If we calculate the ANOVA index after Step 2 in Algorithm 2, we find that the averaged T_{j_2} for the given time period for H_5 equals to 99.99% for the active power and the averaged T_{j_2} for H_5 equals to 99.6% for the reactive power. This means that T_{j_1} for T_{A5} has a much smaller value compared with T_{j_2}. The variation in the output is dominated by the H_5. Here we call T_{A5} an ineffective parameter that cannot be estimated by the proposed method.

To further demonstrate the rationality of using the ANOVA index in observability analysis, let us validate the above experiment under eight different prior values for T_{A_5}, namely

$\{0.1, 0.15, 0.18, 0.19, 0.2, 0.21, 0.22, 0.25\}$. All the other experiment settings remains unchanged. Here, we plot the posterior distributions of T_{A_5}, the true values, the prior values and the MAP values in Fig. 8.25. We can find that the MAP values remain almost the same values with the prior values for all the eight test cases with different levels of deviation. Thus, we have validated the rationality of using ANOVA index to screen out the ineffective dimensions.

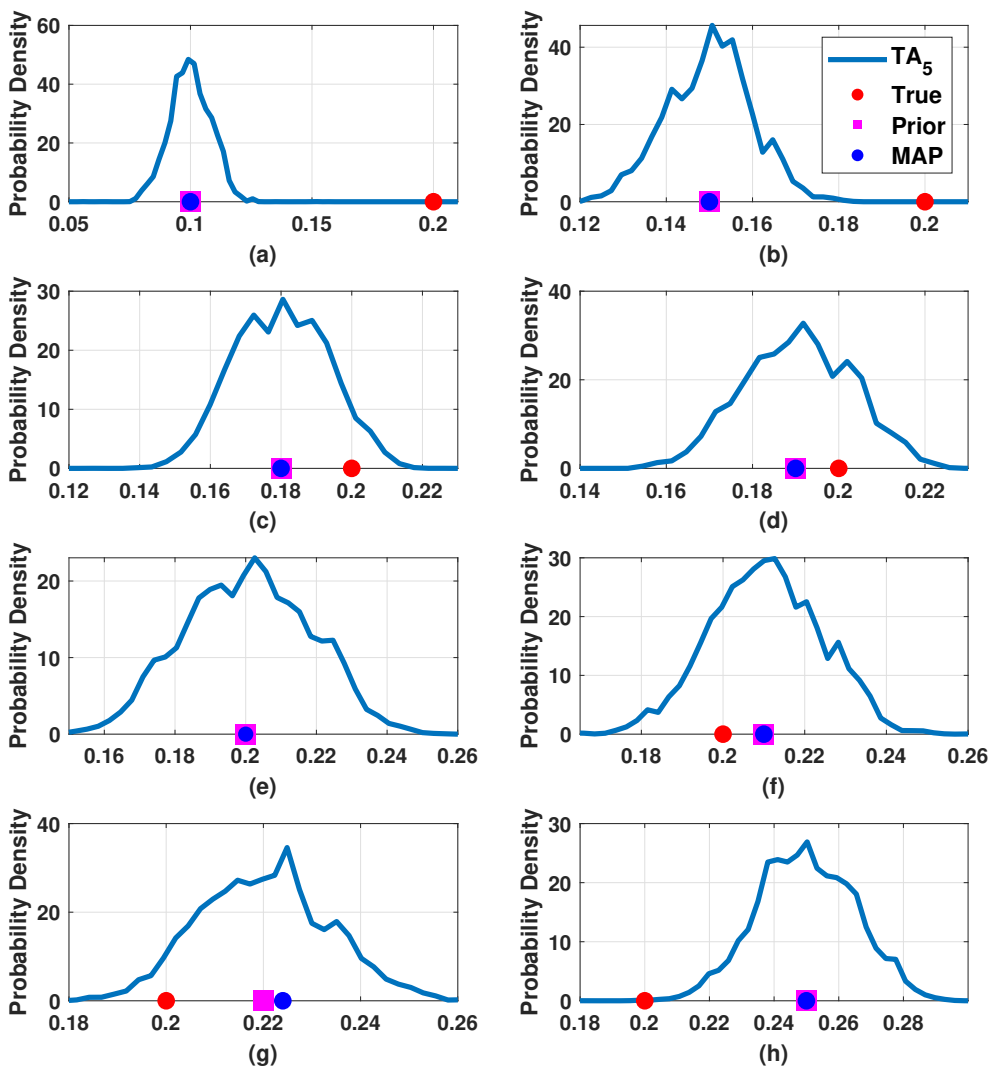


Figure 8.25: Plots for the pdfs, the true values and the MAPs for time constant T_A at (a) prior value of 0.1; (b) prior value of 0.15; (c) prior value of 0.18; (d) prior value of 0.19; (e) prior value of 0.20; (f) prior value of 0.21; (g) prior value of 0.22 and (h) prior value of 0.25.

8.4.8 Conclusions

In this part, a response-surfaced-based Bayesian inference via MCMC for power system dynamic parameter estimation with full pdfs is proposed. The Bayesian inference for the decentralized power system model is first formulated. The use of the MCMC makes this method applicable to non-Gaussian distributions. The PCE surrogate is further built to speed up the MCMC. Simulation results obtained on the New England system demonstrate the good performance of the proposed method for credible power system dynamic security analysis. Future work will be focused on further improving the convergence rate and robustness of the proposed method.

8.5 Case Studies for Real-Time Polynomial-Chaos-Expansion based Dynamic Parameter Estimation under Bayesian Inference Framework using Hybrid MCMC

Two case studies are conducted on the New England system using the aforementioned governor model to show the accuracy, the calculation efficiency and the rationality of the proposed method.

Six Dimension Case

In this part, we are demonstrating the performance of the proposed method with six key parameters that are known to influence the dynamic response of the system and can not be directly measured easily. These are the moment of inertia H, three gains in the exciter, namely, K_A, K_E, K_F, the damping ratio D, and the droop R_D. $\boldsymbol{e}$ are assumed to be identical, independent Gaussian noise with 0.01 standard deviation. The transmission line between Bus 19 and Bus 33 is removed after 0.5 seconds. The time interval is selected as 3 seconds. Let us take Generator 5 as an example. These six parameters are assumed to follow Gaussian distributions with means being the original data provided by the manufacturer, that is, $\{29, 43, 0.9, 0.058, 80, 1.1\}$ and standard deviations being 10% of the means to account for the parameter uncertainties. The true values of these parameters are $\{26, 40, 1, 0.063, 82.5, 1\}$.

For the proposed method, let's set $N_{\text{LMCMC}} = 3,000$, $N_{\text{samples}} = 40,000$ and $N_{\text{interval}} = 1,000$. The simulation results are shown in Fig. 8.27. It is found that the proposed method yields good accuracy using the MAP estimator, providing the estimated values as $\{26.09, 39.7, 1.002, 0.0638, 82.3, 1.00\}$. Even though there are strong correlation between parameters K_A, K_E and K_F, the proposed hybrid MCMC sampler still provides the global optima values. The calculation is finished in just 5 seconds. On the other hand, for the PCE-based M-H algorithm, facing these strong correlations, it needs $1,000,000$ samples to

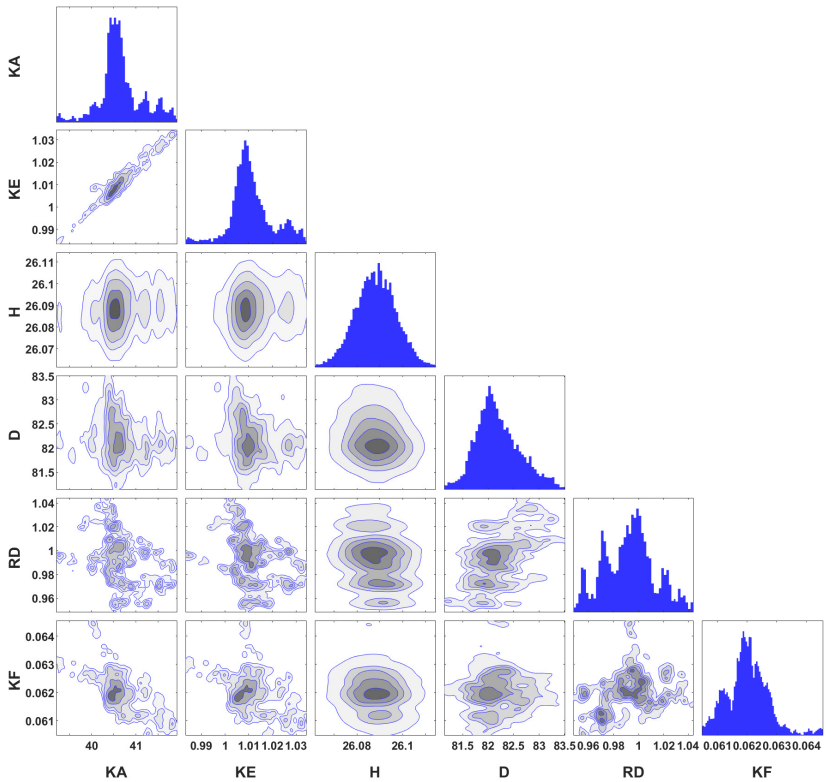


Figure 8.26: 1-D and 2-D posterior marginals of parameters for moment of inertia H, for Gain K_A, for Damping Ratio D, for Droop R_D, for Gain K_E, and for Gain K_F obtained by the PCE-based LMCMC method.

achieve a similar accuracy in the estimation. The simulation will takes around 90 seconds. For the same $1,000,000$ samples, using the time-consuming full dynamic model with traditional M-H algorithm will take almost 5 hours. It means the proposed method achieved three-magnitude speed-up factor compared with the traditional Bayesian inference with full dynamic solver under M-H algorithm. It is also 18 times faster than the PCE-based M-H algorithm. The proposed method also shows great advantages in overcoming the strong correlation between parameters.

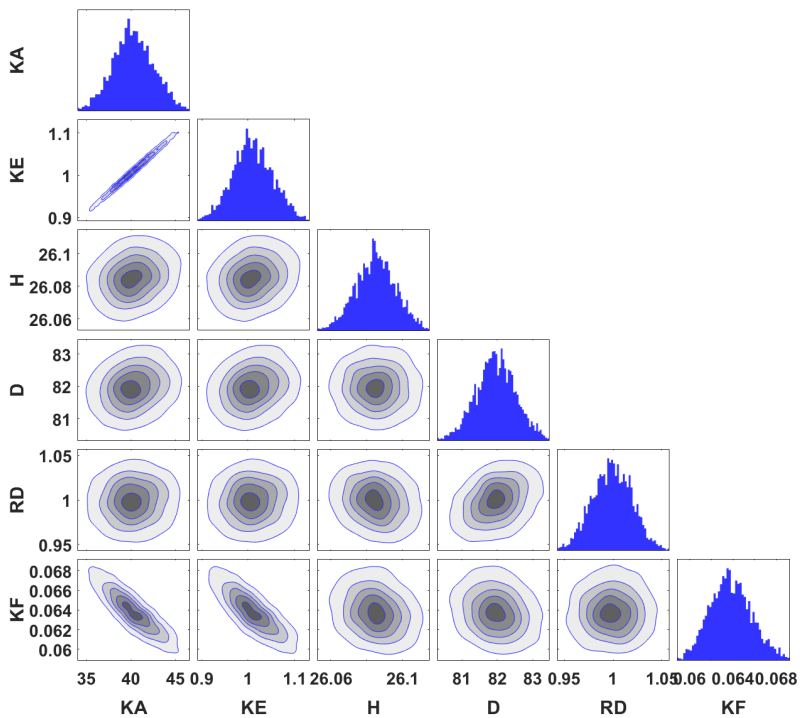


Figure 8.27: 1-D and 2-D posterior marginals of parameters for KA, KE, H, D, R_D and KF obtained by the PCE-based Hybrid MCMC method

High Dimension Case with Limitation in Observability

Not all the dynamic parameters can be estimated from the disturbance caused by the event happened in the system. Different events lead to different observabilities for different parameters [68], [67]. In this case, we are demonstrating a more complicated case with 12 parameters to be estimated. These are the moment of inertia H, three gains in the exciter, namely, K_A, K_E, K_F, the damping ratio D, the droop R_D, four generator parameters, X_d, X'_d, X_q, X'_q and two time constants T_A, T_{CH}. The prior values and the true values of these parameters are listed in Table.8.16. Since the dimension increased, let's set $N_{\text{LMCMC}} = 10,000$, $N_{\text{samples}} = 100,000$ and $N_{\text{interval}} = 1,000$ for the proposed method. The other experiment settings are the same with the previous one.

Unlike the previous case that all the parameter can be estimated, in this case, we need to use the aforementioned ANOVA index to screen out the key parameters with good observability.

Here, we take the screenshot of the ANOVA indexes with respect to all the effective dimensions for the active power P and the reactive power Q with the threshold set as 0.1%. The simulation results are plotted in Fig. 8.28 (a) and (b) separately. The dimensions with the their maximum ANOVA indexes below this threshold for all of the time period are ignored. This is because they have negligible influences on the output P and Q that, therefore, cannot guide the searching path of the MCMC sampler by their corresponding posterior likelihood values $\log(\pi_{\text{post}}(\boldsymbol{m}|\boldsymbol{d}))$. By this way, we identified 10 effective dimensions, which are $\{H$, K_A, K_E, K_F, D, R_D, X_d, X'_d, X_q, $X'_q\}$ and 2 ineffective dimensions, which are $\{T_A$, $T_{CH}\}$. Furthermore, the PCE terms contains these two ineffective dimensions can also be ignored since their PCE coefficients are too small to influence the trajectories of P and Q. Note that, different event may lead to different effective dimensions.

Now, we can estimate these 10 parameters by the proposed method. The simulation results of the 1-D and 2-D posterior marginals of these effective dimensions are plotted on Fig. 8.29. We further use MAP estimator to obtain the estimation results. The comparison studies were conducted among the PCE-based M-H algorithm, the PCE-based LMCMC algorithm and the PCE-based AM algorithm by normalized root-mean square errors (NRMSE), which is $\sqrt{(\sum_{i=1}^{N}((m_i^{est} - m_i^{True})/m_i^{True})^2)}$, and the computing time. As we can see, even the posterior distributions follow the asymmetrical Non-Gaussian distribution, the proposed method can provide very accurate estimation results with a small sample size, compared with the M-H algorithm. As the 2-D posterior marginals in Fig. 8.29 have shown that there are strong correlation between parameters. This finally leads to the biased results obtained by the PCE-based M-H algorithm and the PCE-based LMCMC algorithm. Furthermore, the PCE-based AM algorithm obtained very biased results because the dimension is too high for to the MCMC sampler to obtain a fast convergence. This long burn-in period collapsed the covariance updated in the AM strategy. Here, there is an exception, namely R_D. None of these methods provide an accurate estimation. Let's check the ANOVA index in Fig. 8.28. R_D is identified by ANOVA index for P only. However, it has the smallest ANOVA index. The largest value is just about 2%. This makes the estimation for R_D become difficulty. This phenomenon further demonstrated the rationality of using ANOVA index to screen out key parameters in the proposed method.

8.5.1 Conclusions

In this part, a PCE-based Bayesian inference via hybrid MCMC algorithm for power system dynamic parameter estimation with full pdfs is proposed. The PCE-based surrogate accelerate the full dynamic solver. The hybrid MCMC algorithm can accelerate the traditional M-H algorithm and be able to overcome bad priors and strong correlation between parameters. The simulation results reveal that the proposed can accurately and efficiently estimate dynamic parameters. Future work will be focused on exploring the relationship between ANOVA index and Lie derivative [80] in observability analysis.

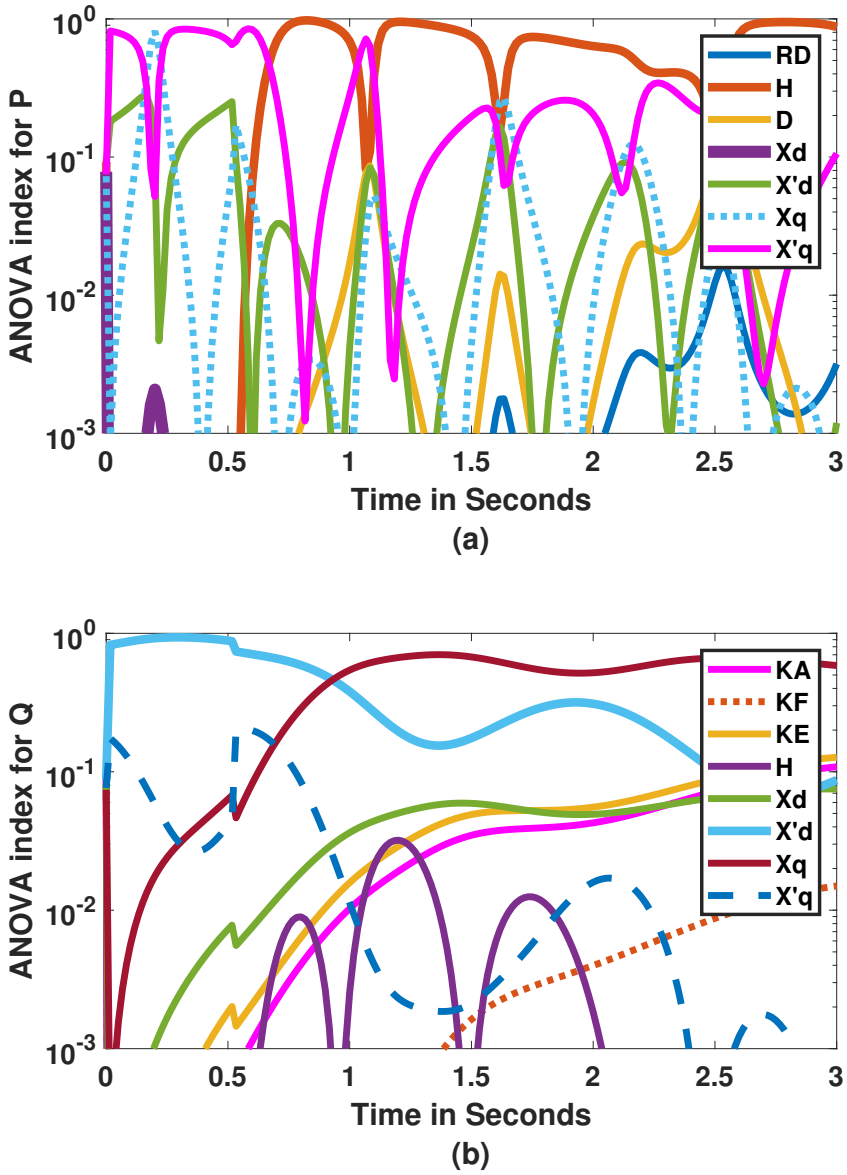


Figure 8.28: Semi-log plots of the ANOVA index of the key parameters (a) for active power P; (b) for reactive power Q.

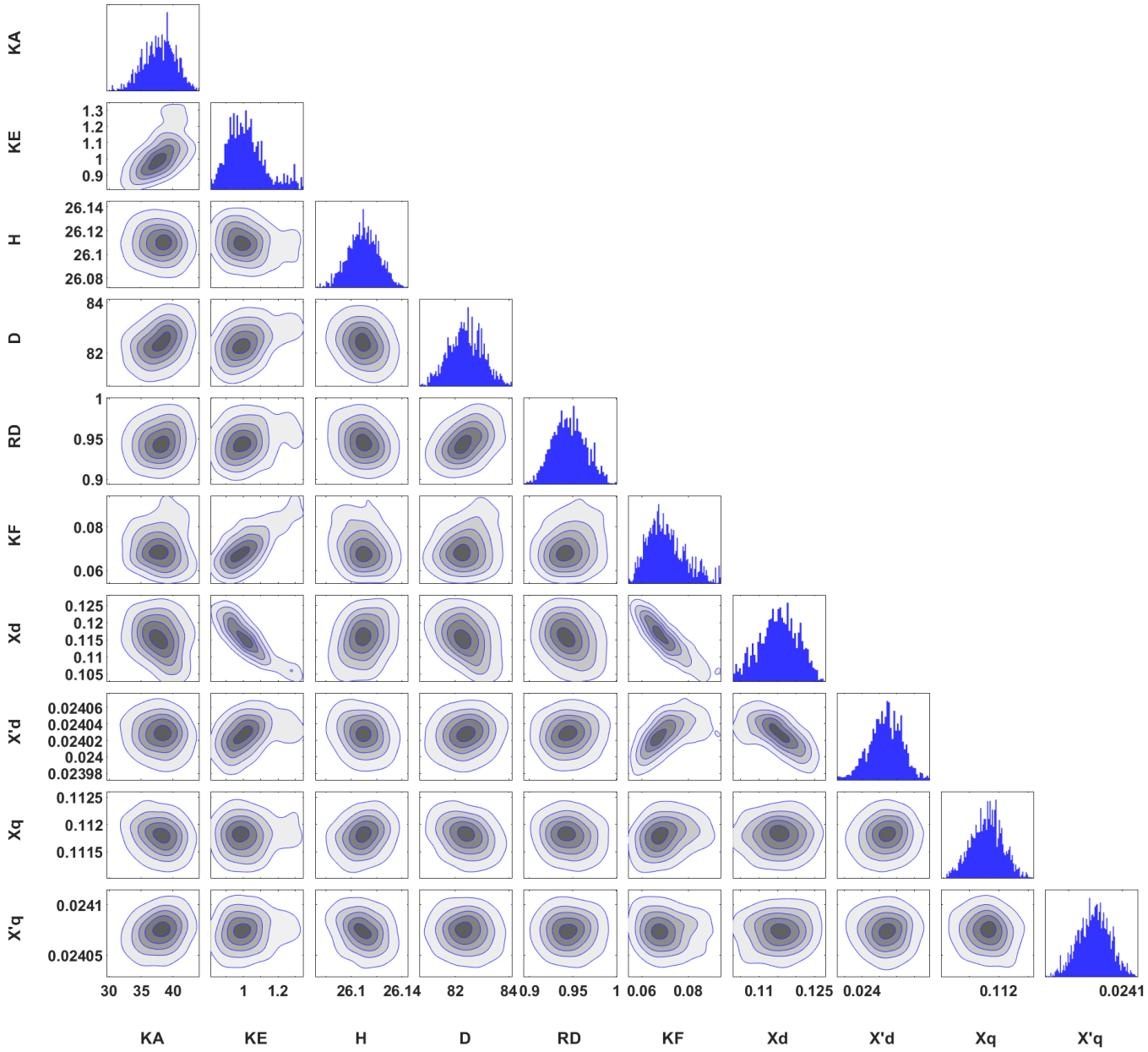


Figure 8.29: 1-D and 2-D posterior marginals of parameters for K_A, K_E, K_H, D, R_D, K_F, X_d, X'_d, X_q and X'_q obtained by the proposed method

Table 8.16: Comparison of the MAPs obtained by Different MCMC methods for Generator 5

	Prior	**True**	**Effective**	**M-H**	**LMCMC**	**CAM**	**Hybrid**
$H(s)$	29	26	**Yes**	26.17	26.09	26.1	26.11
K_A	45	40	**Yes**	39	39.2	40.2	39.4
K_E	0.9	1	**Yes**	0.92	0.92	1.2	1.02
K_F	0.058	0.063	**Yes**	0.059	0.059	0.078	0.066
D	80	82.5	**Yes**	81.0	81.7	82.5	82.5
R_D	1.1	1	**Yes**	0.93	0.91	0.93	0.952
$X_d(p.u)$	0.133	0.1218	**Yes**	0.123	0.1202	0.11	0.119
$X'_d(p.u)$	0.022	0.024	**Yes**	0.02401	0.02403	0.02402	0.02403
$X_q(p.u)$	0.11	0.1127	**Yes**	0.1119	0.1114	0.1124	0.112
$X'_q(p.u)$	0.026	0.024	**Yes**	0.02405	0.02408	0.02408	0.02407
T_A	0.15	0.2	**No**	—-	—-	—-	—-
T_{CH}	0.3	0.35	**No**	—-	—-	—-	—-
Samples	—-	—-	—-	$1e6$	$1e5$	$3e5$	$1e5$
Times(s)	—-	—-	—-	118	38.8	41.4	16.2
NRMSE	—-	—-	—-	0.128	0.139	0.333	0.076

Chapter 9

Future Work

There exist a lot of topics worth further investigation. Below we summarize a few of them.

9.1 Forward Problem

Composite power system vulnerability evaluation to cascading failures using response-surface and Monte-Carlo method For the vulnerability analysis, MC-based method is typically adopted. There are two types of samples for all the MC samples. One type of samples is close to the mean values of the corresponding probability distribution functions. This type of samples accounts for very large proportion in the whole samples. Propagating these samples through the forward solver will not bring any topology change in the power systems. Another type of samples is located at the tail of the probability distributions. The samples means low-probability, high-impact samples. Propagating these types of samples will lead to topology change in the power systems. For our proposed method, the former samples will be propagated through polynomial-chaos-based surrogate model and the latter samples will be propagated through the original full system model. By this way, we cannot only achieve very high accuracy and efficiency in vulnerability evaluation but also be able to handle the topology changes brought by the low-probability, high-impact samples.

High Dimension Dynamic Simulation with Uncertainty In the high dimension dynamic simulation area, the polynomial chaos-ANOVA based method in dealing with parametric uncertainties will be analyzed. It will be compared with the Gaussian Process emulator.

Hierarchical Uncertainty Quantification Propagating the uncertainties from the distribution system to the transmission system is almost an empty in the power system research area due to the computational complexity and the computational burden. The first stage is the uncertainty propagated in the distribution system, the output of the distribution system, also as the connection of the transmission system, will be turned into an input for the

transmission system. Then the second stage uncertainty propagation is conducted. Indeed, this can be even further propagated through interconnected regional power grid.

9.2 Inverse Problem

ANOVA index in observability analysis via Polynomial-chaos-based Kalman filter Based on our previous work in section 6 and 7, we found that the ANOVA index is an effective way to link the parameter space and the measurement space. This enables us to analyze the observability for the parameter space based on ANOVA index. Motivated by this, we are going to use the ANOVA index to build a link between state space and the measurement space. The state space can be obtained by Kalman filter and the ANOVA index is a by-product of the proposed polynomial-chaos-based Kalman filter. Then we can analyze the observability for the state variables.

A Robust Polynomial Chaos Based Kalman Filter based on the Generalized Maximum Likelihood Estimator The polynomial chaos based Kalman filter is not robust to any outlier. Under the Generalized maximum likelihood estimator based Kalman filter framework, the GM-PCKF will be proposed to compare with the former GM-KF,GM-EKF,GM-IEKF, and GM-UKF.

Chapter 10

Conclusion

Power system contains many sources of uncertainties, which are influencing its steady state and dynamic responses. Facing this problem, traditional methods, such as the Monte Carlo method and the Perturbation method, are either too time consuming or suffering from the strong nonlinearity in the system.

To solve these, this dissertation introduced several polynomial chaos based methods to replace the traditional MC based methods. For the probabilistic power flow problem, the adaptively polynomial chaos-ANOVA method is proposed. By choosing the most effective polynomial chaos bases, the calculation efficiency is greatly increased compared with the state-of-the-art generalized polynomial chaos based method without losing accuracy. For the statistical dynamic simulations, the Multi-element polynomial chaos is introduced to compensate the instability of the generalized polynomial chaos based method in the long time simulations. For the state estimation part, a novel polynomial-chaos-based Kalman filter is proposed. It shows great performances in two test systems. For dynamic parameter estimation problems, we proposed polynomial-chaos-based response-surfaced method under the Bayesian framework. Sampled by a hybrid MCMC algorithm, we can conduct dynamic parameter estimation problems accurately and efficiently.

These polynomial chaos based methods have shown great performances in several power system applications. Future work will be focused on developing more polynomial-chaos-based methods in dealing with the uncertainties in the power system forward and inverse problems.

Chapter 11

Bibliography

Bibliography

[1] B. Borkowska, "Probabilistic load flow," *IEEE Trans. Power App. Syst.*, vol. PAS-93, no. 3, pp. 752-759, May 1974

[2] J. L. Rueda, D. G. Colome, I. Erlich, "Assessment and enhancement of small signal stability considering uncertainties," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 198-207, 2009.

[3] C. Safta, R. Chen, "Efficient Uncertainty Quantification in Stochastic Economic Dispatch," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2535 - 2546, June 2016.

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

[5] T. Odun-Ayo and M. L. Crow, "Structure-Preserved Power System Transient Stability Using Stochastic Energy Functions," *IEEE Trans. Circuits Syst.*, vol. 27, no. 3, pp: 1450 - 1458, 2012.

[6] S. V. Dhople, etc, "Analysis of power system dynamics subject to stochastic power injections," *IEEE Trans. Circuits Syst.*, vol. 22, no. 9, pp: 3341-3353, 2013.

[7] 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.

[8] D. Xiu, *et al.*, "Modeling uncertainty in flow simulations via generalized polynomial chaos", *J. Comput. Phys.*, vol. 187, pp. 137-167, 2003.

[9] A. Sandu, C. Sandu, M. Ahmadian, "Modeling multibody dynamic systems with uncertainties, PartI: theoretical and computational aspect", *Multibody System Dynamics*, vol. 15, no. 4, pp: 369 - 391, 2006.

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

[11] I. Hiskens and J. Alseddiqui, "Sensitivity, Approximation, and Uncertainty in Power System Dynamic Simulation,"*IEEE Trans. Power Syst.*, vol. 21, no. 4, pp: 1808 - 1820, 2006.

[12] Kennedy, M. C. and A. O'Hagan, "Bayesian calibration of computer models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 63, no. 3, pp.425-464, 2001.

[13] Y. Xu, L. Mili, and J. Zhao "A novel polynomial-chaos-based Kalman filter," *IEEE Signal Process. Lett.*, vol. 26, no. 1, pp. 9-13, Jan. 2019.

[14] Y. Xu, *et al.* "Response-Surface-based bayesian inference for power system dynamic parameter estimation," *IEEE Trans. Smart. Grid.*, DOI: 10.1109/TSG.2019.2892464, 2019.

[15] Y. Xu, L. Mili, "Probabilistic power flow analysis based on the adaptive polynomial chaos-ANOVA method", *IEEE Power and Energy Society General Meeting*, pp. 1-5, 2018.

[16] Y. Xu, L. Mili *et al.* "Propagating uncertainty in power system dynamic simulations using polynomial chaos," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 338-348, Mar. 2019.

[17] 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.

[18] 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, April 2016.

[19] 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.

[20] W. Li, L. Guan, and D. Zhang, "An adaptive ANOVA-based PCKF for high dimensional nonlinear inverse modeling", *J. Comput. Phys.*, Vol. 258, pp. 752-772, Feb, 2014

[21] W. Gautschi, "On generating orthogonal polynomials," *SIAM J. Sci. Statist. Comput*, vol.3, pp. 289-317, 1982.

[22] F. Miao, V. Vittal, etc, "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.

[23] T. Williams and C. A. 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.

[24] F. Ni, N. Hguyen, and J. Cobben, "asis-Adaptive Sparse Polynomial Chaos Expansion for Probabilistic Power Flow, "*IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 694 - 704, Jan 2017.

[25] N. Wiener, "The homogeneous chaos," *Am. J. Math.*, Vol. 60, no. 4, pp. 897-936, 1938.

[26] X. Sun, etc, "Probabilistic load flow calculation based on sparse polynomial chaos expansion, " *IET Generation, Transmission, Distribution*, April 2018, doi: 10.1049/iet-gtd.2017.0859.

[27] R. Fisher, *Statistical Methods for Research Workers*, Oliver and Boyd, 1925.

[28] W. Hoeffding, "A class of statistics with asymptotically normal distributions", *Annal. Math. Statist.*, Vol. 19, no. 3, pp. 293-325, 1948.

[29] J. Foo, and G.E. Karniadakis, "Multi-element probabilistic collocation method in high dimensions", *J. Comput. Phys.*, Vol. 229, no. 5, pp. 1536-1557, 2010.

[30] Z. Zhang, etc, "Error Estimates for the ANOVA Method with Polynomial Chaos Interpolation: Tensor Product Functions" *SIAM J. Scientific Computing.*, Vol 34, no. 2, pp. 1165-1186, 2012.

[31] X. Yang, M. Choi, G. Lin, and G.E. Karniadakis, "Adaptive ANOVA Decomposition of Incompressible and Compressible Flows", *J. Comput. Phys.*, Vol. 231, no. 4, pp. 1587-1614, 2012.

[32] X. Wan, G. E. Karniadakis, "Multi-element generalized polynomial chaos for arbitrary probability measures," *SIAM J. Sci. Comput*, vol. 28, no. 3, pp.901-928, 2006.

[33] Z. Zhang, etc, "An Adaptive ANOVA-Based Data-Driven Stochastic Method for Elliptic PDEs with Random Coefficient", *Communications in Computational Physics*, Vol. 16, no. 2, pp. 571-598, 2014.

[34] C. Josz, S. Fliscounakis, J. Maeght, and P. Panciatici, "AC power flow data in MATPOWER and QCQP format: iTesla, RTE Snapshots, and PEGASE," http://arxiv.org/abs/1603.01533.

[35] 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.4909-4917,2013.

[36] Z. Qin, W. Li, etc, "Incorporating multiple correlations among wind speeds, photovoltaic powers and bus loads in composite system reliability evaluation," *Appl. Energy*, vol. 110, pp. 285-294, 2013.

[37] Hyungjin Choi, *et al.*, "Propagating Uncertainty in Power-System DAE Models With Semidefinite Programming," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp: 3146-3156, 2017.

[38] Su, Strunz, *et al.*, " Stochastic polynomial-chaos-based average modeling of power electronic systems," *IEEE Trans. Power Electronics.*, vol. 26, no. 4, pp. 1167 - 1171, 2011.

[39] H. Li, , *et al.*, " Voltage sensor validation for decentralized power system monitor using Polynomial chaos theory," *IEEE Trans. Instrumentation and Measurement*, vol. 60, No. 5, pp. 1633 - 1643, 2011.

[40] S. Strunz,*et al.*, " Stochastic formulation of SPICE-type electronic circuit simulation with Polynomial chaos," *ACM Trans. Modeling and Computer Simulation*, vol. 18, no. 4, pp. 1501-1522, 2008.

[41] J. R. Hockenberry *et al.*, "Evaluation of uncertainty in dynamic simulations of power system models: The probabilistic collocation method," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp: 1483 - 1491, 2004.

[42] X. Wan, G. E. Karniadakis, "An adaptive multi-element generalized polynomial chaos method for stochastic differential equations," *Journal of Computational Physics*, vol. 209, no.2, pp.617-642, 2005

[43] P. Prempraneerach, *et al.* , " Uncertainty quantification in simulations of power systems: Multi-element polynomial chaos methods," *Reliability Engineering and System Safety*, Vol. 95, no. 6, , pp 632-646, 2010.

[44] P. W. Sauer and M. A. Pai, *Power System Dynamics and Stability*, Upper Saddle River, NJ: Prentice-Hall, 1998.

[45] https://www2.ee.washington.edu/research/pstca/

[46] R.F Stengel, "Optimal control and estimation," *New York: Dover*, 1994.

[47] S. Julier and J. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE.*, Vol. 92, no. 3, pp. 401 - 422, March 2004.

[48] G. Evensen, "The Ensemble Kalman Filter: theoretical formulation and practical implementation," *Ocean Dynamics*, Vol. 53, pp.343-367, 2003

[49] M. Arulampalam, S. Maskell, N. Gordon and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Trans. Signal Processing.*, vol. 50, no. 2, pp. 174-188, 2002.

[50] O. Cappe, S. Godsill, and E. Moulines, " An Overview of Existing Methods and Recent Advances in Sequential Monte Carlo," *Proceedings of the IEEE.*, vol. 95, no. 5, pp. 899 - 924, May 2007.

[51] I. Arasaratnam S. Haykin and R. Elliott, "Discrete-Time Nonlinear Filtering Algorithms Using Gauss-Hermite Quadrature," *Proceedings of the IEEE.*, vol. 95, no. 5, pp. 953-977, May 2007.

[52] P. Dutta and R. Bhattacharya, "Nonlinear estimation of hypersonic state trajectories in Bayesian framework with polynomial chaos," *Journal of guidance, control, and dynamics.*, vol. 33, no. 6, pp. 1765-1778, 2010.

[53] P. Dutta and R. Bhattacharya, "Nonlinear estimation with polynomial chaos and higher order moment updates," American Control Conference (ACC), pp. 3142-3147, 2010.

[54] E. D. Blanchard, A. Sandu and C. Sandu, " A Polynomial Chaos-Based Kalman Filter Approach for Parameter Estimation of Mechanical Systems," *J. Dyn. Sys., Meas., Control*, vol. 132, no. 6, Nov, 2010)

[55] J. Kolansky and C. Sandu, "Enhanced polynomial chaos-based extended Kalman filter technique for parameter estimation," *J. Comput. Nonlinear Dynam*, vol. 13, no. 2, Nov, 2017.

[56] J. Li and D. Xiu, "A generalized polynomial chaos based ensemble Kalman filter with high accuracy," *J. Comput. Phys*, Vol.228, no. 15, pp. 5454-5469, August 2009.

[57] W. Slika and G. Saad, "Probabilistic identification of chloride ingress in reinforced concrete structures: polynomial chaos Kalman Filter approach with experimental verification," *ASCE J. Engineering Mechanics*, vol. 144, no. 6, pp. 1-15, June 2018.

[58] G. Saad, R. Ghanem, "Characterization of reservoir simulation models using a polynomial chaos based ensemble Kalman filter," *Water Resources*, vol. 45, no. 4, pp. 1-19, 2009.

[59] W. Slika, G. Saad, "A practical polynomial chaos Kalman filter implementation using non-linear error projection on a reduced expansion," *International Journal of Numerical Methods in Engineering*, vol. 112, no. 12, pp. 1869-1885, Dec 2017.

[60] J. B. Zhao, L. Mili, "A robust generalized-maximum likelihood unscented Kalman filter for power system dynamic state estimation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 4, pp. 578-592, 2018.

[61] J. Zhao, M. Netto and L. Mili, "A robust iterated extended Kalman filter for power system dynamic state estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3205-3216, July, 2017.

[62] X. Wan, G. E. Karniadakis, "Long-term behavior of polynomial chaos in stochastic flow simulations," *Comput. Methods Appl. Mech*, vol. 195, no. 41, pp.5582 - 5596, 2006.

[63] G. Kewlani, *et al.*, "A polynomial chaos approach to the analysis of vehicle dynamics under uncertainty," *Veh. Syst. Dyn*, vol. 50, pp.749 - 774, 2012.

[64] J. Sacks, *et al.*, "Design and analysis of computer experiments," *Statistical Science*, vol. 4, no. 4, pp.409-423, 1989.

[65] NERC., "Verification of models and data for generator excitation control system or plant volt/var control functions," 2014.

[66] NERC., "Verification of models and data for turbine/governor and load control or active power/frequency control functions" 2013.

[67] R. Huang *et al.*, "Calibrating parameters of power system stability models using advanced ensemble Kalman filter," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 2895-2905, May 2018.

[68] Z. Huang, P. Du, D. Kosterev, and S. Yang, "Generator dynamic model validation and parameter calibration using phasor measurement at the point of connection," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1939-1949, May 2013.

[69] P. M. Ashton *et al.*, "Inertia estimation of the GB power system using synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 30, no. 2, pp. 701-709, Mar. 2015.

[70] X. Cao *et al.*, "Switching Markov Gaussian models for dynamic power system inertia estimation," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3394-3403, Sept. 2016.

[71] P. Wall and V. Terzjia, "Simultaneous estimation of the time of disturbance and inertia in power systems," *IEEE Trans. Power Del.*, vol. 29, no. 4, pp. 2018-2031, Aug. 2014.

[72] N. Petra *et al.*, "A Bayesian approach for parameter estimation with uncertainty for dynamic power systems," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2735-2743, July 2017.

[73] A. Chakrabortty *et al.*, "A measurement-based framework for dynamic equivalencing of large power systems using wide-area phasor measurement," *IEEE Trans. Power Del.*, vol. 29, no. 4, pp. 2018-2031, Aug. 2014.

[74] M. A. M. Ariff, B. C. Pal, and A. K. Singh, "Estimating dynamic model parameters for adaptive protection and control in power system," *IEEE Trans. Power Syst.*, vol. 30, no. 2, pp. 829-839, Mar. 2015.

[75] I. A. Hiskens, "Power system modeling for inverse problems," *IEEE Trans. Circuits and Syst.*, vol. 51, no. 3, pp. 539-551, Mar. 2004.

[76] S. Guo *et al.*, "Adaptive parameter estimation of power system dynamic model using modal information," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 2854-2861, Nov. 2014.

[77] P. Pegoraro *et al.*, "Bayesian approach for distribution system state estimation with non-gaussian uncertainty models," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 11, pp. 2957-2966, Nov. 2017.

[78] A. K. Singh and B. C. Pal, " Decentralized dynamic state estimation in power systems using unscented transformation," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 794-804, Mar. 2014.

[79] J. Zhao and L. Mili, "Power system robust decentralized dynamic state estimation based on multiple hypothesis testing," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4553-4562, July 2018.

[80] A. Rouhani and A. Abur, "Observabiliy analysis for dynamic state estimation of synchronous machines," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3168-3175, July 2017.

[81] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*, Philadelphia, PA, USA: SIAM, 2005.

[82] Y. Marzouk and H. N. Najm, "Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems," *J. Comput. Phys.*, vol. 228, no. 6, pp. 1862-1902, Apr. 2009.

[83] Y. Marzouk and D. Xiu, "A stochastic collocation approach to Bayesian inference in inverse problems," *Commun. Comput. Phys.*, vol. 6, no. 4, pp. 826-847, Oct. 2009.

[84] M. Girolami and B. Calderhead, " Riemann manifold Langevin and Hamiltonian Monte Carlo methods," *J. R. Stat. Soc. Ser. B*, vol. 73, no.2, 123-214, 2011.

[85] B. Sengupta, K. Friston, and W. Penny, "Gradient-based MCMC samplers for dynamic causal modelling," *Neuroimage*, vol. 125, pp. 1107-1118, Jan 2016.

[86] H. Haario, E. SAKSMAN and J. TAMMINEN, " An Adaptive Metropolis Algorithm," *Bernoulli*, vol. 7, no.2, 223-242, 2001.

[87] H. Haario, M. Laine, A. Mira and E. Saksman, " DRAM: Efficient Adaptive MCMC," *Stat Comput*, vol. 16, 339-354, 2006.

[88] B. Sudret, "Global sensitivity analysis using polynomial chaos expansions", *Reliab. Eng. Syst. Saf.*, vol. 93, no. 7, pp. 964-979, July 2008.