Frequency Prediction of Power Systems in FNET based on State Space Approach and Uncertain Basis Functions

Jin Dong, Student Member, IEEE, Xiao Ma, Student Member, IEEE, Seddik M. Djouadi, Member, IEEE, Husheng Li, Member, IEEE, and Yilu Liu, Fellow, IEEE.

Abstract—In this paper, we discuss the modeling and prediction of power frequency. Power frequency is one of the most essential parameters in the monitoring, control, and protection of power systems and electric equipments because when a significant disturbance occurs in a power system, the frequency varies in time and space. It is critical to employ a dependable model in order to optimize the efficiency and reliability of power systems in the Frequency Monitoring Network (FNET), and thus, prevent frequency oscillation in power grid. This paper describes the use of a state-space model and basis functions to predict power frequency. In the state-space method, expectation maximization (EM) and prediction error minimization (PEM) algorithms are used to dynamically estimate the model’s parameters. In the basis functions method, we employ random basis functions to predict the frequency. The algorithms are easy to implement online, having both high precision and a short response time. Numerical results are presented to demonstrate that the proposed techniques are able to achieve good performance in frequency prediction.

Index Terms—Power frequency, Prediction, Uncertain basis function, Kalman Filtering, FNET, Protection.

I. INTRODUCTION

Contemporary society’s dependence on a reliable power grid cannot be overstated; when power grids fail, significant societal disruptions and economic losses inevitably follow. However, controlling and protecting power systems requires an accurate estimate of frequencies in real time [1]. As is well known, the frequency used in power systems is a universal parameter across the entire interconnected power grid [2], that provides a great deal of information about the power system’s dynamics [3].

If fault locations (grid nodes whose frequency deviations are over the limit) are not predicted accurately, the scale of disruption grows larger and larger with the number of nodes that are integrated over the power grid. In addition, as the size of power systems increases, they also become more and more vulnerable to sharp frequency fluctuation. It is therefore critically important that the modeling and prediction of frequency used for power systems be effective against these threats.

For smart grid, one effective approach to prevent failures from happening is to be able to predict fault locations. In [4], the authors developed a method to identify the most probable failure modes for a given network. Further, in order to detect a line outage and identify a network parameter error, the authors of [5], [6], [7] proposed to take advantage of the phasor information collected from the smart grid. However, as far as we know, no studies have used the analysis and prediction of frequency information to predict failure in the smart grid.

There has been much research devoted to the FNET [8], which is an internet-based, wide-area sensor network consisting of high-precision disturbance recorders and a central processing server. Having high dynamic accuracy, the FNET can track frequencies even in the presence of oscillations [9]; when a significant disturbance occurs in a power system, the frequency will vary in time and space [10]. It is reported that reliability is not directly impacted by slow frequency oscillation unless it deviates enough to cause undesired load shedding, generator tripping, etc [11]. However, the focus of our paper is different as we consider the online prediction for the frequency of 1s or 10s later, which is short-term. Moreover the fluctuations caused by the whole power grid are still there and reflected by the frequency of each node. Therefore, by monitoring frequency deviations, the FNET data can be used to predict system failures and hence prevent wide-area blackouts.

Accurate prediction of frequency is important for two reasons: First, it should be noted that the propagation of frequency oscillations through the power grid, together with the catastrophic cascading failure caused by node failure, is likely to damage electronic equipment and produce power outages in a wide area, thus resulting in substantial economic loss—even possibly jeopardizing national security [12]. Thus, there is a pressing need to predict and restrain node failure and frequency oscillation for power system reliability. Based on the FNET data, we recognize and predict the fluctuation of the nodes, thus, we can use the predictions as benchmark or test values to determine whether there is an abnormality. Second, if the FNET should suddenly fail to provide a clean and continuous measurement for one node, the difference between the predicted values and the collected data from the FNET can indicate a measurement problem or the presence of a spoofing signal introduced during transmission through the Internet; our method will function as an emergency backup to provide optimal predictions to replace “bad” frequencies.

In the past, several frequency estimation methods such as Kalman filtering (KF) [13][14], artificial neural network (ANN) [15][16], least squares (LS) techniques [17], Newton-
type algorithms [18], and adaptive notch filters [19] have been reported in the literature. However, the inclusion of more interconnected power components into the power grid has introduced some challenges for the direct adoption of these techniques. Therefore, most of the aforementioned estimation techniques fail or have degraded performance for predicting power frequency.

Because there is still a lack of fundamental analytical results about frequency prediction, the main contribution of this paper is to propose two general techniques for directly using the collected frequency data from FNET to predict fault locations where the deviation of frequency is over the limit.

A. Methods

First, we present a state-space method for power frequency prediction. We employ two different identification algorithms - expectation maximization (EM) and prediction error minimization (PEM) - to estimate the parameters in the state-space model. This method permits a direct prediction of frequency without having to linearize or use simplifying assumptions on the signal’s model, thus enabling the system to accurately predict small as well as large frequency deviations.

Moreover, we propose a second method based on uncertain basis functions for modeling and prediction purposes. In traditional works, basis functions (e.g., radial basis function [20], and wavelets [21]) are fixed; instead, we assume that the basis functions are random. The modeling and prediction of signals using uncertain basis functions were proposed recently in [22]. Here, we employ this method to predict power frequency. We consider basis functions with two different distributions, Uniform and Gaussian, and show that the prediction errors are small for each distribution.

B. Example

To demonstrate the performance of the proposed stochastic models, we collect a set of real measurement data [23] from the FNET as shown in Fig. 1. For example, there are 1400 sampling points in one randomly picked segment and the samples are quantized into 10 levels for frequency within (59.970 Hz, 60 Hz) as shown in Fig. 2. To simplify the analysis, we define the binary state of each node to be State 0 when the frequency is within (59.985 Hz, 60 Hz) (low frequency deviation) and State 1 otherwise. According to Fig. 2, the probability of being in State 0 or 1 is almost identical. Thus, the proposed prediction algorithms are essential to isolate a failure node predicted to be State 1 before a significant disturbance propagates into the whole power grid.

The remainder of this paper is organized as follows. Section II introduces the FNET system and its application to preventing frequency oscillations. Section III describes state-space method, and Section IV presents the basis functions method. Numerical results are given in Section V to test the proposed approach. Finally, conclusions are presented in Section VI.

II. FNET System

As shown in Fig. 3, the FNET system consists of frequency disturbance recorders (FDRs) and the information management system (IMS). All the frequency measurements are collected at 120 V. Basically, the FDRs perform local GPS-synchronized measurements for voltage, phase angle, and amplitude together with frequencies. These measurements are accurately time-stamped and reported at 0.1-s intervals and then transmitted to a phaser data concentrator (PDC) for data processing and long-term storage. Then, the IMS handles data collection, storage, communication, database operations, and web service.
A variety of applications based on this data have been developed, examples are real-time event and oscillation detection, event location estimation, animated event visualization, and forensic authentication of digital evidence [24]. The FNET system has already shown promise in examining the dynamic characteristics of power grids during transient states and failures [25]. Though [24] has developed techniques to find characteristics of power grids during transient states and system has already shown promise in examining the dynamic forensic authentication of digital evidence [24]. The FNET event location estimation, animated event visualization, and developed, examples are real-time event and oscillation detection, eventually making it possible to isolate a node before failure affects the whole power grid.

III. STATE SPACE MODEL AND PARAMETERS ESTIMATION

In this section, we will first briefly discuss the state-space model and apply it to modeling power frequencies. EM and PEM algorithms are introduced to estimate the parameters of the model.

A. State Space Model

The state-space model has been widely employed in control systems and signal processing since it can be used online and updated after receiving new observations. It consists of a state (or system) equation (1) and a measurement (or output) equation (2). Denoted by \( y_k \) the power frequency measurement at time \( k \), state-space equation in discrete time can be written in the following form:

\[
\begin{align*}
    x_{k+1} &= Ax_k + B_kw_k & (1) \\
    y_k &= Cx_k + D_kv_k & (2)
\end{align*}
\]

where \( x_{k+1} \in \mathbb{R}^{n \times 1} \) (\( \mathbb{R}^{n \times 1} \) denotes the space of real vectors of dimension \( n \times 1 \)) is the state that characterizes the frequency; it is a variable of the time series \( \{x_k\} \) determined by the previous state \( x_k \) and the noise term \( w_k \in \mathbb{R}^{m \times 1} \) introduced at each \( k \); \( A_k \in \mathbb{R}^{n \times n} \) and \( B_k \in \mathbb{R}^{n \times m} \) are corresponding coefficients.

The time varying property of the parameters renders the state-space model able to adapt dynamically to a variety of frequencies. The noise terms \( w_k \) and \( v_k \) can capture small perturbations or uncertainties introduced at each \( k \) which improve the flexibility of the model. In this paper, the noise terms are assumed to be standard Gaussian variables, and the covariances of them are denoted by coefficients \( B_k \) and \( D_k \).

Due to these properties, we propose to use the state-space model to track and predict the behaviors of power frequency in the FNET. The unknown system parameters \( \dot{\theta}_k = \{A_k, B_k, C_k, D_k\} \) and states \( \{x_k\} \) can be estimated through a finite set of received signal measurement data \( Y = \{y_1, y_2, \ldots\} \). Future power frequencies can be predicted recursively based on these estimated parameters.

B. Kalman Filter and EM algorithm

This section describes the procedure employed to estimate model parameters and states associated with (1), using Kalman filtering [26] together with the EM algorithm [27][28]. It has been previously proved that the Kalman filter could optimally estimate the state in the mean square sense and that the filter-based EM algorithm yields a maximum likelihood (ML) parameter estimate, provided that \( w_k \) and \( v_k \) are assumed to be independent zero mean and unit variance Gaussian processes [29].

The Kalman filter can be described as follows [26]:

\[
\begin{align*}
    \hat{x}_{k/k} &= A_{k-1} \hat{x}_{k-1/k-1} + \frac{P_{k/k}^\top C_{k-1}^\top}{P_{k/k-1}} (y_k - C_{k-1} A_{k-1} \hat{x}_{k-1/k-1}) \\
    \hat{x}_{k/k} &= A_k \hat{x}_{k-1/k-1}
\end{align*}
\]

where \( k = 1, 2, \ldots, \) and \( P_{k/k} \) is given by

\[
\begin{align*}
    P_{k/k}^{-1} &= P_{k-1/k-1}^{-1} + A_{k-1}^\top B_{k-1}^\top P_{k-1/k-1} B_{k-1} B_{k-1}^\top A_{k-1} \\
    P_{k/k} &= C_{k-1}^\top D_{k-1}^{-2} C_{k-1} + B_{k-1}^2 - B_{k-1}^2 P_{k/k-1} B_{k-1}^2 \\
    \dot{P}_{k/k} &= A_k P_{k-1/k-1} A_{k-1}^\top + B_{k-1}^2
\end{align*}
\]

where \( B_{k}^2 \) and \( D_{k}^2 \) are assumed to be \( D_{k-1}^2 D_{k-1} \).

System parameters \( \dot{\theta}_k = \{A_k, B_k, C_k, D_k\} \) are estimated using the EM algorithm described by [27], [28]:

\[
\begin{align*}
    \dot{\theta}_k &= E \left( \sum_{i=1}^{k} x_i^\top y_i \left| Y_k \right. \right) \times \left[ E \left( \sum_{i=1}^{k} x_i^\top x_i \left| Y_k \right. \right) \right]^{-1} \\
    \dot{\theta}_k &= \frac{1}{k} E \left( \sum_{i=1}^{k} (x_i - A_i x_{i-1}) (x_i - A_i x_{i-1})^\top \right) | Y_k \rangle \\
    \dot{\theta}_k &= \frac{1}{k} E \left( \sum_{i=1}^{k} (x_i - A_i x_{i-1})^\top A_i^\top \right) | Y_k \rangle \\
    \dot{\theta}_k &= \frac{1}{k} E \left( \sum_{i=1}^{k} (y_i - C_i x_i) (y_i - C_i x_i)^\top \right) | Y_k \rangle \\
    \dot{\theta}_k &= \frac{1}{k} E \left( \sum_{i=1}^{k} (y_i - C_i x_i) (y_i - C_i x_i)^\top \right) C_i^\top | Y_k \rangle \\
    \dot{\theta}_k &= \frac{1}{k} E \left( \sum_{i=1}^{k} (y_i - C_i x_i) (y_i - C_i x_i)^\top \right) C_i^\top | Y_k \rangle
\end{align*}
\]

where \( E(x) \) denotes the expectation operator to variable \( x \), \( E(\cdot|Y_k) \) is the conditional expectation given the measurable set \( Y_k \).

All the variables denoted with "\( \cdot \)" refer to estimated values. Thus the filter-based EM algorithm for computing the MLE \( \dot{\theta}_k \) can be summarized as follows: Choose an initial parameter estimate \( \dot{\theta}_0 \), then compute \( \dot{\theta}_k \) at each iteration according to system (5) for the model.

The system parameters can be computed from the following conditional expectations [27]:

\[
\begin{align*}
    L_k^{(1)} &= E \left( \sum_{i=1}^{k} x_i^\top y_i \left| Y_k \right. \right) \\
    L_k^{(2)} &= E \left( \sum_{i=1}^{k} x_i^\top x_{i-1} \left| Y_k \right. \right) \\
    L_k^{(3)} &= E \left( \sum_{i=1}^{k} [x_i^\top x_{i-1} + x_i^\top x_{i-1}] \left| Y_k \right. \right) \\
    L_k^{(4)} &= E \left( \sum_{i=1}^{k} [y_i^\top y_i + y_i^\top x_i] \left| Y_k \right. \right)
\end{align*}
\]
The interested reader is referred to [30] for more detail on the choice of conditional expectations \( \{ l_k^{(1)}, l_k^{(2)}, l_k^{(3)}, l_k^{(4)} \} \) in (6).

After estimating all the parameters in the state space model, we can achieve a one-step prediction of the power frequency by [29]:
\[
\hat{y}_{k+1} = \hat{C}_k (\hat{A}_k \hat{x}_{k|k} + \hat{B}_k(y_k - \hat{C}_k \hat{x}_{k|k}))
\]  
(7)
where \( \hat{y}_{k+1} \) denotes the predicted power frequency at \( k+1 \) and \( \hat{K}_k \) is the Kalman gain.
\[
\hat{K}_k = (\hat{A}_{k-1} P_{k|k} \hat{C}_{k-1}) (\hat{C}_{k-1} P_{k|k} \hat{C}_{k-1} + \hat{D}_{k-1})^{-1}
\]  
(8)
This algorithm yields parameter estimates with non-decreasing values of the likelihood function, and converges under mild assumptions [31]. It should be noted that there is no exact requirement for the number of measurement data for EM method since it is recursive. However, in order to fully track the model, the system parameter \( B_k \) may grow rapidly, which may cause an unsatisfying prediction performance.

C. PEM Algorithm

In order to improve the prediction performance, we introduce another parameter estimation algorithm - PEM [32], whose objective is to minimize prediction errors. PEM updates the measurement set every time when the new measurement comes in, then, the model is updated with the new measurement set to keep up with time-varying parameters.

In this paper, we focus on one-step-ahead prediction. Given a finite number of measurements, PEM algorithm employs \( N \) stored measurements for the next prediction where the \( N \) can be chosen as the smallest number for which the algorithms have solutions. PEM estimates the parameters by minimizing a least square cost function:
\[
\min J_N = \min \frac{1}{N} \sum_{k=0}^{N-1} \| y_k - \hat{y}_k \|^2
\]  
(9)
where \( \hat{y}_k \) is one-step-ahead predictor at time \( k \) using the measurement data up to time \( k-1 \). The details of the algorithm on solving the above problem can be retrieved from [32].

With the estimated parameters and states of the model, the one-step prediction of the power frequency can be computed by [29]:
\[
\hat{y}_{k+1} = \hat{C}_k (\hat{A}_k \hat{x}_k + \hat{B}_k(y_k - \hat{C}_k \hat{x}_k))
\]  
(10)
where \( \hat{y}_{k+1} \) denotes the predicted power frequency at \( k+1 \); \( \hat{A}_k, \hat{B}_k, \hat{C}_k \) and \( \hat{x}_k \) are parameters and states estimated by PEM at time \( k \).

IV. MODELING AND PREDICTION WITH UNCERTAIN BASIS FUNCTIONS

In this section, we present the uncertain basis function technique employed to model and predict power frequencies.

A. Uncertain Basis Functions

Let \( y_k \) denote current measurements sampled by the device. It is assumed that the model for \( y_k \) can be written as follows:
\[
y_k = \sum_{i=1}^{P} A_i \phi_{i,k} \quad k = 1, 2, \ldots, N
\]  
(11)
where \( \{ \phi_{1,k}, \ldots, \phi_{p,k} \} \) are the basis functions, and \( p < N \), and \( A_i \) is the corresponding coefficient of each basis function. After obtaining a batch of measurements \( \{ y_1, \ldots, y_N \} \), the amplitudes can be estimated by using least squares and minimizing the expected cost function below if the basis functions are known:
\[
J(A) = \sum_{k=1}^{N} \sum_{i=1}^{p} (y_k - A_i \phi_{i,k})^2
\]  
(12)
where \( A = [A_1, \ldots, A_p]^T \) is the amplitude vector to be computed.

However, this assumption for known basis functions is not realistic since they are not given in practice. To tackle this difficulty, one possible solution is to assume that each basis function depends on some unknown parameter vectors \( \theta_i \), where only some statistics of the distribution of \( \theta_i \) are known. Then, the corresponding coefficients can be estimated by minimizing an expected cost function following the technique developed in [22].

In the following discussion, we will review the method developed in [22]. First, assume that each \( \theta_i \) is independent. The basis functions are further represented by \( \phi_{i,k}(\theta_i) \). The expected cost function \( \hat{J}(A) \) is given as:
\[
\hat{J}(A) = \mathbb{E}_\theta \left[ \sum_{k=1}^{N} \sum_{i=1}^{p} (y_k - A_i \phi_{i,k}(\theta_i))^2 \right]
\]  
(13)
where \( \mathbb{E}_\theta \) is the expectation with respect to \( \theta_i \). The measured values are real, so in order to estimate the coefficients \( A \), let \( y = [y_1, \ldots, y_N]^T \), \( \theta_i = [\theta_1, \ldots, \theta_p]^T \), and define an \( N \times p \) matrix \( H \) as below:
\[
H(\theta) = \begin{bmatrix}
\phi_{1,1}(\theta_1) & \phi_{1,2}(\theta_2) & \cdots & \phi_{1,p}(\theta_p) \\
\phi_{2,1}(\theta_1) & \phi_{2,2}(\theta_2) & \cdots & \phi_{2,p}(\theta_p) \\
\vdots & \vdots & \ddots & \vdots \\
\phi_{N,1}(\theta_1) & \phi_{N,2}(\theta_2) & \cdots & \phi_{N,p}(\theta_p)
\end{bmatrix}
\]  
(14)
In order to calculate \( A \), define
\[
G = \mathbb{E}_\theta [H(\theta)]
\]  
(15)
The matrix \( G \) is the expectation of \( H(\theta) \), and
\[
C_H = \mathbb{E}_\theta \left[ (H(\theta) - \mathbb{E}_\theta [H(\theta)])^T (H(\theta) - \mathbb{E}_\theta [H(\theta)]) \right]
\]  
(16)
where \( C_H \) is a type of covariance matrix.
Then, $A$ can be computed from the following equation [22]:

$$A = (G^T G + C_H)^{-1} G^T y$$  \hfill (18)

The term $C_H$ compensates for the uncertainty in the basis functions [22], and due to its independence, it can be written as a $p \times p$ diagonal matrix in (17) (at the bottom of the last page).

In our application, we wish to predict a future frequency using the average basis function:

$$\hat{y}_k = \sum_{i=1}^{p} \hat{A}_i E_{\theta}[\phi_{i,k}(\theta_i)]$$  \hfill (19)

where $\hat{A}_i$s are estimated by (18). This means that we can predict frequency values with $p$ random basis functions. In the next section, basis functions in the following applications are assumed to be governed by two different distributions, and we compute $G$ and $C_H$ for each case.

**B. Applications with Uniform and Gaussian Distributions**

In this section, we compute matrices $G$ and $C_H$ which will be used in the prediction. Let $s_i$ be a special case of $\theta_i$, the basis functions are chosen in the form of

$$\phi_{i,k}(\theta_i) = (k+1)^{s_i-1/N}$$  \hfill (20)

which is one of the most generally used bases [33]. It characterizes both time variations and different bases. Assume that $s_i$ is a random unknown variable that follows some distributions. If $s_i$’s are fixed, then the bases become traditional basis functions. This choice introduces time dependence $k$ as well as randomness.

1) **Uniform Distribution**: Assume that $s_i$ is uniformly distributed over the interval $[m_i/N - B/2, m_i/N + B/2]$, where $m_i$ is a known integer with $1 \leq m_i \leq N$, and $0 < B \leq 1/N$. Note that if we choose $B = 0$, then it reduces to the deterministic case. Under these statistics, we can calculate $G = E[H(\theta)]$ for uniform distribution as follows:

$$[G]_{kj} = \int_{m_i/N - B/2}^{m_i/N + B/2} (k+1)^{s_i-1/N} \frac{1}{B} ds_i$$

$$= \frac{(k+1)^{-1/N}}{B} \int_{m_i/N - B/2}^{m_i/N + B/2} e^{s_i \log(k+1)} ds_i$$  \hfill (21)

$$= \frac{(k+1)^{-1/N}}{B \cdot \log(k+1)} e^{m_i \log(k+1)/N} - e^{-m_i \log(k+1)/N}$$

where $k = 1, \ldots, N$ and $i = 1, \ldots, p$, and $[G]_{kj}$ represents the $k$th row and $i$th column of $G$. For $C_H$, we need to compute $var(\phi_{i,k}(\theta_i))$. Note from probability theory that

$$var(\phi_{i,k}(\theta_i)) = E\left( (k+1)^{2s_i-2/N} \right) - [G]_{kj}^2$$  \hfill (22)

After some calculations, we get

$$E\left( (k+1)^{2s_i-2/N} \right) = \int_{m_i/N - B/2}^{m_i/N + B/2} (k+1)^{2s_i-2/N} \frac{1}{B} ds_i$$

$$= \frac{(k+1)^{-2/N}}{2B \cdot \log(k+1)} e^{2m_i \log(k+1)/N} - e^{-2m_i \log(k+1)/N}$$  \hfill (23)

Then, with (23), (22) and (17), $C_H$ is completely determined.

2) **Gaussian Distribution**: In this section, $s_i$ is assumed to be a Gaussian random variable following the distribution $\mathcal{N}(m_i/N, \sigma_i^2)$. Then, similarly to the previous section, we have

$$[G]_{kj} = \int_{-\infty}^{+\infty} (k+1)^{s_i-1/N} \frac{1}{\sqrt{2\pi \sigma_i^2}} e^{-\frac{(s_i-m_i/N)^2}{2\sigma_i^2}} ds_i$$

$$= \frac{(k+1)^{-1/N}}{\sqrt{2\pi \sigma_i^2}} \int_{-\infty}^{+\infty} e^{s_i \log(k+1)} e^{-\frac{(s_i-m_i/N)^2}{2\sigma_i^2}} ds_i$$

$$= \frac{(k+1)^{-1/N}}{\sqrt{2\pi \sigma_i^2}} \int_{-\infty}^{+\infty} e^{-\frac{(s_i-m_i/N)^2+2m_i \log(k+1)/N}{2\sigma_i^2}} ds_i$$

$$= (k+1)^{-1/N} e^{\frac{1}{2}(\sigma_i^2 \log^2(k+1)+2m_i \log(k+1)/N)}$$  \hfill (24)

and to calculate $C_H$, we only need to compute the variable below as shown in (22)

$$E\left( (k+1)^{2s_i-2/N} \right) = (k+1)^{-2/N}$$

$$\times e^{2\sigma_i^2 \log^2(k+1)+m_i \log(k+1)/N}$$  \hfill (25)

In the above analysis, the mean value of $s_i$ for each distribution is chosen to be $m_i/N$. In the simulation below, we also show the performance with fixed basis functions where $s_i$ is set as $m_i/N$ in (20). The uncertain method can thus be explained as modeling frequency values not by some fixed bases, but rather by random bases with the coefficient having a mean $m_i/N$. This is an original aspect of the method.

**V. Numerical Results**

In this section, we present numerical results that highlight the performance of the proposed methods. According to a set of power frequency data for one specific power station chosen from the FNET, both the state-space model and uncertain basis functions model are applied. For the state-space model, any initial values can be chosen since all these values will converge. For the basis-functions model, we chose $p = 3$ to get a tradeoff between the computations while keeping the performance satisfactory. In addition, the number of measurement data points of the Basis Functions was chosen as the smallest number that can give us a solution. We tuned the parameters for Uniform and Gaussian distributions to have good prediction performance. Three different parameter sets are chosen as in Table I and Table II. Notice that, the mean values of both distributions should be smaller than 1 due to the exponential form of the Basis Functions. We consider a one-step prediction of the power frequency and compute the root mean square error (RMSE) for each method.

**A. State Space Model**

We first employ a test with the real sampled frequency data to demonstrate the prediction performance of the state-space model under EM and PEM algorithms, respectively. Fig. 4 depict a comparison between the measured and the predicted power frequency under both algorithms, respectively.
Fig. 5: The RMSE computed in this case for the same data basis functions is compared with the real measurements in the period. Once we can predict these kinds of event based with actual events to see if there exists a big jump during the period. Therefore, it is more meaningful to check the performance of the prediction with actual events to see if there exists a big jump during the period. Once we can predict these kinds of event based on the historic measurements, we will be able to help prevent and respond to the events in power grids more efficiently. An approach to predict the next frequency deviation in this case is 0.00202. The RMSE of EM/Kalman with 200 measurements is about 0.00270, while it is 0.00204 for PEM.

From Fig. 4, it can be seen that both the EM and PEM algorithms can dynamically track the stochastic behavior of future measurements.

For further comparison, the frequency RMSE is used as a performance measure and is defined as follows:

$$\text{RMSE}(k) = \sqrt{\frac{1}{N} \sum_{k=1}^{N} (y_k - \hat{y}_k)^2}$$  \hspace{1cm} (26)

The RMSE of EM/Kalman with 200 measurements is about 0.00270, while it is 0.00204 for PEM.

### B. Model with Uncertain Basis

Next, we perform another test with the same data set to demonstrate the prediction performance of the uncertain basis functions method. As mentioned before, we take three different parameter sets as examples to validate the methodology. The similar comparisons between the measured and the predicted parameter sets as examples to validate the methodology. The parameters from Parameter set 1 in both Table I and Table II are plugged into the following 1) and 2), respectively.

1) Uniform Distribution: Consider $s_i$ as a random variable in Uniform Distribution, where $s_1 \sim \mathcal{U}(0.10 - B/2, 0.10 + B/2)$, $s_2 \sim \mathcal{U}(0.66 - B/2, 0.66 + B/2)$, and $s_3 \sim \mathcal{U}(0.60 - B/2, 0.60 + B/2)$. Let $B = 1/N = 0.1$, predict the next frequency by (19), where

$$\mathbb{E}_\theta[\theta_i(N+1; \theta_i)] = \frac{(N+2)^{-1/N}}{B \cdot \log(N+2)} e^{\log(N+2)/2} \{e^{\log(N+2)/2} - e^{-\log(N+2)/2}\}$$  \hspace{1cm} (27)

The performance of the prediction of uniformly distributed basis functions is compared with the real measurements in Fig. 5. The RMSE computed in this case for the same data set is 0.00183.

2) Gaussian Distribution: If we assume that $s_i$'s are Gaussian random variables with $s_1 \sim \mathcal{N}(0.1, 0.001)$, $s_2 \sim \mathcal{N}(0.943, 0.001)$, and $s_3 \sim \mathcal{N}(0.925, 0.001)$. Similarly to the above, the prediction can be conducted by (19), where

$$\mathbb{E}_\theta[\theta_i(N+1; \theta_i)] = \frac{(N+2)^{-1/N}}{B \cdot \log(N+2) + 2m_i \log(N+2)/N} \times e^{\frac{1}{2} \sigma_i^2 \log^2(N+2)/2}$$  \hspace{1cm} (28)

The predicted values are plotted in Fig. 6. The RMSE computed in this case is 0.00202.

<table>
<thead>
<tr>
<th></th>
<th>Parameter set 1</th>
<th>Parameter set 2</th>
<th>Parameter set 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_1$</td>
<td>0.1</td>
<td>0.8</td>
<td>0.46</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>0.66</td>
<td>0.11</td>
<td>0.35</td>
</tr>
<tr>
<td>$\mu_3$</td>
<td>0.6</td>
<td>0.5</td>
<td>0.17</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.00183</td>
<td>0.00184</td>
<td>0.00214</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Parameter set 1</th>
<th>Parameter set 2</th>
<th>Parameter set 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_1$</td>
<td>0.1</td>
<td>0.931</td>
<td>0.431</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>0.943</td>
<td>0.144</td>
<td>0.874</td>
</tr>
<tr>
<td>$\mu_3$</td>
<td>0.925</td>
<td>0.425</td>
<td>0.225</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.001</td>
<td>0.002</td>
<td>0.001</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>$\sigma_3$</td>
<td>0.001</td>
<td>0.0015</td>
<td>0.001</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.00202</td>
<td>0.00178</td>
<td>0.00207</td>
</tr>
</tbody>
</table>

C. Performance Evaluation with event jump

As stated in [3], a frequency disturbance (event) has occurred in the system whenever the calculated result of frequency deviation exceeds a predefined limit. Obviously, it is more meaningful to check the performance of the prediction with actual events to see if there exists a big jump during the period. Once we can predict these kinds of event based on the historic measurements, we will be able to help prevent and respond to the events in power grids more efficiently. An
example is the prediction performance using uniform uncertain basis functions, shown in Fig. 7. We can see that both the measured data and predicted frequency keep oscillating and descending from 60.01 Hz to 59.985 Hz in 10 s, which indicates a fairly sharp event in the grid. If we enlarge the sampling interval as in the following section, then we can predict and prevent the event from propagating into the whole grid much earlier.

**D. Performance Evaluation with a Larger Sampling Interval**

In the previous discussions, the frequency was sampled every 0.1 second. In this part, we increased the sampling time to 1 second per frequency. Here, we examined the prediction performance of the proposed methods for a data set with a longer sampling period. Usually, a larger sampling period will indicate that the data used for modeling has a longer history, and a reasonable conclusion would be that the prediction accuracy has become degraded. We plotted the comparison for 200 samples using both the state-space model and the uncertain basis functions; the results are shown in Fig. 8~Fig. 10.

**E. Comparison analysis of all the methods**

To compare the performance, Table III shows the RMSEs obtained from the whole simulation. When the sampling interval is 0.1s, both the state-space method and basis function method yield a high accuracy. As PEM aims at minimizing the sum of the square of prediction errors, it provides better performance in the prediction compared with the EM/Kalman method. The uncertain basis method achieves an even better result than PEM.

In the case of using a larger sampling interval which is 1s, the RMSE of the PEM is 0.01019, while the RMSEs corresponding to Gaussian distributed uncertain basis functions and uniformly distributed uncertain basis functions are 0.00772 and 0.00752, respectively. The accuracy of PEM decreases significantly when the sampling interval is enlarged, while...
the performance of uncertain basis algorithm does not drop very much. The RMSE of the deterministic basis functions was also calculated and found to be 0.00862. The coefficients of deterministic basis functions were estimated using (18). Thus we found that the performance using the uncertain basis had improved around 10% compared with that using the deterministic basis.

VI. CONCLUSION

In this paper, we have considered the modeling and prediction of power frequencies using stochastic models. A state-space method with EM and Kalman filter, and the uncertain basis functions method were proposed for dynamically predicting the behavior of frequencies. In contrast with existing results, where basis functions are assumed to be known before model identification, we assumed them to be random variables obeying certain distribution functions. We validated the stochastic models using real measurement data and different parameter sets. All of the proposed techniques were able to achieve good prediction performance of frequency for the protection of power systems. Notice that, for simplicity and data availability issues, our attention was restricted to the frequency. The suggested state-space and uncertain basis functions methodologies for frequency forecasting, however, can also be modified to take into account other daily exogenous random variables. These include temperature, wind speed, sunshine, and market factors, which might affect the frequency of power grid. Future work will focus on predicting the frequency for multi-machine systems. An interesting problem is analyzing and predicting phasors in the power grid using the techniques proposed in this paper. It is also meaningful to try multi-step prediction in the power grid.

TABLE III: RMSEs of using different methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Short Sampling Interval</th>
<th>Long Sampling Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>EM</td>
<td>0.00270</td>
<td>0.01108</td>
</tr>
<tr>
<td>PEM</td>
<td>0.00204</td>
<td>0.01019</td>
</tr>
<tr>
<td>Deterministic</td>
<td>0.00216</td>
<td>0.00862</td>
</tr>
<tr>
<td>Uniform</td>
<td>0.00183</td>
<td>0.00752</td>
</tr>
<tr>
<td>Gaussian</td>
<td>0.00202</td>
<td>0.00772</td>
</tr>
</tbody>
</table>

REFERENCES


