

# A Two-stage Kalman Filtering Approach for Robust and Real-time Power Systems State Tracking

Jinghe Zhang, *Student Member, IEEE*, Greg Welch, *Member, IEEE*,  
Gary Bishop, and Zhenyu Huang *Senior Member, IEEE*

**Abstract**—As electricity demand continues to grow and renewable energy increases its penetration in the power grid, real-time state tracking becomes essential for system monitoring and control. Recent developments in phasor technology make real-time dynamic state estimation possible with high-speed time-synchronized data provided by synchronized Phasor Measurement Units (PMU).

In this paper we present a two-stage Kalman filtering approach to estimate the *static states* of voltage magnitudes and phase angles, as well as the *dynamic states* of generator rotor angles and generator speeds. Kalman filters achieve optimal performance only when the system noise characteristics have known statistical properties (zero-mean, Gaussian, and spectrally white). However in practice the process and measurement noise models are usually difficult to obtain. Thus in the first stage, we estimate the *static states* from raw PMU measurements, using a lightweight but efficient adaptive Kalman filtering algorithm called Adaptive Kalman Filter with Inflatable Noise Variances (AKF with *InNoVa*), which can identify and reduce the impact of incorrect system modeling and/or bad PMU measurements. In the next stage, the estimated bus voltages are fed into an extended Kalman filter to obtain the *dynamic state estimations*. Simulations demonstrate its robustness to sudden changes of system dynamics and erroneous PMU measurements.

**Index Terms**—Power systems, real-time state estimation, robust state estimation, adaptive Kalman filter, Phasor Measurement Units (PMU), bad data processing

## I. INTRODUCTION

A Quasi-steady-state assumption is typically applied to operational studies, for which the state estimation is at the core. Today's operation is based primarily on a model that largely ignores dynamics in the power grid—electromechanical interaction of generators and dynamic characteristics of loads and control devices are *not* included in operational models. This assumption reduces the computation by several orders of magnitude, enabling operation studies on standard computers within the required operational time intervals. The problem with this assumption is that many studies cannot be performed in the operational environment. The future grid is much less quasi-steady-state compared with the power grid in

the past. In particular, the widespread deployment of renewable generation, smart load controls, energy storage, and plug-in hybrid vehicles will require fundamental changes in the operational concepts and principal components of the grid. In part due to aggressive public policy goals, such as the U.S. state of California's push to generate 33% of its energy from renewable sources by 2020, this evolution will only accelerate, resulting in stochastic operating behaviors and dynamics the grid has never seen nor been designed for.

Nowadays, with the advance of synchrophasor technology, Phasor Measurement Units (PMUs) are becoming increasingly attractive in various time-critical power system applications such as system monitoring, protection, control, and stability assessment [1], [2]. They are able to provide real-time (typically 30 samples/second) synchrophasor data to capture the dynamic characteristics of the power system: in addition to estimating relatively stationary state elements such as bus voltage magnitudes and phase angles (we call these the *static states*), PMUs make it possible to estimate more transient states of a power system such as generator rotor angles and speeds (we call these the *dynamic states*).

Weighted least square (WLS) estimation is a classic method for estimating the *static states* [3]. In [4] the authors studied the use of phasor measurements in WLS state estimation, and how to identify bad data with normalized residual vectors. To deal with unpredictable process changes, [5] proposed a reverse prediction adaptive Kalman filtering algorithm, which adjusted the process noise parameter  $Q$  to improve filtering precision, assuming the measurement model was correct. As a comparison, [6] adjusted the measurement noise parameter  $R$  instead of  $Q$ , to increase the robustness erroneous measurements. An extended Kalman filter (EKF) based approach to estimate the *dynamic states* with dynamic power system models using PMU data was proposed in [7] and [8]. This approach can achieve satisfying and robust performance in terms of tracking the dynamic system states, with the assumption that all PMU measurements are ideal.

In this paper we propose a real-time two-stage Kalman filtering approach for simultaneously estimating the *static states* and the *dynamic states* in a power system. At every time step, stage one takes the raw PMU measurements into the Adaptive Kalman Filter with Inflatable Noise Variances (AKF with *InNoVa*), a novel AKF approach we recently developed to estimate the *static states*. It adjusts  $Q$  and  $R$  on-the-fly, allowing dynamic adaptation to unexpected process model changes and measurement errors. The results are passed on to stage two, which takes uses an Extended Kalman filter to estimate the truly *dynamic states*. Our method is straightforward to

This work supported by U.S. DOE grant DE-SC0002271 by Sandy Landsberg, Program Manager for Applied Mathematics Research; Office of Advanced Scientific Computing Research; DOE Office of Science. The work is also supported in part by U.S. ONR Award# N00014-12-1-0052 by Peter Squire, HPTE Deputy Thrust Manager.

Jinghe Zhang and Gary Bishop are in the Department of Computer Science, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, e-mail: {jing2009, gb}@cs.unc.edu. Greg Welch is in the Institute for Simulation & Training and the Computer Science Division of EECS at the University of Central Florida, FL, USA, e-mail: welch@ucf.edu. Zhenyu Huang is with the Pacific Northwest National Laboratory, Richland, WA, USA, e-mail: zhenyu.huang@pnl.gov.

implement, and shown to be effective and robust under various experimental scenarios.

The remainder of this article is organized as follows. Section II presents a review of traditional Kalman filtering techniques. Section III presents the principles and implementation details of our two-stage Kalman filtering approach. Section IV presents results that demonstrate robustness under different unusual but important conditions. The conclusions and acknowledgements are stated in Section V and VI.

## II. TRADITIONAL KALMAN FILTERS

In this section we give a brief description of the classic Kalman filters. Readers can refer to [9] for a more detailed introduction.

### A. The Kalman Filter (KF)

The Kalman filter (KF) has been used in a wide range of applications from economic analysis to radar tracking. It is an efficient recursive filter that estimates the state of a process in a way that minimizes the mean of the squared error when the process and measurement models are accurate and obey certain statistical and spectral properties.

An assumed *linear* system can be modeled as a pair of linear stochastic process and measurement equations

$$x_k = Ax_{k-1} + w_{k-1} \quad (1)$$

$$z_k = Hx_k + v_k \quad (2)$$

where  $x \in \mathcal{R}^n$  is the state vector,  $z \in \mathcal{R}^m$  is the measurement vector,  $A$  is a  $n \times n$  matrix that relates the state at the previous time step  $k-1$  to the state at the current step  $k$  in the absence of either a driving function or process noise<sup>1</sup>, and  $H$  is a  $m \times n$  matrix that relates the state to the measurement  $z_k$ . The process noise  $w_k$  and measurement noise  $v_k$  are assumed to be mutually independent random variables, spectrally white, and with normal probability distributions

$$p(w) \sim N(0, Q) \quad (3)$$

$$p(v) \sim N(0, R), \quad (4)$$

where the process noise covariance  $Q$  and measurement noise covariance  $R$  matrices are often assumed to be constant.

We define  $\hat{x}_k^- \in \mathcal{R}^n$  to be the *a priori* state estimate at time step  $k$  given the knowledge of the process prior to  $k$ , so  $e_k^- \equiv x_k - \hat{x}_k^-$  is called the *a priori* estimate error and  $P_k^- \equiv E[e_k^- e_k^{-T}]$  is called the *a priori* estimate error covariance.

Similarly, We define  $\hat{x}_k \in \mathcal{R}^n$  to be the *a posteriori* state estimate at time step  $k$  given measurement  $z_k$ , so  $e_k \equiv x_k - \hat{x}_k$  and  $P_k \equiv E[e_k e_k^T]$  are called the *a posteriori* estimate error and the *a posteriori* estimate error covariance respectively.

The Kalman filter estimates the state by minimizing the *a posteriori* estimate error covariance, in a recursive prediction-correction manner. The prediction step is realized by a set of time update equations:

$$\text{Predict} \begin{cases} \hat{x}_k^- = A\hat{x}_{k-1} \\ P_k^- = AP_{k-1}A^T + Q \end{cases} \quad (5)$$

The time update equations are responsible for projecting forward (in time) the previous state  $x_{k-1}$  and error covariance estimates  $P_{k-1}$  to obtain the *a priori* estimates for the next time step  $k$ .

The correction step is carried out by a set of measurement update equations:

$$\text{Correct} \begin{cases} K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \\ \hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \\ P_k = (I - K_k H)P_k^- \end{cases} \quad (6)$$

where  $K$  is a  $n \times m$  matrix called the Kalman gain matrix,  $z_k$  is the actual measurement at time step  $k$ ,  $H\hat{x}_k^-$  is the predicted measurement at time step  $k$ . Kalman gain  $K$  reflects how we trust the actual measurement  $z_k$  versus the predicted measurement  $H\hat{x}_k^-$ . From its expression, one can tell that larger values of  $R$  place more weight on the predicted values while smaller values of  $R$  place more weight on the measured values. The measurement update equations are instrumental in the feedback, i.e. for incorporating a new measurement into the *a priori* estimate to obtain an improved *a posteriori* estimate.

### B. The Extended Kalman Filter (EKF)

In reality, the process to be estimated and (or) the measurement relationship to the process are often *nonlinear*. This is certainly true when estimating the dynamic states of a power system. A nonlinear system can be modeled using nonlinear stochastic process and measurement equations corresponding to the linear equations (1) and (2):

$$x_k = a(x_{k-1}, w_{k-1}) \quad (7)$$

$$z_k = h(x_k, v_k). \quad (8)$$

One can approximate the states and measurements by

$$\tilde{x}_k = a(\hat{x}_{k-1}, 0) \quad (9)$$

$$\tilde{z}_k = h(\tilde{x}_k, 0). \quad (10)$$

These nonlinear functions can then be linearized about the point of interest  $x$  in the state space. To do so one need to compute the Jacobian matrices

$$A = \left. \frac{\partial a(x)}{\partial x} \right|_x, W = \left. \frac{\partial a(x)}{\partial w} \right|_x, H = \left. \frac{\partial h(x)}{\partial x} \right|_x, V = \left. \frac{\partial h(x)}{\partial v} \right|_x \quad (11)$$

where  $A$  and  $W$  are the partial derivatives of  $a$  with respect to  $x$  and  $w$ ,  $H$  and  $V$  are the partial derivatives of  $h$  with respect to  $x$  and  $v$ , respectively, at each time step.

An Extended Kalman filter (EKF) is essentially a Kalman filter modified to linearize the estimation about the current mean and covariance. Similar to the KF equations, the EKF can be expressed as follows:

$$\text{Predict} \begin{cases} \hat{x}_k^- = a(\hat{x}_{k-1}, 0) \\ P_k^- = A_k P_{k-1} A_k^T + W_k Q_{k-1} W_k^T \end{cases} \quad (12)$$

$$\text{Correct} \begin{cases} K_k = P_k^- H_k^T (H_k P_k^- H_k^T + V_k R_k V_k^T)^{-1} \\ \hat{x}_k = \hat{x}_k^- + K_k(z_k - h(\hat{x}_k^-, 0)) \\ P_k = (I - K_k H_k)P_k^- \end{cases} \quad (13)$$

<sup>1</sup>In practice, the matrix  $A$  may change with each time step, but it is assumed to be constant here.

### III. TWO-STAGE KALMAN FILTERING APPROACH

#### A. Stage One

The inputs to stage one are the raw measurement data collected from PMUs installed in the system. The outputs are the relatively static states: bus voltage magnitudes and phase angles, which can be estimated using the available measurements and system models. The same system model from our previous research [10] is used in this particular power system state estimation problem. Per [4], we use a rectangular coordinate formulation where the real and imaginary parts of bus voltages are considered state variables, to avoid numerical startup problems associated with current phasors.

For the process model, the power system is assumed to be reasonably stable, hence a quasi-static model of a power system has been employed for the base process model. Note that this assumption could be incorrect (sub-optimal) but our models dynamically adapt as indicated below.

For the measurement model, we assume all measurements are provided by PMUs. In addition to higher precision and higher rate measurements, PMUs have the advantage that they measure bus voltage phasors and line current phasors, so the relationship between measurements and states is linear. This is in contrast to conventional power system state estimators which use real and reactive power as measurements, so the measurement-state relationship is non-linear. We also assume that a PMU installed at a specific bus can measuring not only the bus voltage phasor, but also the current phasors along all the lines incident to the bus. A more detailed measurement modeling can be found in [10].

Ideally, the chosen base models are the “true” models, with accurately modeled noise characteristics. However when the chosen models do not match the actual system behavior, and/or PMU measurements contain significant errors, the state estimates could deviate from the true states rapidly. Thus we developed *AKF with InNoVa* in the first stage to deal with unknown system dynamics and erroneous PMU measurements.

With our adaptive Kalman filter we treat all un-modeled PMU measurement errors as noise. By definition, these errors are unknown and unpredictable, so they cannot be reflected in the measurement noise covariance  $R$ . Similarly, we treat deviations from the true process model as un-modeled noise that is not reflected in the process noise covariance  $Q$ . Multiple/switching process models are not considered for now.

Throughout we refer to the conventional Kalman filter innovation  $(z_k - H\hat{x}_k^-)$  as the *a priori innovation*, to distinguish it from the *a posteriori innovation*  $(z_k - H\hat{x}_k)$ . Traditionally, to assess the performance of a filter people examine the *a priori innovation*  $\mathcal{I}_k^- = (z_k - H\hat{x}_k^-)$ , which should have a normal distribution with zero mean and covariance  $S_k = HP_k^-H^T + R$  [11]. When the implemented process model does not match reality, the mean of the innovation can shift, and the magnitude grow, such that eventually the normalized *a priori innovation* (normalized by its covariance  $S_k$ ) exceeds a predetermined threshold. However, it is usually impossible to determine whether the shift/growth is caused by a process model mismatch, a measurement model mismatch, or both. Note that  $Q$  and  $R$  are already blended (indistinguishable)

in  $S_k$ . This is why we want to investigate the *a posteriori innovation*  $\mathcal{I}_k = (z_k - H\hat{x}_k)$  as well: its ideal covariance  $T_k$  can be used to identify un-modeled PMU measurement errors.

*Lemma:* Ideally, the *a posteriori innovation*  $\mathcal{I}_k = (z_k - H\hat{x}_k)$  should be normally distributed with zero mean and covariance  $RS_k^{-1}R$ .

*Proof:* According to the *a posteriori* state estimate in (6) by incorporating the measurement, we have

$$\begin{aligned}\hat{x}_k &= \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \\ H\hat{x}_k &= H\hat{x}_k^- + HK_k(z_k - H\hat{x}_k^-) \\ z_k - H\hat{x}_k &= (z_k - H\hat{x}_k^-) - HK_k(z_k - H\hat{x}_k^-) \\ z_k - H\hat{x}_k &= (I - HK_k)(z_k - H\hat{x}_k^-),\end{aligned}\quad (14)$$

where  $I$  is the identity matrix. Then the mean of the *a posteriori innovation* is

$$E(z_k - H\hat{x}_k) = (I - HK_k)E(z_k - H\hat{x}_k^-) = 0, \quad (15)$$

and the covariance of the *a posteriori innovation* is

$$\begin{aligned}\text{cov}(z_k - H\hat{x}_k) &= (I - HK_k)\text{cov}(z_k - H\hat{x}_k^-)(I - HK_k)^T \\ &= (I - HK_k)S_k(I - HK_k)^T.\end{aligned}\quad (16)$$

Next we will show that  $I - HK_k = RS_k^{-1}$ :

$$\begin{aligned}I - HK_k &= I - HP_k^-H^T(HP_k^-H^T + R)^{-1} \\ &= R(HP_k^-H^T + R)^{-1} \\ &= RS_k^{-1}.\end{aligned}\quad (17)$$

Combining (16) and (17) we can now write

$$\begin{aligned}\text{cov}(z_k - H\hat{x}_k) &= RS_k^{-1}S_k(RS_k^{-1})^T \\ &= R(S_k^{-1})^T R^T \\ &= RS_k^{-1}R,\end{aligned}\quad (18)$$

because the *a priori innovation* covariance  $S_k$  and measurement noise covariance  $R$  are both symmetric matrices. We define  $T_k = RS_k^{-1}R$  as the *a posteriori innovation* covariance. ■

The basic idea of our approach is as follows. To simplify the notation we omit the time step count  $k$ . Within each filtering cycle, after the prediction step, we compute the ideal *a priori innovation* covariance  $S$  and the normalized *a priori innovation* vector  $\tilde{\mathcal{I}}^-$  where

$$\tilde{\mathcal{I}}_i^- = |\mathcal{I}_i^-|/\sqrt{S_{ii}}. \quad (19)$$

If  $\tilde{\mathcal{I}}_i^- > \tau$  for some threshold  $\tau$ , then  $i \in \text{Out}$ , where *Out* holds the outlier indices. In our experiments, we used  $\tau = 3$  (measurement units). First we assume the outliers are caused by unknown process noise, so we want to “inflate”  $Q$  by a diagonal matrix  $\Delta Q$ , such that  $P^-$  is also inflated by  $\Delta Q$ . Thus  $S$  is consequently inflated by  $\Delta S = H(\Delta Q)H^T$  and

$$\tilde{\mathcal{I}}_i^- = |\mathcal{I}_i^-|/\sqrt{S_{ii} + \Delta S_{ii}} \leq \tau \quad i = 1, 2, \dots, n, \quad (20)$$

where

$$\Delta S_{ii} = \sum_{j=1}^n H_{ij}^2 \Delta Q_j = (H(i, :) \cdot H(i, :))([\Delta Q_1, \dots, \Delta Q_n]^T), \quad (21)$$

and “.” denotes dot product. We use a linear programming approach to solve the optimization problem:

$$\begin{aligned} & \min \sum_{i=1}^n \Delta Q_i \quad (22) \\ \text{s.t. } & \Delta S_{ii} = (H(i, :) \cdot H(i, :))([\Delta Q_1, \dots, \Delta Q_n]^T) \\ & \geq (|\mathcal{I}_i^-|/\tau)^2 - S_{ii}, \quad \forall i \in \text{Out} \\ & \Delta Q_1 \geq 0, \Delta Q_2 \geq 0, \dots, \Delta Q_n \geq 0. \end{aligned}$$

The inflated  $Q$  is then incorporated in the correction step. Similarly, we compute the ideal *a posteriori* innovation covariance  $T$  and the normalized *a posteriori* innovation vector  $\tilde{\mathcal{I}}$  where

$$\tilde{\mathcal{I}}_i = |\mathcal{I}_i|/\sqrt{T_{ii}}. \quad (23)$$

If  $\tilde{\mathcal{I}}_i > \tau$ , then  $i \in \text{MeasOut}$  where  $\text{MeasOut}$  holds the measurement outlier indices, indicating abnormal measurements.

Now we can separate the measurement “noise” elements from the process “noise” elements. Let us denote  $\text{ProcOut} = \text{Out} \setminus \text{MeasOut} = \{i : i \in \text{Out} \text{ and } i \notin \text{MeasOut}\}$  as the set difference between  $\text{Out}$  and  $\text{MeasOut}$ . If  $\text{MeasOut}$  is not empty, we will recalculate the  $\Delta Q$  as in optimization problem (22), except only for  $\forall i \in \text{ProcOut}$ , and update  $Q$  to  $Q + \Delta Q$ . As for the measurements, we “inflate”  $R$  in this way: for  $\forall i \in \text{MeasOut}$ ,  $R_{ii}$  is inflated to  $\lambda_i R_{ii}$ . As a result,  $T = RS^{-1}R$  is also inflated such that the  $i$ th diagonal element is now  $\lambda_i^2 T_{ii}$  and

$$\tilde{\mathcal{I}}_i = |\mathcal{I}_i|/\lambda_i \sqrt{T_{ii}} \leq \tau, \quad i = 1, 2, \dots, n. \quad (24)$$

It is relatively straightforward to compute the  $\lambda_i$  values by

$$\lambda_i = (|\mathcal{I}_i|/\sqrt{T_{ii}})/\tau, \quad i \in \text{MeasOut}. \quad (25)$$

Finally, with the inflated  $Q$  and  $R$ , we recompute the correction step (6) to obtain a more robust state estimation. The  $Q$  and  $R$  are also updated and carried on to the next cycle.

Furthermore,  $Q$  and  $R$  should also be deflatable if the abnormal process/measurement problems are only temporary and eventually resolved. Our solution is to employ an exponential decay process to enable automatic deflation of the parameters over each cycle. The decay time constant can be customized by users, according to their expectation and the specific circumstances.

## B. Stage Two

In stage two, the estimation results from stage one are fed directly into an Extended Kalman Filter (EKF) as “measurements.” The output are the dynamic states: generator rotor angles and generator speed. As mentioned before, measurements can be expressed in terms of the state variables either using the rectangular or the polar coordinates. In this method, our measurements are in their rectangular forms, *i.e.* the real and imaginary parts of all bus voltages, so the process and measurement models need to be modified as follows.

## C. The Process Model

Without loss of generality, in a power system that consists of  $n$  generators, let us consider the generator  $i$  which is connected to the generator terminal bus  $i$ . We use a classical model for the generator composed of a voltage source  $|E_i| \angle \delta_i$  with constant amplitude behind an impedance  $X'_{d_i}$ . The nonlinear differential-algebraic equations regarding the generator  $i$  can be written as

$$\begin{cases} \frac{d\delta_i}{dt} = \omega_B(\omega_i - \omega_0) \\ \frac{d\omega_i}{dt} = \frac{\omega_0}{2H_i} (P_{m_i} - \frac{|E_i|}{X'_{d_i}} \sin \delta_i |V_i| \cos \theta_i \\ \quad + \frac{|E_i|}{X'_{d_i}} \cos \delta_i |V_i| \sin \theta_i - D_i(\omega_i - \omega_0)) \\ = \frac{\omega_0}{2H_i} (P_{m_i} - \frac{|E_i|}{X'_{d_i}} \sin \delta_i \text{Re}(V_i) \\ \quad + \frac{|E_i|}{X'_{d_i}} \cos \delta_i \text{Im}(V_i) - D_i(\omega_i - \omega_0)) \end{cases} \quad (26)$$

where state variables  $\delta_i$  and  $\omega_i$  are the generator rotor angle and speed respectively,  $\omega_B$  and  $\omega_0$  are the speed base and the synchronous speed in per-unit (pu) quantities,  $P_{m_i}$  is the mechanical input,  $H_i$  is the machine inertia<sup>2</sup>,  $D_i$  is the generator damping coefficient and  $V_i = |V_i| \angle \theta_i$  is the phasor voltage at the generator terminal bus  $i$  (which is a function of  $\delta_1, \delta_2, \dots, \delta_n$ ).

For the state vector  $x = [\delta_1, \omega_1, \delta_2, \omega_2, \dots, \delta_n, \omega_n]^T$ , the corresponding *continuous time* change in state can be modeled by the linearized equation

$$\frac{dx}{dt} = A_c x + w_c, \quad (27)$$

where  $w_c$  is an  $2n \times 1$  continuous time process noise vector with  $2n \times 2n$  noise covariance matrix  $Q_c = E[w_c w_c^T]$ , and  $A_c$  is an  $2n \times 2n$  continuous time state transition Jacobian matrix, whose entries for  $i \in \{1, \dots, n\}$  and  $j \in \{1, \dots, n\}$  ( $i \neq j$ ) are the corresponding partial derivatives

$$A_{c[2i-1, 2i-1]} = 0 \quad (28)$$

$$A_{c[2i-1, 2i]} = \omega_B \quad (29)$$

$$\begin{aligned} A_{c[2i, 2i-1]} = & -\frac{\omega_0 |E_i|}{2H_i X'_{d_i}} \left[ \cos \delta_i \text{Re}(V_i) + \sin \delta_i \frac{\partial \text{Re}(V_i)}{\partial \delta_i} \right. \\ & \left. + \sin \delta_i \text{Im}(V_i) - \cos \delta_i \frac{\partial \text{Im}(V_i)}{\partial \delta_i} \right] \end{aligned} \quad (30)$$

$$A_{c[2i, 2i]} = -\frac{\omega_0}{2H_i} D_i \quad (31)$$

$$A_{c[2i-1, 2j-1]} = 0 \quad (32)$$

$$A_{c[2i-1, 2j]} = 0 \quad (33)$$

$$A_{c[2i, 2j-1]} = -\frac{\omega_0 |E_i|}{2H_i X'_{d_i}} \left[ \sin \delta_i \frac{\partial \text{Re}(V_i)}{\partial \delta_j} - \cos \delta_i \frac{\partial \text{Im}(V_i)}{\partial \delta_j} \right] \quad (34)$$

$$A_{c[2i, 2j]} = 0. \quad (35)$$

<sup>2</sup>The mechanical power  $P_{m_i}$  and machine inertia  $H_i$  should not be confused with the error covariance  $P$  and measurement Jacobian  $H$  from the preceding section. While potentially confusing these are the variables used by popular convention in the respective fields.

Hence the update of the state vector  $x$  from time step  $(k-1)$  to  $k$  over duration  $\Delta t$  has the complete corresponding *discrete-time* state transition matrix

$$A = I + A_c \cdot \Delta t. \quad (36)$$

The discrete-time process noise covariance  $Q$  can be formulated by integrating the continuous time process equation (27) over the time interval  $\Delta t$  as described before:

$$Q = \int_0^{\Delta t} e^{A_c t} Q_c e^{A_c^T t} dt \quad (37)$$

Hence the process model is written as

$$x_k = Ax_{k-1} + w_{k-1} \quad (38)$$

$$= (I + A_c \cdot \Delta t)x_{k-1} + w_{k-1}, \quad (39)$$

where  $w$  is the process noise with normal probability distribution  $p(w) \sim N(0, Q)$ .

#### D. The Measurement Model

The expanded system nodal equation can be expressed as:

$$Y_{exp} \begin{pmatrix} E \\ V \end{pmatrix} = \begin{pmatrix} Y_{GG} & Y_{GL} \\ Y_{LG} & Y_{LL} \end{pmatrix} \begin{pmatrix} E \\ V \end{pmatrix} = \begin{pmatrix} I_G \\ 0 \end{pmatrix}, \quad (40)$$

where  $E$  is the vector of internal generator complex voltages,  $V$  is the vector of bus complex voltages,  $I_G$  represents electrical currents injected by generators,  $Y_{exp}$  is called the expanded nodal matrix, which includes loads and generator internal impedances:  $Y_{GG}$ ,  $Y_{GL}$ ,  $Y_{LG}$  and  $Y_{LL}$  are the corresponding partitions of the expanded admittance matrix.

According to the expanded system nodal equation, all node voltages phasors can be expressed in terms of the internal generator voltages and angles by using the bus voltage reconstruction matrix  $R_V$ :

$$V = -Y_{LL}^{-1} Y_{LG} E = R_V E \quad (41)$$

Thus in a system with  $n$  generators and  $m$  buses, we have

$$\begin{pmatrix} V_1 \\ V_2 \\ \vdots \\ V_m \end{pmatrix} = \begin{pmatrix} R_{V_{11}} & R_{V_{12}} & \cdots & R_{V_{1n}} \\ R_{V_{21}} & R_{V_{22}} & \cdots & R_{V_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ R_{V_{m1}} & R_{V_{m2}} & \cdots & R_{V_{mn}} \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_n \end{pmatrix}, \quad (42)$$

where  $V_i = |V_i| \angle \theta_i$  is the phasor voltage at bus  $i$ , and  $E_j = |E_j| \angle \delta_j$  is the voltage source at generator  $j$ .

In rectangular form, we have

$$\begin{aligned} \text{Re}(V_i) &= |R_{V_{i1}}| |E_1| \cos(\angle R_{V_{i1}} + \delta_1) \\ &+ |R_{V_{i2}}| |E_2| \cos(\angle R_{V_{i2}} + \delta_2) \\ &+ \cdots + |R_{V_{in}}| |E_n| \cos(\angle R_{V_{in}} + \delta_n) \end{aligned} \quad (43)$$

$$\begin{aligned} \text{Im}(V_i) &= |R_{V_{i1}}| |E_1| \sin(\angle R_{V_{i1}} + \delta_1) \\ &+ |R_{V_{i2}}| |E_2| \sin(\angle R_{V_{i2}} + \delta_2) \\ &+ \cdots + |R_{V_{in}}| |E_n| \sin(\angle R_{V_{in}} + \delta_n) \end{aligned} \quad (44)$$

The ‘‘measurements’’ obtained from the stage-one AKF can now be written as

$$z = h(x) + v, \quad (45)$$

where  $z = [\text{Re}(V_1), \text{Im}(V_1), \dots, \text{Re}(V_m), \text{Im}(V_m)]^T$  is the  $2m \times 1$  measurement vector that consists of real and imaginary parts of each estimated bus voltage,  $x = [\delta_1, \omega_1, \delta_2, \omega_2, \dots, \delta_n, \omega_n]^T$  is the state vector and  $v$  is the normally distributed measurement noise with covariance  $P$  estimated by the stage-one AKF.

After linearization, the measurement model is

$$z = Hx + v, \quad (46)$$

where  $H$  is the corresponding Jacobian matrix defined as

$$H = \begin{pmatrix} \frac{\partial \text{Re}(V_1)}{\partial \delta_1} & \frac{\partial \text{Re}(V_1)}{\partial \omega_1} & \cdots & \frac{\partial \text{Re}(V_1)}{\partial \delta_n} & \frac{\partial \text{Re}(V_1)}{\partial \omega_n} \\ \frac{\partial \text{Im}(V_1)}{\partial \delta_1} & \frac{\partial \text{Im}(V_1)}{\partial \omega_1} & \cdots & \frac{\partial \text{Im}(V_1)}{\partial \delta_n} & \frac{\partial \text{Im}(V_1)}{\partial \omega_n} \\ \frac{\partial \text{Re}(V_2)}{\partial \delta_1} & \frac{\partial \text{Re}(V_2)}{\partial \omega_1} & \cdots & \frac{\partial \text{Re}(V_2)}{\partial \delta_n} & \frac{\partial \text{Re}(V_2)}{\partial \omega_n} \\ \frac{\partial \text{Im}(V_2)}{\partial \delta_1} & \frac{\partial \text{Im}(V_2)}{\partial \omega_1} & \cdots & \frac{\partial \text{Im}(V_2)}{\partial \delta_n} & \frac{\partial \text{Im}(V_2)}{\partial \omega_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial \text{Re}(V_m)}{\partial \delta_1} & \frac{\partial \text{Re}(V_m)}{\partial \omega_1} & \cdots & \frac{\partial \text{Re}(V_m)}{\partial \delta_n} & \frac{\partial \text{Re}(V_m)}{\partial \omega_n} \\ \frac{\partial \text{Im}(V_m)}{\partial \delta_1} & \frac{\partial \text{Im}(V_m)}{\partial \omega_1} & \cdots & \frac{\partial \text{Im}(V_m)}{\partial \delta_n} & \frac{\partial \text{Im}(V_m)}{\partial \omega_n} \end{pmatrix} \quad (47)$$

More specifically, for  $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, n$ , we have

$$\frac{\partial \text{Re}(V_m)}{\partial \delta_n} = -|R_{V_{mn}}| |E_n| \sin(\angle R_{V_{mn}} + \delta_n) \quad (48)$$

$$\frac{\partial \text{Im}(V_m)}{\partial \delta_n} = |R_{V_{mn}}| |E_n| \cos(\angle R_{V_{mn}} + \delta_n) \quad (49)$$

$$\frac{\partial \text{Re}(V_m)}{\partial \omega_n} = \frac{\partial \text{Im}(V_m)}{\partial \omega_n} = 0 \quad (50)$$

#### IV. CASE STUDY

Our two-stage Kalman filtering approach has been tested on several multi-machine systems under various conditions. The ideal testing cases are omitted in this paper, because we are more interested in the filtering performances under adverse circumstances.

We will show the testing results for a 16-generator-68-bus system (Fig. 1), representing New England Test System and New York Power System (NETS-NYPS). Five seconds are simulated in steps of 0.01 seconds. At  $t = 1.1$  seconds a three-phase fault at bus 29 occurs, and is then cleared at 0.05 seconds. This event represents a large disturbance (an emergency) causing voltage oscillations and rotor speed changes. PMU data is assumed to contain 1% random noise. In the following four subsections, four abnormal cases are simulated to illustrate the robustness of our proposed algorithm. In each case we executed three two-stage Kalman filtering (KF) scenarios for a more informative comparison: Scenario 1 uses traditional KF at stage one, while Scenario 2 uses naive robust KF (RKF) (using largest normalized innovation test to identify and exclude bad measurements, without adjusting model parameters) and Scenario 3 uses AKF with InNoVa. Without loss of generality, we visualize the voltage tracking results for bus 22 in 3D. At stage two, the outputs of these three filters are processed by separate EKFs to estimate the system’s dynamic states. We chose generator 58 to demonstrate and compare the dynamic states tracking results.

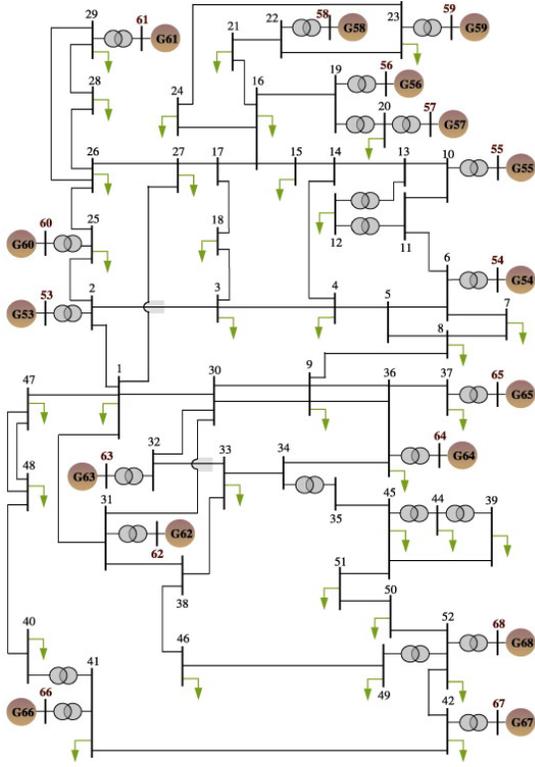


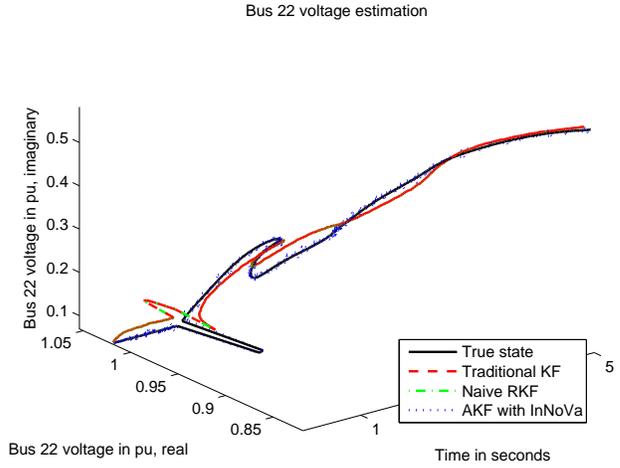
Fig. 1. The 16-generator-68-bus system

### A. Case 1: small $Q$ and $R$ initialization

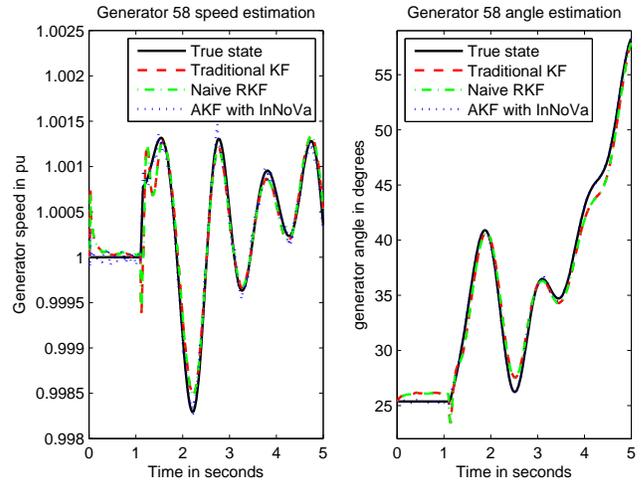
Here we initialize each process element with a very small variance ( $2.5e^{-5}$ ) as if the user was very confident the system was stable. We then initialize the measurement covariance 1%, but simulate an actual PMU measurement noise of 2%. Fig. 2 shows the two-stage tracking results of the three scenarios. In this and the following simulations, we always use a black solid line for the true state, a red dash-dash line for Scenario 1, a green dash-dot line for Scenario 2, and a blue dot-dot line for Scenario 3—our proposed approach. From Fig. 2(a) we can tell that the KF and naive RKF are both strongly affected by the small  $Q$ —the estimated voltages have slow and small oscillations. Yet the AKF with InNoVa closely tracks the true state. We have observed that bus 29 voltages have the largest process variance (0.1180) which is no surprise because the fault is simulated at bus 29. The average measurement noise has grown to 1.4% in 5 seconds. Fig. 2(b) illustrates the corresponding dynamic state estimations. It can be seen the Scenario 3 outperforms the others.

### B. Case 2: small $Q$ with a malfunctioning device

Here we initialize  $Q$  as above, and this time set  $R$  correctly. We then introduce a problem: a malfunctioning PMU at bus 22. The simulated voltage measurements provided by this PMU contain systematic errors, with a normal distribution  $\mathcal{N}(1, 0.1^2)$ . Fig. 3 illustrates the results. Fig. 3(a) shows that the naive RKF has identified and dropped the bad measurements. However it is still impacted by the small  $Q$  initialization. The AKF with InNoVa not only adjusted  $Q$  (bus 29 still has the largest variance), but also identified the bad measurement. Although our algorithm does not exclude the



(a) Performance comparison of the three scenario after stage one in case 1: small  $Q$  and  $R$  initialization



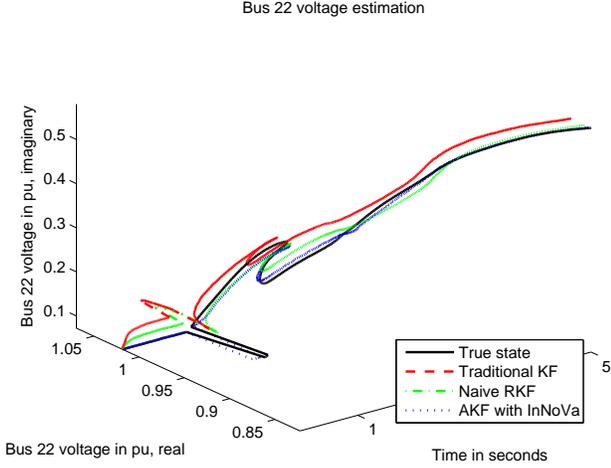
(b) Performance comparison of the three scenarios after stage two in case 1

Fig. 2. Performance comparison of the three scenario in case 1

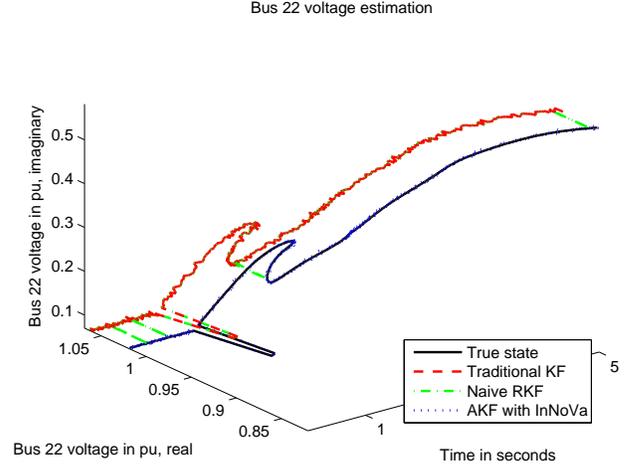
measurement—we seek to avoid unobservable conditions by maintaining sufficient redundancy—the noise of this erroneous measurement becomes 41.21%, which is significantly larger than the others (1%). Hence this measurement is not weighted as heavily, and has a negligible impact on the estimated states. Fig. 3(b) also shows that while Scenario 2 does improve, Scenario 3 stays the closest to the truth.

### C. Case 3: large $Q$ with a malfunctioning device

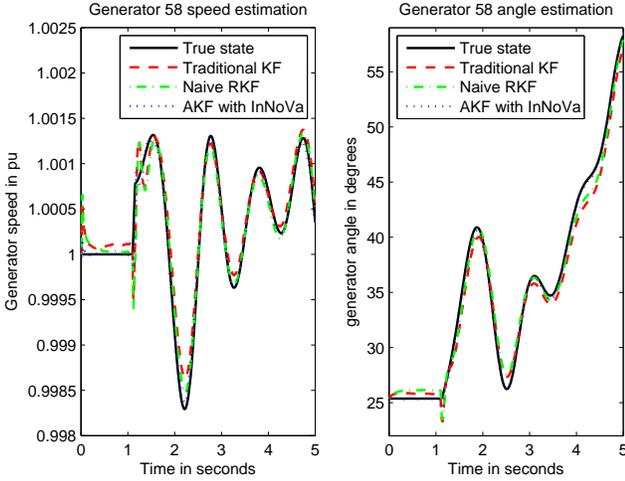
Now we increase  $Q$  significantly: the process variance of each state element is initialized to 1. This time the voltage measurement errors at bus 22 are even larger, with distribution  $\mathcal{N}(2.8, 0.1^2)$ . The results are depicted in Fig. 4. Interestingly in Fig. 4(a), the naive RKF rarely detects the systematic error. Because  $Q$  is set so large (hence  $S$  is so large) the normalized *a priori innovations* almost always fall under the threshold. However the error is always detected by the *a posteriori innovations* testing in AKF with InNoVa. Thus, as shown in Fig. 4(a) and 4(b), Scenario 3 still performs the best.



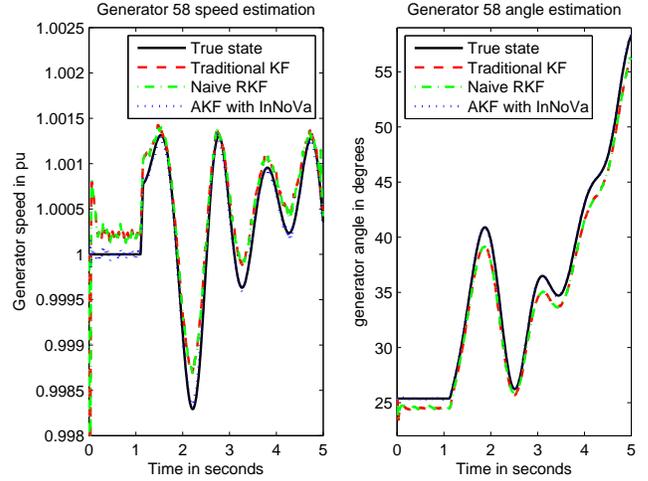
(a) Performance comparison of the three scenarios after stage one in case 2: small  $Q$  initialization with a malfunctioning device



(a) Performance comparison of the three scenarios after stage one in case 3: large  $Q$  with a malfunctioning device



(b) Performance comparison of the three scenarios after stage two in case 2



(b) Performance comparison of the three scenarios after stage two in case 3

Fig. 3. Performance comparison of the three scenarios in case 2

Fig. 4. Performance comparison of the three scenarios in case 3

#### D. Case 4: large $Q$ with different types of interferences

With the same large  $Q$  as in case 3, case 4 has more complicated noise interferences: at  $t = 1$ , bus 22 voltage measurements are increased by a constant value of 4 for 0.1 second; at  $t = 2$ , bus 23 voltage measurements are interfered by a normally distributed noise  $\mathcal{N}(0, 5^2)$  for 0.2 second; at  $t = 3$ , line 22 – 21 current measurements are interfered by a uniformly distributed noise  $\mathcal{U}(0, 10)$  for 0.3 second; at  $t = 4$ , line 23 – 22 current measurements are decreased by a constant value of  $15i$  for 0.4 second. Note that although errors are treated as Gaussian white noises, they do not have to be Gaussian and white. The tracking results can be seen in Fig. 5.

It is clear the traditional KF is the most vulnerable to these disturbances. When the disturbance is large enough (e.g.,  $t = 1$ ), the naive RKF is capable of detecting the bad measurements and calibrating the estimates, but otherwise it underperforms. Fortunately, the AKF with InNoVa remains robust throughout. Moreover, as the corresponding measurement noise variances suddenly become abnormally large, the inflated  $R$  can be used to signal humans of a need to inspect

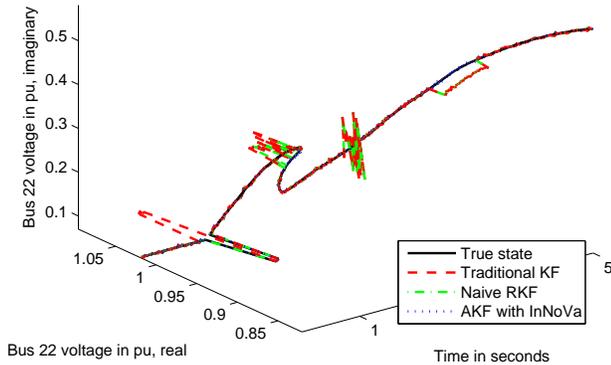
and repair devices. Fig. 5(a) and 5(b) both show that various noise disturbances do not affect Scenario 3.

## V. CONCLUSIONS

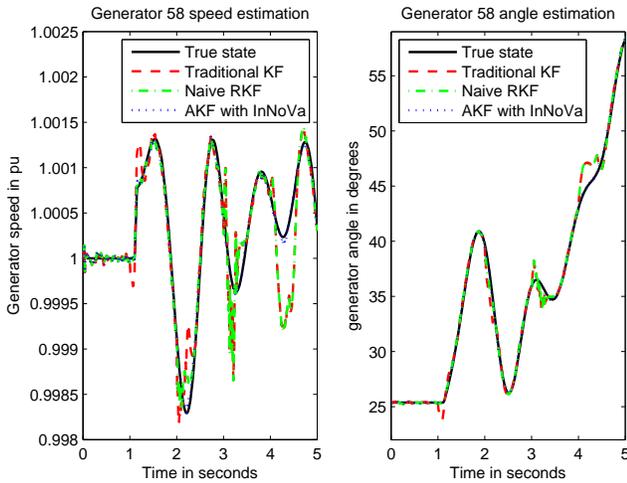
This paper presents a more comprehensive dynamic power system state estimation technique with a two-stage Kalman filtering approach. With real-time phasor measurements provided by PMUs, this algorithm enables on-line robust power system state estimations.

In stage one, *AKF with InNoVa* is designed to deal with incorrect system modeling as well as bad measurements, by adjusting noise modeling parameters on-the-fly. It utilizes a normalized *a priori innovation* test and a normalized *a posteriori innovation* test to help separating the process and measurement factors when facing terrible estimations. The adjusted noise parameters provide useful information about the system: the inflation of process noise covariance  $Q$  indicates fast changing state or even wrong model, while the inflation of measurement noise covariance  $R$  implies potentially bad PMU measurements. An exponential decay process is employed to

Bus 22 voltage estimation



(a) Performance comparison of the three scenarios after stage one in case 4: large  $Q$  with several noise interfered devices



(b) Performance comparison of the three scenarios after stage two in case 4

Fig. 5. Performance comparison of the three scenarios in case 4

enable automatic deflation of the parameters if the problems are resolved. This lightweight yet efficient algorithm gives better bus voltage phasor estimates, hence benefits stage two EKF in generating more accurate generator state estimation results. The output of both stages are essential for myriad time-critical analyses and applications.

## VI. ACKNOWLEDGEMENT

At the Pacific Northwest National Laboratory we acknowledge Ning Zhou, Pengwei Du, and Ruisheng Diao for helpful background discussions. This work is supported by U.S. Department of Energy grant DE-SC0002271 “Advanced Kalman Filter for Real-Time Responsiveness in Complex Systems,” PIs Zhenyu Huang at PNNL and Greg Welch at UNC (UCF). At DOE we acknowledge Sandy Landsberg, Program Manager for Applied Mathematics Research; Office of Advanced Scientific Computing Research; DOE Office of Science. This work is also supported in part by U.S. ONR Award# N00014-12-1-0052 by Peter Squire, HPTE Deputy Thrust Manager.

## REFERENCES

- [1] A. Phadke, J. Thorp, and M. Adamiak, “A new measurement technique for tracking voltage phasors, local system frequency, and rate of change of frequency” Power Apparatus and Systems, IEEE Transactions on, vol. PAS-102, no. 5, pp. 1025-1038, may 1983.
- [2] A. G. Phadke, J. S. Thorp, and K. J. Karimi, “State estimation with phasor measurements” Power Systems, IEEE Transactions on, vol. 1, no. 1, pp. 233-238, feb. 1986.
- [3] Ali Abur and Antonio Gomez Exposito “Power System State Estimation: Theory and Implementation”. Published by Marcel Dekker, 2004.
- [4] Jun Zhu and Abur, A. “Bad Data Identification When Using Phasor Measurements”. Power Tech, 2007 IEEE Lausanne, July 2007.
- [5] Zhongzhi Li, Xuegang Wang “Reverse Prediction Adaptive Kalman Filtering Algorithm for Maneuvering Target Tracking”. Journal of Computational Information Systems 6:10 (2010) 3257-3265.
- [6] K. R. Shih, S. J. Huang “Application of a Robust Algorithm for Dynamic State Estimation of a Power System”. IEEE Power Engineering Review, vol. 22, no. 1, 2002.
- [7] Zhenyu Huang, Kevin Schneider, and Jarek Nieplocha “Feasibility Studies of Applying Kalman Filter Techniques to Power System Dynamic State Estimation” The 8th International Power Engineering Conference IPEC2007, 3-6 December 2007, Singapore.
- [8] Zhenyu Huang, Pengwei Du, Dmitry Kosterev, and Bo Yang, “Application of Extended Kalman Filter Techniques for Dynamic Model Parameter Calibration” IEEE Power and Energy Society General Meeting 2009, Calgary, Canada, Institute of Electrical and Electronics Engineers, Piscataway, NJ, 2009
- [9] Greg Welch, and Gary Bishop “An introduction to the Kalman Filter”. TR 95-041, Department of Computer Science, University of North Carolina at Chapel Hill, April 2004.
- [10] Jinghe Zhang, Greg Welch, and Gary Bishop “Observability and Estimation Uncertainty Analysis for PMU Placement Alternatives”. 2010 North American Power Symposium (NAPS 2010), Arlington, TX, U.S.A., September 26-28, 2010.
- [11] Frank L. Lewis “Optimal Estimation: With an Introduction to Stochastic Control Theory”. published by Wiley-Interscience, April 1986.



**Jinghe Zhang** is a Ph.D candidate in computer science at the University of North Carolina (UNC) at Chapel Hill. Her current research interests include optimal sensor placement for the power grid, and large scale estimation in general. Zhang has a BS from the University of Science and Technology of China, and an MS from the University of Idaho, both in mathematics.



**Greg Welch** is a research professor of computer science at the University of North Carolina (UNC) at Chapel Hill. His current research interests include stochastic estimation, virtual and augmented reality, human tracking systems, 3D telepresence, projector-based graphics, and computer vision. Welch has a PhD in computer science from UNC-Chapel Hill. He is a member of the IEEE Computer Society and the ACM.



**Gary Bishop** is a professor of computer science at the University of North Carolina (UNC) at Chapel Hill. His current research interests include applications of computer technology to address the needs of people with disabilities, and systems for man-machine interaction. Bishop has a PhD in computer science from UNC-Chapel Hill.



**Zhenyu Huang** is a staff research engineer at the Pacific Northwest National Laboratory, Richland, WA, and a licensed professional engineer in the state of Washington. His research interests include power system stability and control, high-performance computing applications, and power system signal processing. Huang (M’01, SM’05) received his B. Eng. from Huazhong University of Science and Technology (1994) and Ph.D. from Tsinghua University (1999).