

Polynomial-Chaos-Based Decentralized Dynamic Parameter Estimation Using Langevin MCMC

Yijun Xu, *Student Member*, Xiao Chen, Lamine Mili, *Life Fellow*, Can Huang, *Senior Member*, and Mert Korkali, *Senior Member*

Abstract—This paper develops a polynomial-chaos-expansion (PCE)-based approach for decentralized dynamic parameter estimation. Under Bayesian inference framework, the non-Gaussian posterior distributions of the parameters can be obtained through Markov Chain Monte Carlo (MCMC). However, the latter method suffers from a prohibitive computing time for large-scale systems. To overcome this problem, we develop a decentralized generator model with the PCE-based surrogate, which allows us to efficiently estimate some generator parameter values. Furthermore, the gradient of the surrogate model can be easily obtained from the PCE coefficients. This allows us to use the gradient-based Langevin MCMC in lieu of the traditional Metropolis-Hasting algorithm so that the sample size can be greatly reduced. Simulations carried out on the New England system reveal that the proposed method can achieve a speedup factor of three orders of magnitude as compared to the traditional method without losing the accuracy.

Index Terms—Dynamic parameter estimation, polynomial chaos expansion, Bayesian inference, Langevin MCMC.

I. INTRODUCTION

PARAMETER estimation for dynamic generator models is of paramount importance for power system dynamic security analysis. However, the traditional offline stage-testing-based methods are very costly, time-consuming, and labor-intensive [1]. This has prompted a growing number of researchers to develop online, low-cost phasor-measurement-unit (PMU)-based methods. To overcome this difficulty, some of these methods estimate the generator moment of inertia [2], [3], while others estimate other generator parameters using the Kalman filtering approach [1], [4], [5]. Unfortunately, the former methods calculate only point estimated values of the inertia parameters without providing their confidence intervals while the latter methods suffer from the weaknesses of Kalman filters such as the Gaussian assumption for the process and measurement noise and the slow convergence rate [1].

Unlike the above methods, Bayesian-inference based methods provide maximum-a-posteriori (MAP) parameter estimated

values along with the corresponding posterior probability distributions [6], [7]. Furthermore, they are not limited to Gaussian assumptions. However, the posterior distribution in the Bayesian inference is typically obtained by MCMC-based methods, which require prohibitive computing times. This prompted some researchers to propose methods that keep the approximated Gaussian assumption while taking the local optimization approach [6]. Unfortunately, in this approach, the local optima may be highly biased when the initial guess is far from the true value or the posterior distribution is non-Gaussian.

To overcome the aforementioned weaknesses, this paper develops a polynomial-chaos-expansion (PCE)-based Bayesian inference algorithm for power system dynamic parameter estimation by applying the Bayesian framework for the parameter estimation of the decentralized generator model [8]. This eliminates the uncertainties of the line and the transformer models and the loads within a centralized system model [8], [9]. To deal with the non-Gaussian posterior distribution, we resort to the MCMC method while speeding up its convergence rate via the gradient-based Langevin MCMC. It turns out that the gradient information of the latter can be obtained easily from the PCE surrogate, which greatly accelerates the dynamical solver. Simulation results carried out on the New-England test system reveal that our proposed method can accurately and simultaneously estimate several key generator dynamic model parameters with almost three-magnitude improvement in computing speed compared to the traditional MCMC-based Bayesian inference method.

II. PROBLEM FORMULATION

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

1) *Decentralized Generator Model*: The decentralized synchronous generator model, its the measurement model, and the notations are all following the work of Zhao and Mili [8], where the two-axis model with a IEEE-DC1A exciter and a TGOV1 turbine-governor is considered [10]. Here, when a disturbance occurs in the system, the local PMU at the i th 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 i th generator as model outputs, which are expressed as

$$P_{ei} = V_{di}I_{di} + V_{qi}I_{qi} + e_{Pi}, \quad (1)$$

$$Q_{ei} = -V_{di}I_{qi} + V_{qi}I_{di} + e_{Qi}, \quad (2)$$

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

X. Chen, C. Huang, and M. Korkali are with Lawrence Livermore National Laboratory, Livermore, CA 94550 USA (e-mail: chen73@llnl.gov; huang38@llnl.gov; korkali1@llnl.gov).

This work was sponsored by the U.S. Department of Energy through its Laboratory Directed Research and Development program and Advanced Modeling Grid Research program. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

where e_{P_i} and e_{Q_i} are the measurement noise. By this way, once we capture the local voltage phasor $V_i \angle \theta_i$ for the i th 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 i th generator model's output using the differential and algebraic equations given in [8]. This is the decentralized model used in this paper.

2) *Bayesian Inference*: Tarantola [12] formulated the Bayesian inference model as

$$\mathbf{d} = \mathbf{f}(\mathbf{m}) + \mathbf{e}, \quad (3)$$

where \mathbf{d} contains the observations, consisting of d_1 that contains the active power measurements, P_{ei} , and d_2 that contains the reactive power measurements, Q_{ei} ; $\mathbf{m} \in \mathbb{R}^N$ are the parameters to be estimated; N is the number of parameters to be estimated that depends on the specific applications; $\mathbf{f}(\cdot)$ is the vector-valued forward function that represents the aforementioned differential-algebraic equations, which map the model parameters \mathbf{m} to the observations \mathbf{d} ; $\mathbf{e} \in \mathbb{R}^2$ stands for the measurement error vector whose components are assumed to be mutually independent random variables with the joint probability density functions π_e defined as $\pi_e = \prod_{i=1}^2 \pi_{e_i}(e_i)$. 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 \mathbf{m} is given by $\pi_{\text{prior}}(\mathbf{m}) = \prod_{i=1}^N \pi_i(m_i)$. Note that \mathbf{e} and \mathbf{m} are also assumed to be mutually independent. Given the observation \mathbf{d} , the posterior pdf $\pi_{\text{post}}(\mathbf{m}|\mathbf{d})$ for the parameters \mathbf{m} is derived as [12]

$$\pi_{\text{post}}(\mathbf{m}|\mathbf{d}) \propto \pi_{\text{like}}(\mathbf{d}|\mathbf{m})\pi_{\text{prior}}(\mathbf{m}). \quad (4)$$

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

$$\pi_{\text{like}}(\mathbf{d}|\mathbf{m}) = \prod_{i=1}^2 \pi_{e_i}(d_i - f_i(\mathbf{m})). \quad (5)$$

Given a set of parameters contained in \mathbf{m} , we obtain the trajectories of P_{ei} and Q_{ei} from the forward solver $\mathbf{f}(\cdot)$. 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

$$\log \pi_{e_i}(d_i - f_i(\mathbf{m})) = \sum_{t=0}^{t_{\text{end}}} \log \pi_{e_i}^t(d_i^t - f_i^t(\mathbf{m})). \quad (6)$$

Thanks to the high 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 (4) can be put into the following form:

$$\log(\pi_{\text{post}}(\mathbf{m}|\mathbf{d})) \propto \sum_{i=1}^2 \log \pi_{e_i}(d_i - f_i(\mathbf{m})) + \sum_{i=1}^N \log(\pi_i(m_i)), \quad (7)$$

yielding the MAP estimator defined as

$$\hat{\mathbf{m}}_{\text{MAP}} = \arg \min_{\mathbf{m}} -\log(\pi_{\text{post}}(\mathbf{m}|\mathbf{d})). \quad (8)$$

Note that due to the nonlinearity of $\mathbf{f}(\cdot)$, even if the prior assumption is Gaussian, $\pi_{\text{post}}(\mathbf{m}|\mathbf{d})$ 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 \mathbf{m} . The Langevin MCMC that we adopted will be described in Section IV.

III. PCE-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 by a PCE-based surrogate.

A. Review of the Generalized Polynomial Chaos Expansion

Introduced by Wiener and further developed by Xiu and Karniadakis [13], the generalized polynomial chaos expansion has been shown to be a cost-effective tool in modeling response surfaces [14]–[16]. 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 $\boldsymbol{\xi} = [\xi_1, \xi_2, \dots, \xi_N]$ be a vector of random variables following a standard probability distribution (e.g. the Gaussian or the beta distribution), to which, as shown in [13], a unique orthogonal polynomial is associated. Let $\Phi_i(\xi_1, \xi_2, \dots, \xi_N)$ denote this procedure's corresponding polynomial chaos basis and let a_i denote the i th polynomial chaos coefficient. Formally, we have

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

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. A relatively low maximum polynomial chaos order (typically 2) is found to provide output results with enough accuracy [15]–[17]. From the polynomial chaos coefficients, the mean, μ , and the variance, σ^2 , of the output z can be determined as $\mu = a_0$, $\sigma^2 = \sum_{i=1}^{N_P} a_i^2 E[\Phi_i^2]$, where $E[\cdot]$ is the expectation operator.

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

$$\int \Phi_r(\xi) \Phi_s(\xi) d\lambda = \begin{cases} 0 & \text{if } r \neq s, \\ > 0 & \text{if } r = s. \end{cases} \quad (10)$$

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

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

2) *Construction of the Polynomial Chaos Basis*: A set of multi-dimensional polynomial chaos basis functions can be constructed as the tensor product of the one-dimensional polynomial chaos basis associated with each input random variable. Formally, we have $\Phi(\boldsymbol{\xi}) = \Phi(\xi_1) \otimes \Phi(\xi_2) \otimes \cdots \otimes \Phi(\xi_N)$, where $\Phi(\xi_i)$ denotes the one-dimensional polynomial chaos basis for the i th random variable.

3) *Collocation Points*: Collocation points can be regarded as a finite sample of $\boldsymbol{\xi} = [\xi_1, \xi_2, \dots, \xi_N]$ that are chosen to approximate the polynomial chaos coefficients. The elements of the collocation points are generated by using the union of the zeros and the roots of one higher-order, one-dimensional polynomial for every random variable [13], [16]. Here, sparse-grid method is suggested to generate collocation-point combinations [13].

B. Building PCE-based Surrogate for Dynamic Power Systems

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

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

$$m_i = F_i^{-1}(T(\xi_i)), \quad (11)$$

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 of (9).
(c) Construct M combinations of collocation points and put them into the polynomial chaos basis ($M \times (N_P + 1)$) matrix \mathbf{H}_{pc} . Formally, we have

$$\mathbf{H}_{pc} = \begin{pmatrix} \Phi_0(\boldsymbol{\xi}_1) & \Phi_1(\boldsymbol{\xi}_1) & \cdots & \Phi_{N_P}(\boldsymbol{\xi}_1) \\ \Phi_0(\boldsymbol{\xi}_2) & \Phi_1(\boldsymbol{\xi}_2) & \cdots & \Phi_{N_P}(\boldsymbol{\xi}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_0(\boldsymbol{\xi}_M) & \Phi_1(\boldsymbol{\xi}_M) & \cdots & \Phi_{N_P}(\boldsymbol{\xi}_M) \end{pmatrix}; \quad (12)$$

- (d) Compute the power system dynamic model output for the selected collocation points to get the $(M \times 1)$ output Z matrix given by $Z = (z(t, \boldsymbol{\xi}_1) \ z(t, \boldsymbol{\xi}_2) \ \cdots \ z(t, \boldsymbol{\xi}_M))^T$.
(e) Estimate the unknown coefficients A based on the collocation points that are selected and the model output from

$$Z = \mathbf{H}_{pc}A, \quad (13)$$

where A is the $(N_P + 1)$ coefficient vector expressed as

$$A = (a_0(t) \ a_1(t) \ \cdots \ a_{N_P}(t))^T. \quad (14)$$

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

$$\hat{A} = \arg \min_{\hat{A}} r^T r, \quad (15)$$

which yields $\hat{A} = (\mathbf{H}_{pc}^T \mathbf{H}_{pc})^{-1} \mathbf{H}_{pc}^T Z$.

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 in the polynomial form.

C. Incorporating the Polynomial Chaos Expansion into the Bayesian Inference Framework

In the gPC-based Bayesian inference [14], [15], we use the approximated gPC solution in (9) to replace the exact forward solver solution $\mathbf{f}(\mathbf{m})$ in (7) as

$$\log(\pi_{\text{post}}(\boldsymbol{\xi}|\mathbf{d})) \propto \sum_{i=1}^2 \log \pi_{e_i}(d_i - z_i(\boldsymbol{\xi})) + \sum_{i=1}^N \log(\pi_i(\xi_i)). \quad (16)$$

Once a sample point $\boldsymbol{\xi}$ is proposed in the 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 (11) to map $\boldsymbol{\xi}$ back into \mathbf{m} to obtain the posterior distribution of \mathbf{m} .

IV. APPLICATION OF THE PROPOSED METHOD

This section shows the proposed PCE-based Bayesian inference via the LMCMC algorithm in the dynamic parameter estimation problem.

A. Review of the MCMC Methods

1) *Metropolis-Hastings (MH) algorithm*: As the most classical MCMC algorithm, starting from the initial guess, \mathbf{m}_0 , the MH method employs a proposal distribution, $q(\mathbf{m}_k, \cdot)$, at each sample point \mathbf{m}_k to generate a proposed sample point \mathbf{m}_{k+1} . Once generated, the sample point is either accepted or rejected by the MH method. This procedure is then applied to the next sample point, yielding a chain of sample points from the posterior pdf $\pi_{\text{post}}(\mathbf{m})$ [12]. To further stabilize the numerical computation of the MH algorithm, the posterior pdf is typically transformed into the log-form as suggested in [7].

2) *The Langevin MCMC algorithm (LMCMC)*: Unlike the MH algorithm that simply uses the proposed distribution function driven by random variables to approach the posterior distribution asymptotically, the LMCMC algorithm employs a combination of two mechanisms to generate the samples of a random walk [18]. It first generates new samples via the Langevin dynamics mechanism, which uses the evaluations of the gradient of the target probability density function, then it accepts or rejects new samples using the MH mechanism. This process can be described by

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \frac{\varepsilon^2}{2} \nabla \log(\pi_{\text{post}}(\mathbf{m}_k|\mathbf{d})) + \varepsilon \boldsymbol{\rho}. \quad (17)$$

Here $\varepsilon > 0$ is a fixed step size, $\nabla \log(\pi_{\text{post}}(\mathbf{m}_k|\mathbf{d}))$ 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 MH mechanism. This Langevin dynamics enables the LMCMC a better convergence rate than the MH algorithm.

B. The Proposed PCE-based LMCMC

Here, let us introduce the LMCMC under the PCE framework. From (16), we can see that the original complicated dynamic solver at \mathbf{m} space is simplified into the polynomial forms at ξ space. This can greatly simplify the derivation of the gradients of the posterior likelihood. Let us assume that $e \sim \mathcal{MVN}(\mathbf{0}, \Sigma_e)$, where $\Sigma_e = \text{diag}\{\sigma_{e_1}^2, \sigma_{e_2}^2\}$. Under Gaussian assumption for the prior \mathbf{m} , Normalized Hermite polynomials are chosen for ξ , which means $\xi \sim \mathcal{MVN}(\mathbf{0}, \Sigma_\xi)$ and Σ_ξ is a $N \times N$ identity matrix. From (6) and (16), we have

$$\log(\pi_{\text{post}}(\xi|\mathbf{d})) \propto \sum_{i=1}^2 \sum_{t=0}^{t_{\text{end}}} -\frac{(d_i^t - z_i^t(\xi))^2}{2\sigma_{e_i}^2} + \sum_{i=1}^N -\frac{(\xi_i)^2}{2}. \quad (18)$$

Then, we derive the gradients as

$$\nabla \log(\pi_{\text{post}}(\xi|\mathbf{d})) = \sum_{i=1}^2 \sum_{t=0}^{t_{\text{end}}} \frac{(d_i^t - z_i^t(\xi))}{\sigma_{e_i}^2} \nabla z_i^t(\xi) + \mathbf{g}(\xi), \quad (19)$$

where

$$\mathbf{g}(\xi) = [-\xi_1 \quad -\xi_2 \quad \cdots \quad -\xi_N]^\top \quad (20)$$

represents the gradients from prior information. Here, let us derive the gradient of $z_i^t(\xi)$. Following [17], let us reorganize (9) with $P = 2$ at time point t as

$$z_i^t = a_0^t + \left(\sum_{i=1}^N a_{i,i}^t \xi_i + \sum_{i=1}^N a_{i,i}^t \frac{\xi_i^2 - 1}{\sqrt{2}} \right) + \sum_{j=1}^N \sum_{1 \leq i < j} a_{i,j}^t \xi_i \xi_j. \quad (21)$$

From this simple polynomial form, we can obtain ∇z_i^t via

$$\begin{aligned} \nabla z_i^t &= \begin{bmatrix} \frac{\partial z_i^t}{\partial \xi_1} \\ \frac{\partial z_i^t}{\partial \xi_2} \\ \vdots \\ \frac{\partial z_i^t}{\partial \xi_N} \end{bmatrix} = \begin{bmatrix} a_{1,1}^t \\ a_{2,2}^t \\ \vdots \\ a_{N,N}^t \end{bmatrix} + \sqrt{2} \begin{bmatrix} a_{1,1}^t \xi_1 \\ a_{2,2}^t \xi_2 \\ \vdots \\ a_{N,N}^t \xi_N \end{bmatrix} \\ &+ \begin{bmatrix} 0 & a_{1,2}^t & \cdots & a_{1,N-1}^t & a_{1,N}^t \\ a_{1,2}^t & 0 & \cdots & a_{2,N-1}^t & a_{2,N}^t \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{1,N-1}^t & a_{2,N-1}^t & \cdots & 0 & a_{N-1,N}^t \\ a_{1,N}^t & a_{2,N}^t & \cdots & a_{N-1,N}^t & 0 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_{N-1} \\ \xi_N \end{bmatrix}. \end{aligned} \quad (22)$$

This form remains the same for both the active power P and the reactive power Q with different PCE coefficients. Based on (17), we can propose a new sample ξ_{k+1} from the current sample ξ_k via the PCE-based LMCMC as

$$\xi_{k+1} = \xi_k + \frac{\varepsilon^2}{2} \nabla \log(\pi_{\text{post}}(\xi_k|\mathbf{d})) + \varepsilon \rho. \quad (23)$$

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

C. Application in the Decentralized Parameter Estimation

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.

Algorithm 1 Polynomial-Chaos-Expansion-based Bayesian Inference using the Langevin MCMC

- 1: Choose the initial guess of the parameters \mathbf{m}_0 from the manufactured data as the Bayesian prior $\mathbf{m}_{\text{prior}}$;
 - 2: Mapping \mathbf{m} into ξ via (11), build the PCE surrogates as the response surface of the decentralized dynamic model;
 - 3: Compute $\log(\pi_{\text{post}}(\xi_0|\mathbf{d}))$ from the PCE surrogate via (16);
 - 4: **for** $k = 0, \dots, N_{\text{samples}} - 1$ **do**
 - 5: Compute $\nabla \log(\pi_{\text{post}}(\xi_k|\mathbf{d}))$ from (19);
 - 6: Generate a new sample ξ_{k+1} from (23);
 - 7: Compute $\log(\pi_{\text{post}}(\xi_{k+1}|\mathbf{d}))$ from the PCE surrogate via (16);
 - 8: Calculate the correction factor $c = \frac{q(\xi_{k+1}, \xi_k)}{q(\xi_k, \xi_{k+1})}$;
 - 9: Compute $\alpha(\xi_k, \xi_{k+1})$ defined as

$$\alpha(\xi_k, \xi_{k+1}) = \log(\min\{1, \frac{\pi_{\text{post}}(\xi_{k+1}|\mathbf{d})}{\pi_{\text{post}}(\xi_k|\mathbf{d})} \cdot c\});$$
 - 10: Draw $u \sim \mathcal{U}([0, 1])$;
 - 11: **if** $\log(u) < \alpha(\xi_k, \xi_{k+1})$ **then**
 - 12: Accept: Set $\xi_{k+1} = \xi_{k+1}$;
 - 13: **else**
 - 14: Reject: Set $\xi_{k+1} = \xi_k$;
 - 15: **end if**
 - 16: **end for**
 - 17: Mapping samples of ξ back to \mathbf{m} , then plot the PDF of $\pi_{\text{post}}(\mathbf{m}|\mathbf{d})$, and find the MAP points.
-

V. CASE STUDIES

Case studies are conducted on the New England system using the aforementioned governor model as described in [10]. The accuracy and calculation efficiency of the proposed PCE-based LMCMC are compared with the PCE-based MH algorithm and traditional full dynamic model with the MH algorithm.

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, which cannot be directly measured easily. These are the moment of inertia H , the three gains of the exciter, namely, K_A , K_E , K_F , the damping ratio D , and the droop R_D . 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 the Gaussian distributions with means being the original data provided by the manufacturer, that is, $\{29, 39, 0.9, 0.058, 80, 1.1\}$ and the standard deviations being 10% of the means to account for the parameter uncertainties. The true values of these parameters are $\{26, 40, 1, 0.063, 82.5, 1\}$. The simulation results are shown in Fig. 1. It is found out that the proposed method yields good accuracy using the MAP estimator, which provides the estimated values as $\{26.09, 40.3, 1.008, 0.062, 82.2, 1.00\}$. Even in presence of local optima, the MCMC sampler still provides the global optimum values. However, due to the strong correlation between some parameters, such as the 2-D posterior marginals between K_A and K_E , the sample size required in MCMC is large. Thanks to the PCE surrogate, every sample can be computed with negligible times. Using the MH algorithm that needs 10^6 samples, the traditional Bayesian method with full dynamic model takes almost 5 hours to complete the computation while the PCE-based MH algorithm takes only 2 minutes. Because of the fast convergence rate of the LMCMC, the PCE-based LMCMC can converge using 5×10^4 samples with a solution time of 15 seconds. This means that the proposed method achieves a three-orders-of-magnitude speedup factor compared to the traditional Bayesian inference with full dynamic solver under the MH algorithm. It is also 7 times faster than the PCE-based MH algorithm.

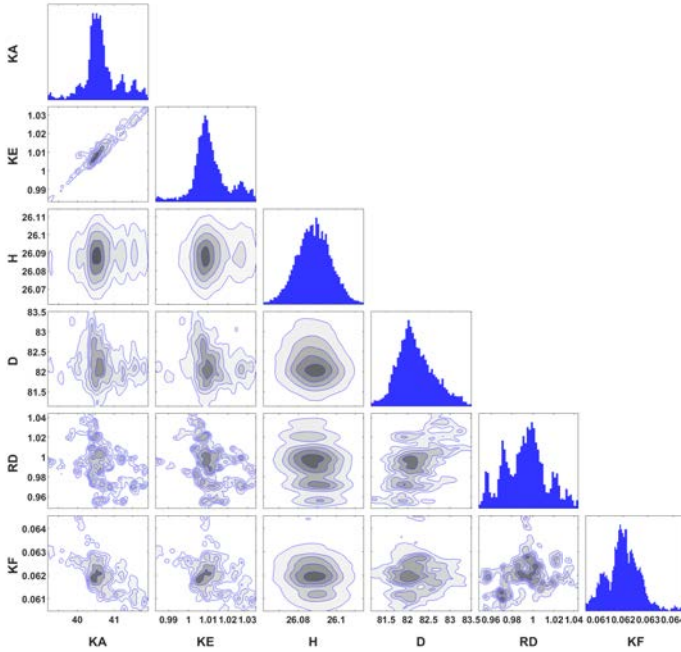


Fig. 1. 1-D and 2-D posterior marginals of parameters for Gain K_A , for Gain K_E , for moment of inertia H , for Damping Ratio D , for Droop R_D , and for Gain K_F obtained by the proposed method.

VI. CONCLUSIONS

In this paper, a PCE-based Bayesian inference using the LMCMC is proposed for power system dynamic parameter

estimation with full description of the pdfs. Specifically, the Bayesian inference for the decentralized power system model is first formulated. It is then made applicable to non-Gaussian distributions via the MCMC. The latter is then sped up using the PCE surrogate while the LMCMC reduces the sample size required by the MH algorithm. The PCE surrogate allows us to easily calculate the gradients for the Langevin dynamics. The simulation results demonstrate the good performance of the proposed method. Future work will be focused on further improving the robustness of the proposed method to outliers.

REFERENCES

- [1] 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.
- [2] 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.
- [3] P. Wall and V. Terzija, "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.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] Y. Xu *et al.*, "Propagating uncertainty in power system dynamic simulations using polynomial chaos," *IEEE Trans. Power Syst.*, 2018, DOI: 10.1109/TPWRS.2018.2865548.
- [10] P. W. Sauer and M. A. Pai, *Power System Dynamics and Stability*, Upper Saddle River, NJ: Prentice-Hall, 1998.
- [11] A. Rouhani and A. Abur, "Observability analysis for dynamic state estimation of synchronous machines," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3168-3175, July 2017.
- [12] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*, Philadelphia, PA, USA: SIAM, 2005.
- [13] D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University Press, Princeton, NJ, 2010.
- [14] 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.
- [15] 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.
- [16] 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.
- [17] Y. Xu, and L. Mili, "Probabilistic Power Flow Analysis based on the Adaptive Polynomial Chaos-ANOVA Method," *IEEE Power and Energy Society General Meeting (PESGM)*, 2018.
- [18] B. Sengupta, K. Friston, and W. Penny, "Gradient-based MCMC samplers for dynamic causal modelling," *Neuroimage*, vol. 125, pp. 1107-1118, Jan. 2016.