UNCERTAINTY-DRIVEN ADAPTIVE ESTIMATION WITH APPLICATIONS IN ELECTRICAL POWER SYSTEMS

Jinghe Zhang

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Computer Science

Chapel Hill
2013

Approved by:

Gregory F. Welch

Gary Bishop

Ning Zhou

Zhenyu Huang

Montek Singh

ABSTRACT

JINGHE ZHANG: Uncertainty-Driven Adaptive Estimation with Applications in Electrical Power Systems
(Under the direction of Gregory F. Welch)

From electrical power systems to meteorology, large-scale state-space monitoring and forecasting methods are fundamental and critical. Such problem domains pose challenges from both computational and signal processing perspectives, as they typically comprise a large number of elements, and processes that are highly dynamic and complex (e.g., severe nonlinearity, discontinuities, and uncertainties). This makes it especially challenging to achieve real-time operations and control.

For decades, researchers have developed methods and technology to improve the accuracy and efficiency of such large-scale state-space estimation. Some have devoted their efforts to hardware advances—developing advanced devices with higher data precision and update frequency. I have focused on methods for enhancing and optimizing the state estimation performance.

As uncertainties are inevitable in any state estimation process, uncertainty analysis can provide valuable and informative guidance for on-line, predictive, or retroactive analysis. My research focuses primarily on three areas:

1. Grid Sensor Placement. I present a method that combines off-line steady-state uncertainty and topology analysis for optimal sensor placement throughout the grid network.

2. Filter Computation Adaptation. I present a method that utilizes on-line state uncertainty analysis to choose the best measurement subsets from the available (large-scale) measurement data. This allows systems to adapt to dynamically available computational resources.

3. Adaptive and Robust Estimation. I present a method with a novel on-line measurement uncertainty analysis that can distinguish between suboptimal/incorrect system modeling and/or

erroneous measurements, weighting the system model and measurements appropriately in real-time as part of the normal estimation process.

We seek to bridge the disciplinary boundaries between Computer Science and Power Systems Engineering, by introducing methods that leverage both existing and new techniques. While these methods are developed in the context of electrical power systems, they should generalize to other large-scale scientific and engineering applications.

Dedicated to

My Family

ACKNOWLEDGEMENTS

When I first stepped into Sitterson Hall, our computer science department at UNC Chapel Hill, I did not know what to expect from my new graduate student life. And honestly, I did not expect that all things can fall into place so smoothly for me, for which I am very thankful.

Since I started writing this acknowledgements section, I tried to list all the people who have helped my graduate career, and then eventually I realized that would be mission-impossible. Because there are so many people who have inspired, supported and helped me, in one way or another. So as I write this acknowledgement to express my sincere gratitude, I apologize in advance to anyone who has been left out, and hope that I can at least remember those who have influenced me the most along the way. Many thanks to:

- Dr. Greg Welch, for giving me the opportunity to work as a research assistant, for being my advisor and friend, and for being the most enthusiastic, insightful and knowledgeable computer scientist that I have had the pleasure of working with;
- Dr. Gary Bishop for constantly inspiring me with his creative ideas, profound knowledge and cheerful personality, and for passing on “rationalizing everything" ability to me;
- Dr. Montek Singh for agreeing to serve on my committee, and for his support and insightful feedbacks;
- Dr. Zhenyu Huang for supporting me in all possible ways, and for hosting us during our visits to the Pacific Northwest National Lab (PNNL);
- Dr. Ning Zhou for his thoughtful advice and help on my research progress, and for providing helpful background information;
- Dr. Pengwei Du and Dr. Ruisheng Diao, for providing “real-time technique support" whenever I have questions about power systems;

- Janet Jones for helping me and looking out for me before her retirement, and for being a friend indeed (and in need) ever after;
- Dawn Andres for making my travel and commuting so easy and worry-free;
- Tianren Wang and Rukun Fan for being my friends and teammates. We have been helping, supporting and sharing ideas with each other unconditionally.

On a more personal note, I would like to thank:

- Yuchen Lu, my husband, for his love, patience and support. No matter what future holds, he will always be the one I can turn to. Thanks for being my rock!
- My parents for being an influence throughout my life. They are proud of me as much as I am proud of them;
- My newborn baby Eugene, for making my family and my life complete.

Financial support for this work came from:

- 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. At DOE we acknowledge Sandy Landsberg, Program Manager for Applied Mathematics Research; Office of Advanced Scientific Computing Research; DOE Office of Science;
- The Office of Naval Research award N00014-09-1-0813, “3D Display and Capture of Humans for Live-Virtual Training." We acknowledge Dr. Roy Stripling, Program Manager at ONR.
- The Office of Naval Research award N00014-12-1-0052, “3D Display and Capture of Humans for Live-Virtual Training." We acknowledge Peter Squire, HPTE Deputy Thrust Manager at ONR.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 Motivations and Topics

From electrical power grids to economics, from microcontroller chips to spacecrafts, systems are everywhere and a part of everything. Either complex or simple, a system at a given time point can be determined by a set of system variables. The smallest possible such set is known as the *internal state*, or simply the *state*. Because the set of state variables can fully specify the system at any give time, it is an essential element in system analysis and control.

While the true values of state variables may never be known exactly, they can be estimated using the available measurements and established state models. From an engineering perspective, numerous professionals have devoted their efforts to developing advanced measuring technologies and data processing technologies, so that measuring and computing devices are constantly developing towards higher accuracy and greater efficiency.

Other than designing better measuring instruments, processing data on larger and faster computational facilities, what else can we do to improve the estimation qualities? Here we propose three topics from a computational and system-wide perspective:

- Grid Sensor Placement. When the measurement sensors are expensive/limited, it is important to install them at the most informative locations throughout the system network, so that they can provide measurement data with optimal observability and reliability.
- Filter Computation Adaptation. When the massive measurement data processing and

limited computational resources become obstacles in real-time state estimation, we hope to reduce the computational complexity of the estimator. Strategic reduction of the number of measurements used per estimate allows us to do so without sacrificing the estimation quality.

- Adaptive and Robust Estimation. In a hostile environment where the hypothetical models do not match the actual systems, and/or the measurements contain significant errors, the estimated states could deviate from the true states rapidly. The influenced estimation has to be corrected as much as possible.

Our research on these topics, with applications in power systems, can serve as a bridge between Computer Science and Power Engineering. Furthermore, they can be extended to other large-scale complex systems as well. The background, previous work and our proposed ideas related to these three topics will be explained in the following subsections.

1.1.1 Grid Sensor Placement

In our particular project we are dealing with advanced measurement sensors across the power grid. Traditionally, power grid measurements are provided by *remote terminal units* (RTUs) at the substations. RTU measurements include real/reactive power flows, power injections, and magnitudes of bus voltages and branch currents.

A novel *phasor measurement technology* was developed in the 1980s [2], [3]. The crucial devices, called *phasor measurement units*(PMUs), are able to measure both magnitude and phase angle of the electrical waves in a power grid. With the global positioning system (GPS), PMU provide high-speed and high-accuracy sensor data with precise time synchronization. These synchronized phasor measurements are commonly referred to as *synchrophasors*. As a comparison, RTU measurements are much slower and are not well synchronized.

While PMU measurements currently cover fewer than 1% of the nodes in the U.S. power grid, the power industry actively invests in advancing the technology and installing more units. However at the present stage, the installation must be selective because of resource constraints.

Previous PMU placement research has focused primarily on network topology, with the goal of finding configurations that achieve full network observability with a minimum number of PMUs [4, 5, 6, 7]. However we have noticed two issues about this problem: 1) The observability of the entire system is not an ąřall or nothingąś problem. PMUs are added to the grid incrementally, and it will be useful to have some guidance on where to install them. 2) The power system networks usually have complex topologies, so that when we obtain more than one “optimal placement" solution with the same number of PMUs, one will have to make a selection among these solutions.

To address these issues, recently we introduced an evaluation method that utilizes stochastic models of the signals and measurements to characterize the observability and corresponding uncertainty of power system states, for any given configuration of PMUs [8]. This evaluation method includes a novel observability checking algorithm and estimation uncertainty analysis. The observability checking classifies all the observable buses into one or more levels, according to the ąřdirectnessąś of their observations: Level 1 contains the most directly observable buses (PMU buses), while the buses with more “indirect" (Section 2.8.1) observations are labeled with higher levels. Then based on the previous work [9, 10], we compute a stochastic estimate of the asymptotic or steady-state error covariance, P^{∞}. P^{∞} is used as a quantitative metric for the performance evaluation. We generalize our approach for any PMU placement layout without requirements for achieving full observability.

Because PMUs can provide real-time synchrophasor data to the Supervisory Control And Data Acquisition (SCADA) system to capture the dynamic characteristics of the power system, they have been found useful in estimating more transient states [11]. Thus we have extended our evaluation method to both *static* (bus voltage magnitudes and phase angles) and *dynamic*

(generator rotor speed and rotor angles) state estimation applications. In [12], a new approach is presented to design an optimal PMU placement according to estimation uncertainties of the power system *dynamic* states. We restate the optimal PMU placement problem so that the solution is not only a minimum set of PMUs that can cover the entire power system, but also the one helps dynamic state estimation the most. So if an integer programming solver returns multiple placement solutions, which all achieve full observability with the same minimum number of PMUs, we apply steady-state uncertainty analysis to evaluate the solutions. The optimal PMU placement is hence selected according to the evaluation results. This approach can provide planning engineers with a new tool to help in the selection between PMU placement alternatives.

1.1.2 Filter Computation Adaptation

Dynamic state estimation (DSE) techniques, such as the Kalman filters [13], have been the subject of extensive research and applications. Especially in modern power systems, DSE is a fast developing research area. Compared to traditional steady state estimators, DSE techniques are able to track dynamic state variables both efficiently and accurately.

However, in large-scale and wide-area inter-connected power systems, the combination of high computational complexity and slow processing speed presents a significant challenge. To help address this challenge, it is a good idea to employ faster response filters with reduced measurement-space.

In our recent work we have set our sight on measurement-space reduction: we attempt to select and process the measurement subsets which can improve state calculations the most. As inspired by the previous single-constraint-at-a-time (SCAAT) method [14], an approach we call Lower Dimensional Measurement-space (LoDiM) state estimation was developed in [15, 16].

The LoDiM approach employs an estimator such as the Kalman filter, but with a twist: while the filter is running in the foreground, it also performs a measurement selection procedure in the background to optimally reduce the measurement dimension. The measurement selection procedure utilizes the principal component analysis (PCA) of the error covariance, to determine the subset of the state space to update most urgently (e.g. with larger estimation uncertainties than others). We prefer to target at the principal uncertainties first, thus the measurements are ranked according to their abilities to calibrate these uncertainties. Only a smaller subset that can reduce the principal uncertainties most efficiently is then processed by the foreground filter during each cycle. As a result, during each cycle it yields much smaller computational load and hence lower latency. Furthermore, we have also extended our approach to Reduced Measurement-space Dynamic State Estimation (ReMeDySE), for estimating the *dynamic state* of power systems [17].

Although we present LoDiM in the context of the Kalman filter, the associated measurement selection procedure is not filter-specific, i.e. it can be used with other state estimation methods, as long as the state uncertainties are dynamically estimated along with the states. Our approach is not application-specific either: it can be applied to other large-scale, real-time on-line and computationally intensive state tracking systems with modern parallel computation techniques.

1.1.3 Adaptive and Robust Estimation

Kalman filters achieve optimal performance only when the system noise characteristics have known statistics that obey certain properties (zero-mean, Gaussian, and spectrally white). However in practice, it is difficult to obtain the process and measurement noise models. When the theoretical models do not match the actual models, and/or the measurements contain significant errors, the estimated state can deviate from the true state rapidly. When people notice

estimated state deviate from true state, it is usually impossible to determine whether the deviation is caused by a process model mismatch, a measurement model mismatch, or both.

To address this problem, we propose an efficient adaptive approach called the Adaptive Kalman Filter (AKF) with Inflatable Noise Variances (InNoVa). Primary simulations have demonstrated the robustness of this approach facing sudden changes of system dynamics and erroneous measurements, with the ability to adjust noise modeling parameters on-the-fly. It is capable of doing these because besides the regular *normalized innovation test*, we also employ a *normalized residual test* to help separate the process and measurement factors.

Furthermore, we have considered incorporating AKF with InNoVa into the dynamic state estimation algorithm from previous research [11], to also help with the estimation of power system dynamic states. Thus we developed a two-stage Kalman filter 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. In the first stage, we estimate the static states from raw PMU measurements, using AKF with InNoVa, which is able to identify and reduce the impact of incorrect system modeling and/or bad PMU measurements. Then in the next stage, the estimated bus voltages are fed into a second extended Kalman filter to obtain the dynamic state estimation. Case studies show the inherent advantages of our two-stage Kalman filter technique over the extended Kalman filter alone.

1.2 Thesis Statement

Combinations and comparisons of process and measurement uncertainties can be used to assess/optimize measurement unit placement, improve computational efficiency and flexibility, and increase system robustness. Specifically,

1. steady-state uncertainty analysis can be combined with topological analysis

to evaluate or optimize measurement unit placement in a network;

2. sensitivity-to-uncertainty ratio ranking can be used to select measurement subsets to optimally adapt to the available computational resources; and

3. innovation-residual tests can be used to identify apparent outlier measurements, and adapt the process/measurement models.

1.3 Contributions

- Our grid sensor placement approach is able to evaluate and compare any candidate sensor placement configuration in complex networks, regardless of achieving full observability.

  While off-line steady-state uncertainty analysis has aided in optimal sensor placement for continuous-space systems such as the 3D tracking system, the combination of this with a network topology analysis enables sensor placement evaluation and comparison for discrete-space systems such as the power system.

- Our Lower Dimensional Measurement-space (LoDiM) state estimation algorithm, adapting to the computation budget available, can help many large-scale, real-time on-line and computation-intensive state tracking systems because

  1) it features a dynamic measurement selection procedure: a small measurement subset is dynamically and strategically chosen for each cycle. The smaller measurement-space incorporate less information per cycle, but has higher reporting rates;

  2) the measurement selection procedure is not filter-specific, i.e. it can be used with other state estimation methods such as particle and unscented filters.

- Our Adaptive Kalman Filter with Inflatable Noise Variances (AKF with InNoVa) enables robust state estimation. Its ability of adjusting noise modeling parameters on-the-fly can efficiently

1) identify and adapt to incorrect system modeling;

2) identify and reduce the impact of erroneous measurements.

1.4 Thesis Organization

There are three main topics in this thesis. They are individually studied, yet closely related. Among which, I investigate uncertainty-aware grid sensor placement stratagies in Chapter 2. Next, Chapter 3 discusses how to use uncertainty-guided measurement selection approach to reduce computational burden. Then I present our lately developed adaptive and robust state estimation algorithms, based on uncertainty analyses, in Chapter 4. Finally, Chapter 5 summarizes our work, and proposes possible plans for the future research.

For each thesis topic, I will state the structure of its own chapter more specifically in the “Chapter Organization" section.

CHAPTER 2
GRID SENSOR PLACEMENT

The synchronized phasor measurement unit (PMU), developed in the 1980s [2, 3], is considered one of the most important devices in the future of power systems. While PMU measurements currently cover fewer than 1% of the nodes in the U.S. power grid, the power industry has gained the momentum to advance the technology and install more units. However, with limited resources, the installation must be selective.

Previous PMU placement research has focused primarily on the network topology, with the goal of finding configurations that achieve full network observability with a minimum number of PMUs. Here we introduce a new approach that also includes stochastic models for the signals and measurements, to characterize the observability and corresponding uncertainty of any given configuration of PMUs, whether that configuration achieves full observability or not. Utilization of this approach to designing optimal PMU placements, according to estimation uncertainties of the power system static states/dynamic states, are also discussed in details. This approach can provide planning engineers with a new tool to help choose between PMU placement alternatives.

2.1 Introduction

State estimation plays a crucial role in determining the health of a power system. State can be estimated using the available measurements and system model. Traditionally, remote terminal units (RTU) at the substations provide power grid measurements. The most commonly used RTU measurements in state estimation are:

- Line power flow measurements: the real and reactive power flow along the transmission lines or transformers.
- Bus power injection measurements: the real and reactive power injected at the buses.
- Voltage magnitude measurements: the voltage magnitudes of the buses.

Under certain circumstances such as state estimation of distribution systems, the line current magnitude measurements may be considered, which provide the current flow magnitudes along the transmission lines or transformers.

With the ability to impact future power system monitoring and operation methods, phasor measurement technology is one of the key enabling technologies for the future “smart grid". Phasor measurement units (PMUs) make two additional types of measurements available:

- Voltage phasor measurements: the phase angles and magnitudes of voltage phasors at system buses.
- Current phasor measurements: the phase angles and magnitudes of current phasors along transmission lines or transformers.

Recent development of phasor measurement technologies offers high-speed sensor data (typically 30 or 60 samples/second) with precise time synchronization. This is in comparison with traditional RTU measurements in SCADA system, which have cycle times of 2 to 4 seconds and are not well synchronized.

PMUs are becoming increasingly attractive in various power system applications such as system monitoring, protection, control, and stability assessment. Widely scattered in the power system network and synchronized from the common global positioning system (GPS) radio clock, PMUs can provide real-time synchrophasor data to the SCADA system to capture the dynamic characteristics of the power system, and hence facilitate time-critical applications

such as dynamic state estimation and dynamic stability analysis [11]. Compared to estimating relatively stationary state elements such as bus voltage magnitudes and phase angles, *dynamic state estimation* seeks to estimate more transient states of a power system, i.e. generator rotor angles and speeds. With properly placed and sufficient numbers of PMUs, we can estimate the system status in real-time using time stamped synchrophasors.

There has been significant work in selecting the best locations to install PMUs [4, 5, 6, 7]. Many algorithms have been developed, primarily with the aim of utilizing a minimum number of PMUs to ensure full network observability. However, two issues caught our attention:

- First, the observability of the entire system is not an "all or nothing" problem. Due to various resource limitations, we do not have the luxury of installing a complete set of PMUs throughout the power grid. Yet we are not starting from scratch either. In practice people are installing PMUs to the grid incrementally, thus it would be useful to have some guidance on where to install them.
- Second, in the previous research, the network topology was the primary focus. In practice the networks usually have complex topologies where more than one solution for the same minimum number of PMUs will appear. In such cases, the planning engineers need to make a choice from these solutions.

We focus on bridging these gaps. Specifically our contributions include:

- We develop a stochastic model that captures state estimation uncertainties, to facilitate PMUs installation assessment.
- We design an optimal PMU placement evaluation algorithm by incorporating uncertainty estimates into topological considerations for the specific network.
- We present an approach to the comparison among alternative configurations via quantitative measure of expected uncertainties.

2.2 Chapter Organization

The rest of this chapter is organized as follows. Section 2.3 introduces the background. In Section 2.4 and Section 2.5, I review the basic concepts of state space models and steady state estimation briefly. Section 2.6 discusses traditional power system state estimation method. Section 2.7 describes the PMU devices and their impact on power system state estimation. Then for static (Section 2.8) and dynamic (Section 2.9) state estimation respectively, we present our approach to determine the effects of any given PMU configuration, and restate the optimal PMU placement problem such that the solution is not only a minimum set of PMUs that can cover the entire power system, but also the set that improves state estimation the most. Our methods are simulated on multi-machine system models to demonstrate how it might help engineers make decisions on different candidate plans.

2.3 Background and Related Work

The phasor measurement unit (PMU) was first developed and utilized in [2, 3]. Considering partially observable systems (with an inadequate number of PMUs) the authors in [18] presented an estimation algorithm based on singular value decomposition (SVD), which did not require fully observable system prior to estimation.

Optimal PMU placement for full observability was studied in [4]. They developed an algorithm for finding the minimum number of PMUs required for power system state estimation, using simulated annealing optimization and graph theory in formulating and solving the problem.

In [7] the authors focused on the analysis of network observability and PMU placement when using a mixed measurement set. They developed an optimal placement algorithm for

PMUs using integer programming. In [5], a strategic PMU placement algorithm was developed to improve the bad data processing capability of state estimation by taking advantage of the PMU technology. Furthermore, [6] re-studied the PMU placement problem and presented a generalized integer linear programming formulation.

Our work, however, is taking a different approach: we first characterize the observability and corresponding uncertainty of the power system buses for any given configuration of PMUs, whether that configuration achieves full observability or not. Then we also presented our work on connecting the optimal PMU placement with power system static as well as dynamic state estimation: we provide a method for evaluating any candidate PMU placement design, which is especially useful when there are multiple placement candidates [8], [12].

2.4 State Space Models

The state space models are the most basic yet extensively used mathematical models in all kinds of state estimation problems. Here I present a brief review to introduce the relevant mathematical notations used throughout this dissertation. For more detailed description, the readers can refer to [19].

An assumed *linear* system can be modeled as a set of linear stochastic process equation and measurement equation:

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1} \tag{2.1}$$

$$z_k = Hx_k + v_k \tag{2.2}$$

where $x \in \mathcal{R}^n$ is the state 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, without either a driving function or process noise[1], $u \in \mathcal{R}^l$ is the optional control input vector. B is a $n \times l$ matrix that relates the control input

[1]In practice, the matrix A may change with each time step, but it is assumed to be constant here.

vector to the state x_k, $z \in \mathcal{R}^m$ is the measurement vector, H is a $m \times n$ matrix that relates the state x_k to the measurement z_k. $w \in \mathcal{R}^n$ and $v \in \mathcal{R}^m$ are process and measurement noise vectors respectively.

The process noise w and measurement noise v are assumed to be mutually independent random variables, white, and with normal probability distributions:

$$p(w) \sim N(0, Q) \tag{2.3}$$

$$p(v) \sim N(0, R) \tag{2.4}$$

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

In reality, the process to be estimated and (or) the measurement relationship to the process is usually *non-linear*. Especially when our objective is to estimate the dynamic states of the power system. A non-linear system can be modeled using non-linear stochastic process and measurement equations

$$x_k = a(x_{k-1}, u_{k-1}) + w_{k-1} \tag{2.5}$$

$$z_k = h(x_k) + v_k. \tag{2.6}$$

where $x \in \mathcal{R}^n$ is the state vector, a is a non-linear function that relates the state at the previous time step $k-1$ to the state at the current step k. u is the (optional) driving function, considered as parameters in function a. $z \in \mathcal{R}^m$ is the measurement vector. h is another non-linear function that relates the state x_k to the measurement z_k. w and v again represent the zero-mean process and measurement noise as defined before.

These non-linear functions can be linearized about the point of interest x in the state space. To do so one need to compute the state Jacobian and measurement Jacobian matrices

$$A = \frac{\partial a(x)}{\partial x}|_x \tag{2.7}$$

$$H = \frac{\partial h(x)}{\partial x}|_x \tag{2.8}$$

where A and H are the partial derivatives of a and h, with respect to x.

For any state estimate $\hat{x}_k$, we define the estimate error as $e_k \equiv x_k - \hat{x}_k$ and estimate error covariance as $P_k \equiv E[e_k e_k^T]$, where E denotes statistical expectation. For the state estimate $\hat{x}_k$ at the time step k, the estimate error covariance P_k contains important information, reflecting the uncertainty of the estimation.

2.5 Steady State Estimation

On-line methods such as the Kalman filter [13, 19] can be used to estimate time varying state and error covariance, by minimizing the *a posteriori* estimate error covariance P_k in a recursive predictor-corrector fashion. More details can be found in Section 3.4. In these on-line scenarios the estimate error covariance P_k changes over time. However the *steady-state* error covariance

$$\begin{aligned} P^\infty &= \lim_{k\to\infty} E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T] \\ &= \lim_{k\to\infty} E[e_k e_k^T] \end{aligned} \tag{2.9}$$

can be computed off-line in closed form. Here x and $\hat{x}$ represent the true and estimated states respectively, and E denotes statistical expectation. In fact to compute the steady-state uncertainty one does not actually need to estimate x_k and $\hat{x}_k$. Instead one can estimate P^∞ directly from state-space models of the system using stochastic models for the various noise sources.

The rest of this section reviews the method for estimating P^∞ based on previous work [9, 10], which employs the Discrete Algebraic Riccati Equation (DARE). The DARE represents a closed-form solution to the steady-state covariance P^∞ [20]. Assuming the process and measurement noise elements are uncorrelated, it can be written in this form:

$$P^\infty = AP^\infty A^T + Q - AP^\infty H^T (R + HP^\infty H^T)^{-1} HP^\infty A^T \tag{2.10}$$

The derivation of the DARE will be explained in Section 3.4.1.

The DARE solution P^{∞} can be calculated using the MacFarlane-Potter-Fath "Eigenstructure Method" [20]. Given the model parameters A, Q, H, and R from the last subsection, the $2n \times 2n$ discrete-time Hamiltonian matrix is written as

$$\Psi = \begin{pmatrix} A + QA^{-T}H^{T}R^{-1}H & QA^{-T} \\ A^{-T}H^{T}R^{-1}H & A^{-T} \end{pmatrix}. \tag{2.11}$$

Then after calculating n characteristic eigenvectors $[e_1, e_2, ..., e_n]$ of Ψ, we form a matrix

$$\begin{pmatrix} B \\ C \end{pmatrix} = [e_1, e_2, ..., e_n] \tag{2.12}$$

Finally, we use B and C to compute the steady-state error covariance

$$P^{\infty} = BC^{-1}. \tag{2.13}$$

Note that if the Jacobians A and H are functions of the state, we will generally have to compute them at each point of interest in the state space. The noise covariance matrices R and Q might be constant, or might also vary as a function of the state.

For each point of interest in the state space, P^{∞} indicates the expected asymptotic state estimation uncertainty corresponding to the candidate design modeled by the specific A, Q, H, and R. Intuitively, given PMU placements leading to the same level of observability, the lower the overall uncertainty is, the more we prefer this design.

2.6 Traditional Power System State Estimation

Research in power system state estimation can be traced back to the late 1960ąś. The conventional state estimator provides the estimates of the power system *static* states, i.e. bus voltage phasors (magnitudes and phase angles), based on measurements obtained from the

SCADA system. Usually, people use weighted least squares (WLS) estimators to find the best estimates of the states. In this subsection, we give a brief review of traditional power system state estimation method and model, based on the classic textbook [21]. We refer this procedure as a *static state estimation* because it obtains the voltage phasors at all buses at a single point in time. In other words, WLS estimator only focuses on the measurement equations (2.2, 2.6) in the state space model, without considering the process equations (2.1, 2.5).

2.6.1 Weighted Least Squares Estimator

The measurements used by the estimator are the real and reactive power flows and bus injections, and the voltage magnitudes $|V|$, as provided by SCADA data. Traditionally the phase angles, θ, are not known. It is assumed that bad data have been eliminated from this measurement set by the usual methods. The measurements are usually formulated in vector notation as

$$z = h(x) + v \tag{2.14}$$

where z is the measurement vector, x is the state vector, v is the vector of the measurement errors from noise, and the vector function h is the nonlinearity between the power measurements and the state.

The WLS estimator aims at minimizing the objective function J, which is formulated as the sum of the squares of the weighted residuals:

$$J(x) = [z - h(x)]^T R^{-1} [z - h(x)] \tag{2.15}$$

where R is the measurement error covariance matrix. At the kth iteration of the WLS solution, the residual vector r is defined as

$$r^k = z - h(x^k). \tag{2.16}$$

To minimize the objective function J, the first-order optimality conditions must be satisfied, that is

$$g(x) = \frac{\partial J(x)}{\partial x} = -H^T(x)R^{-1}[z - h(x)] = 0, \tag{2.17}$$

where

$$H(x) = \frac{\partial h(x)}{\partial x} \tag{2.18}$$

is the Jacobian matrix obtained by taking partial derivatives of h with respect to x.

A non-linear optimization algorithm, such as the Newton-Raphson technique, is used to iteratively converge to a solution. Assuming the Newton-Raphson method and the first-order Taylor expansion are applied, along with an initial value of x, the step at each iteration k becomes:

$$x^{k+1} = x^k + [G(x^k)]^{-1}H^T(x^k)R^{-1}[z - h(x^k)], \tag{2.19}$$

where $G(x^k)$ is the *gain matrix* given by

$$G(x^k) = \frac{\partial g(x^k)}{\partial x} = H^T(x^k)R^{-1}H(x^k). \tag{2.20}$$

The gain matrix is positive definite and symmetric provided that the system is fully observable, and it is normally sparse.

The estimator is considered to have converged when the reduced-order measurement mismatch vector $H^T(x^k)R^{-1}[z - h(x^k)]$ has reached some prescribed threshold.

2.6.2 The Measurement Function and Measurement Jacobian

As mentioned before, the most commonly used measurements provided by traditional SCADA system are the line power flows, bus power injections, bus voltage magnitudes, and

sometimes line current magnitudes. These measurements can be expressed in terms of the state variables either using the rectangular or the polar coordinates—the conventional choice is polar coordinates. Here we briefly describe the measurement equations provided by [21].

If a system contains N buses, the state vector will have $(2N-1)$ elements: N bus voltage magnitudes and $(N-1)$ bus voltage phase angles, as one bus is chosen as the reference bus with its phase angle set equal to an arbitrary value such as 0. This bus is also known as the *slack bus* or *swing bus*.

Typically we choose bus 1 as the reference bus. The state vector will have the following form:

$$x = [\theta_2 \theta_3 ... \theta_N V_1 V_2 ... V_N]^T. \tag{2.21}$$

Assuming the general two-port π-model for the power system transmission lines (as shown in Fig. 2.1),

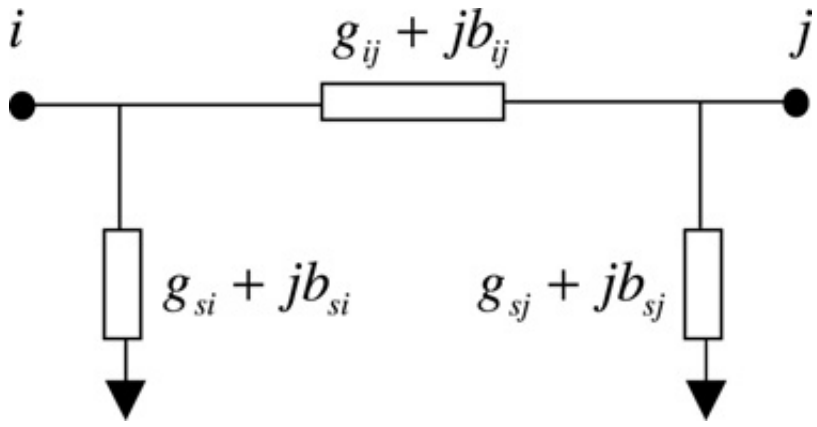


Figure 2.1: The two-port π-model of a transmission line

the above mentioned measurements can be expressed in the following equations:

- Real power injection at bus i:

$$P_i = V_i \sum_{j=1}^{N} V_j (G_{ij} \cos\theta_{ij} + B_{ij} \sin\theta_{ij}) \tag{2.22}$$

- Reactive power injection at bus i:

$$Q_i = V_i \sum_{j=1}^{N} V_j (G_{ij} \sin\theta_{ij} - B_{ij} \cos\theta_{ij}) \tag{2.23}$$

- Real power flow from bus i to bus j:

$$P_{ij} = V_i^2 (g_{si} + g_{ij}) - V_i V_j (g_{ij} \cos\theta_{ij} + b_{ij} \sin\theta_{ij}) \tag{2.24}$$

- Reactive power flow from bus i to bus j:

$$Q_{ij} = -V_i^2 (b_{si} + b_{ij}) - V_i V_j (g_{ij} \sin\theta_{ij} - b_{ij} \cos\theta_{ij}) \tag{2.25}$$

- Line current flow magnitude from bus i to bus j:

$$I_{ij} = \frac{\sqrt{P_{ij}^2 + Q_{ij}^2}}{V_i} \tag{2.26}$$

  Because in practice the shunt admittance $g_{si} + jb_{si}$ is much smaller comparing to the series admittance $g_{ij} + jb_{ij}$, the shunt admittance is usually ignored:

$$I_{ij} = \sqrt{(g_{ij}^2 + b_{ij}^2)(V_i^2 + V_j^2 - 2V_i V_j \cos\theta_{ij})} \tag{2.27}$$

- Voltage magnitude at bus i:

$$V_i = V_i \tag{2.28}$$

where

V_i is the voltage magnitude at bus i,

θ_i is the phase angle at bus i,

$\theta_{ij} = \theta_i - \theta_j$ is the phase angle difference between buses i and j,

$g_{si} + jb_{si}$ is the shunt admittance connected at bus i,

$g_{ij} + jb_{ij}$ is the series admittance of the transmission line connecting buses i and j,

$G_{ij} + jB_{ij}$ is the ijth element of the complex bus admittance matrix.

The measurement Jacobian matrix H then has the following form:

$$H = \begin{pmatrix} \frac{\partial P_{inj}}{\partial \theta} & \frac{\partial P_{inj}}{\partial V} \\ \frac{\partial P_{flow}}{\partial \theta} & \frac{\partial P_{flow}}{\partial V} \\ \frac{\partial Q_{inj}}{\partial \theta} & \frac{\partial Q_{inj}}{\partial V} \\ \frac{\partial Q_{flow}}{\partial \theta} & \frac{\partial Q_{flow}}{\partial V} \\ \frac{\partial I_{mag}}{\partial \theta} & \frac{\partial I_{mag}}{\partial V} \\ 0 & \frac{\partial V_{mag}}{\partial V} \end{pmatrix}. \tag{2.29}$$

where

P_{inj} and Q_{inj} are the real and reactive power injection measurements,

P_{flow} and Q_{flow} are the real and reactive power flow measurements,

I_{mag} and V_{mag} are the line current flow magnitude and voltage magnitude measurements.

The elements of this measurement Jacobian matrix H can be derived from the measurement equations conveniently:

- Partial derivatives of real power injection at bus i:

$$\frac{\partial P_i}{\partial \theta_i} = \sum_{j=1}^{N} V_i V_j(-G_{ij}\sin\theta_{ij} + B_{ij}\cos\theta_{ij}) - V_i^2 B_{ii} \tag{2.30}$$

$$\frac{\partial P_i}{\partial \theta_j} = V_i V_j(G_{ij}\sin\theta_{ij} - B_{ij}\cos\theta_{ij}) \tag{2.31}$$

$$\frac{\partial P_i}{\partial V_i} = \sum_{j=1}^{N} V_j(G_{ij}\cos\theta_{ij} + B_{ij}\sin\theta_{ij}) + V_i G_{ii} \tag{2.32}$$

$$\frac{\partial P_i}{\partial V_j} = V_i(G_{ij}\cos\theta_{ij} + B_{ij}\sin\theta_{ij}) \tag{2.33}$$

- Partial derivatives of reactive power injection at bus i:

$$\frac{\partial Q_i}{\partial \theta_i} = \sum_{j=1}^{N} V_i V_j(G_{ij}\cos\theta_{ij} + B_{ij}\sin\theta_{ij}) - V_i^2 G_{ii} \tag{2.34}$$

$$\frac{\partial Q_i}{\partial \theta_j} = V_i V_j(-G_{ij}\cos\theta_{ij} - B_{ij}\sin\theta_{ij}) \tag{2.35}$$

$$\frac{\partial Q_i}{\partial V_i} = \sum_{j=1}^{N} V_j(G_{ij}\sin\theta_{ij} - B_{ij}\cos\theta_{ij}) - V_i B_{ii} \tag{2.36}$$

$$\frac{\partial Q_i}{\partial V_j} = V_i(G_{ij}\sin\theta_{ij} - B_{ij}\cos\theta_{ij}) \tag{2.37}$$

- Partial derivatives of real power flow from bus i to bus j:

$$\frac{\partial P_{ij}}{\partial \theta_i} = V_i V_j(g_{ij}\sin\theta_{ij} - b_{ij}\cos\theta_{ij}) \tag{2.38}$$

$$\frac{\partial P_{ij}}{\partial \theta_j} = -V_i V_j(g_{ij}\sin\theta_{ij} - b_{ij}\cos\theta_{ij}) \tag{2.39}$$

$$\frac{\partial P_{ij}}{\partial V_i} = -V_j(g_{ij}\cos\theta_{ij} + b_{ij}\sin\theta_{ij}) + 2(g_{ij} + g_{si})V_i \tag{2.40}$$

$$\frac{\partial P_{ij}}{\partial V_j} = -V_i(g_{ij}\cos\theta_{ij} + b_{ij}\sin\theta_{ij}) \tag{2.41}$$

- Partial derivatives of reactive power flow from bus i to bus j:

$$\frac{\partial Q_{ij}}{\partial \theta_i} = -V_i V_j (g_{ij} \cos\theta_{ij} + b_{ij} \sin\theta_{ij}) \tag{2.42}$$

$$\frac{\partial Q_{ij}}{\partial \theta_j} = V_i V_j (g_{ij} \cos\theta_{ij} + b_{ij} \sin\theta_{ij}) \tag{2.43}$$

$$\frac{\partial Q_{ij}}{\partial V_i} = -V_j (g_{ij} \sin\theta_{ij} - b_{ij} \cos\theta_{ij}) - 2(b_{ij} + b_{si}) V_i \tag{2.44}$$

$$\frac{\partial Q_{ij}}{\partial V_j} = -V_i (g_{ij} \sin\theta_{ij} - b_{ij} \cos\theta_{ij}) \tag{2.45}$$

- Partial derivatives of line current flow magnitude from bus i to bus j, shunt admittance ignored:

$$\frac{\partial I_{ij}}{\partial \theta_i} = \frac{g_{ij}^2 + b_{ij}^2}{I_{ij}} V_i V_j \sin\theta_{ij} \tag{2.46}$$

$$\frac{\partial I_{ij}}{\partial \theta_j} = -\frac{g_{ij}^2 + b_{ij}^2}{I_{ij}} V_i V_j \sin\theta_{ij} \tag{2.47}$$

$$\frac{\partial I_{ij}}{\partial V_i} = \frac{g_{ij}^2 + b_{ij}^2}{I_{ij}} (V_i - V_j \cos\theta_{ij}) \tag{2.48}$$

$$\frac{\partial I_{ij}}{\partial V_j} = \frac{g_{ij}^2 + b_{ij}^2}{I_{ij}} (V_j - V_i \cos\theta_{ij}) \tag{2.49}$$

- Partial derivatives of voltage magnitude at bus i:

$$\frac{\partial V_i}{\partial V_i} = 1 \tag{2.50}$$

$$\frac{\partial V_i}{\partial V_j} = 0 \tag{2.51}$$

$$\frac{\partial V_i}{\partial \theta_i} = 0 \tag{2.52}$$

$$\frac{\partial V_i}{\partial \theta_j} = 0 \tag{2.53}$$

2.7 The Phasor Measurement Units (PMUs)

Synchronized by GPS satellite clock, PMUs have achieved a level of precision that typically exceeds the conventional measurements, which makes phasor telemetry a valuable source of

measurements. Nowadays, PMUs are becoming more and more attractive in various power system applications such as system monitoring, protection, control, and stability assessment.

The benefit of PMUs is especially evident when it comes to power system state estimation. Conventional power system state estimation uses real and reactive power as measurements to estimate the bus phasor voltages. As a result, the relationship between measurements and states is non-linear. The solution is always obtained by linearizing the model and solving it in an iterative fashion.

PMUs alleviate this problem by providing the phasors of voltages and currents measured at a given substation. Hence, by measuring the bus phasor voltages and line phasor currents, the relationship between measurements and states becomes linear. When the measurement equation is linear, the estimation algorithm is direct and significantly faster than the non-linear ones.

2.7.1 Observability Checking

For modeling power systems, we consider two types of measurements: PMU measurements and *zero injection* measurements.

A PMU installed at a specific bus is capable of measuring not only the bus voltage phasor, but also the current phasors along all the lines incident to the bus. Hence in addition to the phasor voltage at this bus, we are able to compute the phasor voltages of all of its neighboring buses.

If a bus has neither generators nor loads, it is a *zero injection bus* [21]. The inclusion of zero injection buses has been proven effective in reducing the number of PMUs needed for an optimal placement plan [7]. Because the sum of flows on all the associated branches to the bus is zero, this equality normally corresponds to a pseudo measurement called the *zero injection*

measurement [7], which is useful in system state estimation. According to Kirchhoff's current law, we know that at any bus in the power grid, the sum of currents flowing into that bus is equal to the sum of currents flowing out of that bus. Thus, for an injection-measured bus and its n neighbors ($(n+1)$ buses total), if the phasor voltages at any n out of the $(n+1)$ buses are known (observable), then the remainder can also be computed and hence become observable.

Consistent with previous studies, we only consider isolated zero injection buses in our test systems. Meaning that, once a zero injection bus location is used in our algorithm, other zero injection buses adjacent to it (if there is any) will not be taken into account. This is because although zero injection buses can help reduce the PMUs needed for a system, the reduction of PMUs will compromise the system reliability. More detailed explanation can be found in [22].

Based on these assumptions, we have developed the observability checking algorithm shown in Algorithm 1.

2.8 My Method: PMU Placement for Power System Static State Estimation

2.8.1 PMU Placement Evaluation

Previous studies consider a PMU placement optimal in the sense that it *maximizes observability* while *minimizing cost*. The authors in [6, 7] presented a numerical formulation of this problem, and developed an optimal placement algorithm for PMUs using integer linear programming. This formulation allows easy analysis of the network observability and concerns about the full coverage of the system. However, the integer linear programming approach may be insufficient for determining the optimal locations of PMUs. The reason is that very often there will be multiple optimal solutions with the same minimum number of PMUs. In such

Algorithm 1: Observability checking for a given set of PMU locations and zero injection buses.

Input: PMU buses location P_bus and zero injection buses location Z_bus
Output: Observable buses location O_bus
$O\_bus = \emptyset$;
$O\_bus = P\_bus$ (the buses that can be measured directly by PMUs);
for $i \leftarrow 1$ **to** $|P\_bus|$ **do**
  $O\_bus = O\_bus \cup neighbors(P\_bus_i)$
(the $|\cdot|$ operator denotes the cardinality of the set);
(adding the buses that can be estimated through at least one adjacent PMU bus);
$flag = 1$;
while *flag* **do**
  $flag = 0$;
  for $j \leftarrow 1$ **to** $|Z\_bus|$ **do**
    if
    $(|Z\_bus_j \cup neighbors(Z\_bus_j)) \cap O\_bus| = |Z\_bus_j \cup neighbors(Z\_bus_j)| - 1$
    then
      $O\_bus = O\_bus \cup Z\_bus_j \cup neighbors(Z\_bus_j)$;
      $flag = 1$;
      (adding the buses that can be estimated using the zero injection information. According to the Kirchhoff's law: for a zero injection bus and its neighborhood, if all but one is known, then the unknown one can be estimated);

cases, one would need a means for comparing the different solutions, as some will actually offer lower estimate uncertainty and latency.

Built on the previous work by Allen *et al.* [9, 10], our approach is to use a stochastic estimate of the *asymptotic* or *steady-state* error covariance, as a quantitative metric for the performance evaluation. Our approach is generalized for any PMU placement layout, without requirements for achieving full observability.

To estimate the steady-state error covariance, we first consider the relevant state space models.

The Process Model

For modeling the state and covariance changes over time, there are several ways to identify the parameters A and Q, both *a priori* and on-line [10, 20]. Following the methodology described in [23], a quasi-static model of an assumingly stable power system is used in our approach. The oscillations in the state variables of this model are assumed small, thus the states of the power system at time step $(k+1)$ are the same as those at time step k, except for some zero mean, white Gaussian noise. Hence our process model reduces to

$$x_k = x_{k-1} + w_{k-1}. \tag{2.54}$$

The Measurement Model

Traditionally, real and reactive power measurements are used in the power system state estimation, resulting in non-linear measurement models and iterative algorithm to solve for the state of the system. Now with the help of PMUs, the measurement model becomes linear: we simply let the phasor bus voltages be the state, while phasor bus voltages and phasor currents be the measurements. In this way we are able to use non-iterative estimation algorithms, which speeds up the computation signficantly.

The measurement model has the general form

$$z = Hx + v, \tag{2.55}$$

where $x \in \mathcal{R}^n$ is the complex state vector, $z \in \mathcal{R}^m$ is the complex measurement vector, and H is a $m \times n$ complex matrix that relates the state to the measurement z_k. The measurement noise elements in v are assumed to be mutually independent random variables, white, and with normal probability distributions

$$p(v) \sim N(0, R). \tag{2.56}$$

For computational convenience, we convert the complex valued measurement model into a real valued measurement model as in [18]:

$$\begin{pmatrix} Re(z) \\ Im(z) \end{pmatrix} = \begin{pmatrix} Re(H) & -Im(H) \\ Im(H) & Re(H) \end{pmatrix} \begin{pmatrix} Re(x) \\ Im(x) \end{pmatrix} + \begin{pmatrix} Re(v) \\ Im(v) \end{pmatrix} \tag{2.57}$$

where $Re(\cdot)$ and $Im(\cdot)$ stand for the real and imaginary parts of the corresponding vector/matrix respectively.

After the procedure of observability checking, all the observable buses are divided into one or more levels, according to the "directness" of their observations: with Level 1 containing the most directly observable buses, and the buses with more "indirect" observations labeled with higher levels. Note that if the power grid only uses PMU measurements, the observable buses belong to either Level 1 or Level 2; however, if both PMU measurements and injection measurements (which could be zero injection pseudo-measurements) are taken into account, there may be more than two levels of observable buses.

Level 1 buses

If the bus i is in Level 1 (meaning that there is a PMU placed on this bus), the measurement equation is

$$z_i = x_i + v_i, \tag{2.58}$$

where z_i is the measured complex voltage at bus i, x_i is the "true" complex voltage at bus i and v_i is the complex measurement noise of this PMU. Thus for all the buses in Level 1, we can write the measurement equation in the matrix form

$$z_V = I \cdot x_{L_1} + v_V, \tag{2.59}$$

where z_V is the complex voltage measurement subvector, I is the identity matrix, x_{L_1} is the complex state subvector ("true" complex voltages at all Level 1 buses) and v_V is the voltage measurement noise subvector.

Level 2 buses

If the bus i is in Level 2 (meaning that there is at least one PMU placed on an adjacent bus), then for each PMU placed at some adjacent bus j, the measurement equation will be

$$z_{ji} = \begin{pmatrix} Y_{ji} & -Y_{ji} \end{pmatrix} \begin{pmatrix} x_j \\ x_i \end{pmatrix} + v_j, \tag{2.60}$$

where z_{ji} is the measured complex current at bus j (towards bus i), Y_{ji} is the admittance of line (j, i), x_j and x_i are the "true" complex voltages at bus j and i respectively, and v_j is the complex measurement noise of this PMU. So for all the buses in Level 1 and Level 2, we have the measurement equation

$$z_C = \begin{pmatrix} Y_{CL_1} & Y_{CL_2} \end{pmatrix} \begin{pmatrix} x_{L_1} \\ x_{L_2} \end{pmatrix} + v_C, \tag{2.61}$$

where z_C is the complex current measurement subvector, Y_{CL_1} and Y_{CL_2} are the line admittance matrices that relates Level 1 and Level 2 bus voltages to z_C respectively, x_{L_1} and x_{L_2} are the complex state subvector ("true" complex voltages at all Level 1 and Level 2 buses), and v_C is the current measurement noise subvector. In this equation, the measurement matrix $\begin{pmatrix} Y_{CL_1} & Y_{CL_2} \end{pmatrix}$ has each row sum up to zero. If the number of PMUs installed is sufficient, or the power grid is well-connected, it is quite possible that a Level 2 bus has more than one PMU-installed neighbor buses. In such cases, the measurement equation has fused the contributions from each adjacent PMU.

Level 3 or above buses

If the bus i is in Level 3 or above (meaning that we can compute the bus voltage, but there is no PMU placed on itself or any adjacent bus), then we must use the information about net injection current measurements, and Kirchhoff's Current Law, to calculate the complex bus voltage. Generally there are two possibilities:

- an injection measurement is taken at bus i; or
- an injection measurement is taken at bus j, which is adjacent to bus i.

Overall, assuming there are totally l levels of observable buses, the measurement equation can be written as

$$z_I = \begin{pmatrix} Y_{IL_1} & Y_{IL_2} & Y_{IL_3} \cdots & Y_{IL_l} \end{pmatrix} \begin{pmatrix} x_{L_1} \\ x_{L_2} \\ \vdots \\ x_{L_l} \end{pmatrix} + v_I, \tag{2.62}$$

where z_I is the complex injection current measurement subvector (zero pseudo-measurements), Y_{IL_1} through Y_{IL_l} are the node admittance matrices of all observable buses that are related to z_I, x_{L_1} through x_{L_l} represent the complex state subvector for all observable buses, and v_I is the current measurement noise subvector.

Combining the above three cases, the complete measurement model can be expressed as

$$\begin{pmatrix} z_V \\ z_C \\ z_I \end{pmatrix} = \begin{pmatrix} I & 0 & 0\cdots & 0 \\ Y_{CL_1} & Y_{CL_2} & 0\cdots & 0 \\ Y_{IL_1} & Y_{IL_2} & Y_{IL_3}\cdots & Y_{IL_l} \end{pmatrix} \begin{pmatrix} x_{L_1} \\ x_{L_2} \\ \vdots \\ x_{L_l} \end{pmatrix} + \begin{pmatrix} v_V \\ v_C \\ v_I \end{pmatrix}. \tag{2.63}$$

The measurement model can be converted into real valued model, using the technique mentioned before.

2.8.2 Optimal PMU Placement

Conventionally, the PMU placement is considered optimal in the sense that it makes *all buses* in the system observable with a *minimum number of PMUs*. This is especially true for large-scale complex systems. The authors in [6, 7] presented a numerical formulation of this optimization problem:

$$\min \sum_{i}^{n(buses)} w_i \cdot X_i \tag{2.64}$$

$$s.t. \quad f(X) = C \cdot X \geq \hat{1}$$

where $n(buses)$ is the number of buses in the system, w_i is the cost of the PMU installed at bus i, X is a binary decision variable vector with entries X_i defined as

$$X_i = \begin{cases} 1 & \text{if a PMU is installed at bus } i \\ 0 & \text{otherwise} \end{cases}, \tag{2.65}$$

$\hat{1}$ is a vector whose entries are all ones, $f(X)$ is a vector function and C is the binary connectivity matrix with entries

$$C_{[k,m]} = \begin{cases} 1 & \text{if } k = m \\ 1 & \text{if buses } k \text{ and } m \text{ are connected} \\ 0 & \text{otherwise} \end{cases} . \tag{2.66}$$

This formulation allows easy analysis of the network observability and concerns about the full coverage of the system. We first solve this optimization problem using integer linear programming techniques. However, the integer linear programming approach is insufficient for determining the optimal locations of PMUs. The reason is that very often there will be multiple “optimal" solutions with the same minimum number of PMUs. In such cases, one would need a means for comparing the different solutions, as some will actually offer lower estimate uncertainty and latency.

Thus the next step is to evaluate the solution(s) using steady-state uncertainty analysis as described in section 2.5. Naturally, we prefer the solution with lower uncertainties. If there is no “absolute winner", (i.e. at each bus, the uncertainty value computed for this solution is lower than any other solution), the user need to define a criterion to judge each solution, and we will discuss more in the next section. Briefly, our approach for optimal PMU placement using uncertainty evaluation can be described by Algorithm 2.

2.8.3 Simulation Results

In this section, we first apply our observability checking and estimation uncertainty analysis to two simulated multi-machine systems. We then apply our approach to evaluate four different optimal PMU placement alternatives, and find the best one.

Algorithm 2: Optimal PMU Placement for Static State Estimation

```
Build the optimization problem as shown in (2.64) based on the network topology;
if there is more than one solution to this problem then
    for each solution set of PMUs do
        Evaluate A according to section 2.8.1;
        Evaluate Q according to section 2.8.1;
        Evaluate H according to section 2.8.1;
        Evaluate R according to section 2.8.1;
        Compute Ψ using (2.11);
        Compute the n eigenvectors of Ψ using (2.12);
        Compute P^∞ using (2.13);
    Apply the uncertainty values to user-defined criterion;
    Determine the optimal PMU placement strategy;
else
    Return the only possible solution set of PMUs;
```

Observablity Checking and Estimation Uncertainty

Given the sets of PMU buses and zero injection buses, Algorithm 1 checks observability by returning a list of observable buses classified at different levels. Then the steady state error covariance P^{∞} of these observable buses is computed, with each diagonal element P_{ii}^{∞} (i.e. the variance) indicating the estimation uncertainty of the corresponding bus. The estimation uncertainty of any unobservable bus i is $P_{ii}^{\infty} = \infty$. To better illustrate the results, we define $1/\sqrt{P_{ii}^{\infty}}$ to be the estimation "certainty" (information) of the corresponding bus. In this way, the certainty of any unobservable bus is zero; as for any observable bus, the higher certainty value it has, the better it can be estimated.

We simulated our algorithm on one small system consisting of three machines in a looped network of nine buses as shown in Fig. 2.2.

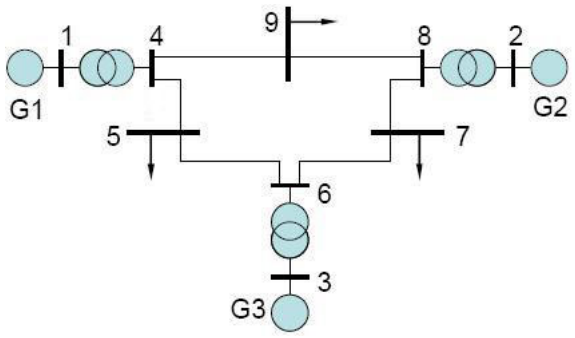


Figure 2.2: The 3-machine 9-bus system

Table 2.1 presents the observability analysis of three simulated cases. Fig. 2.3, 2.4 and 2.5 illustrates the certainty plots respectively. In all the following certainty plots, the green-colored numbers in cirles represent the PMU buses, while the blue-colored numbers in squares represent the zero-injection buses.

Table 2.1: Observability Checking for the 3-machine 9-bus System

Case	PMU buses	0-inj buses	Observable buses
1	1, 2	None	*L1*: 1, 2 *L2*: 4, 8
2	1, 2, 3	None	*L1*: 1, 2, 3 *L2*: 4, 6, 8
3	1, 2, 3	9	*L1*: 1, 2, 3 *L2*: 4, 6, 8 *L3*: 9

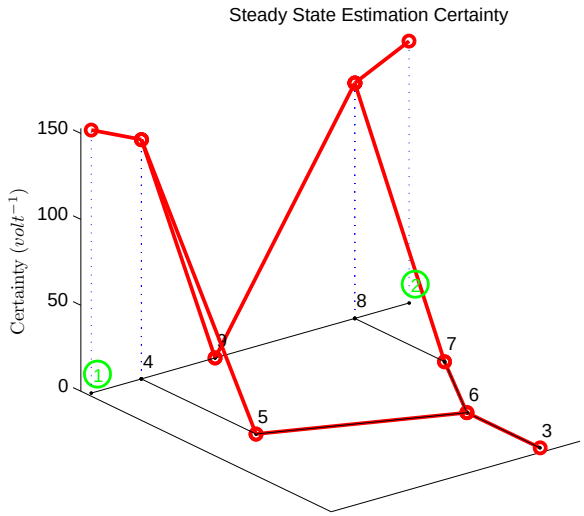


Figure 2.3: The certainty analysis plot of case 1 from Table 2.1, in the 3-machine 9-bus system

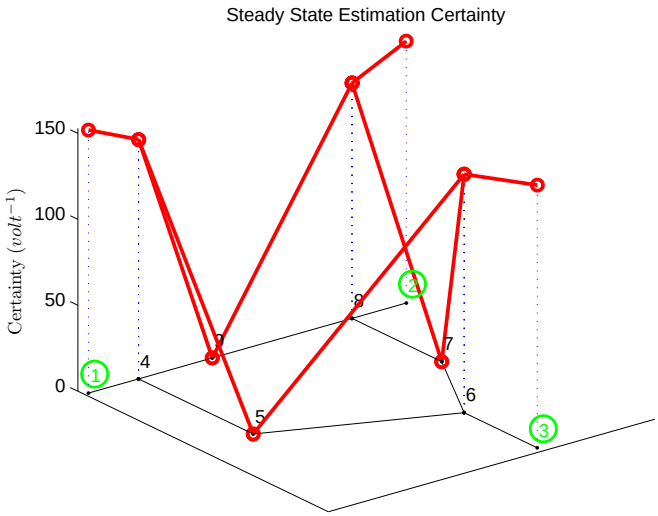


Figure 2.4: The certainty analysis plot of case 2 from Table 2.1, in the 3-machine 9-bus system

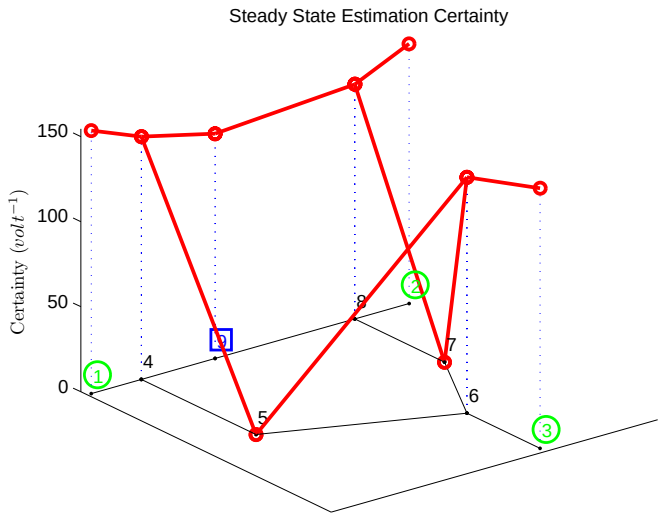


Figure 2.5: The certainty analysis plot of case 3 from Table 2.1, in the 3-machine 9-bus system

Then similarly, we simulated our method on a larger 16-machine 68-bus system as shown in Fig. 2.6, representing the interconnected New England Test System and New York Power System (NETS-NYPS) [24], with also three cases shown in Table 2.2 and the certainty analysis

of each one in Fig. 2.7, 2.8 and 2.9.

Table 2.2: Observability Checking for the 16-machine 68-bus System

Case	PMU buses	0-inj buses	Observable buses
1	1, 11, 21,	None	*L1*: 1, 11, 21 *L2*: 2, 6, 10, 12, 16, 22, 27, 30, 31, 47
2	1, 11, 21, 31, 41, 51	None	*L1*: 1, 11, 21, 31, 41, 51 *L2*: 2, 6, 10, 12, 16, 22, 27, 30, 38, 40, 42, 45, 47, 50, 62, 66
3	1, 11, 21, 31, 41, 51	12, 38, 40, 42, 50	*L1*: 1, 11, 21, 31, 41, 51 *L2*: 2, 6, 10, 12, 16, 22, 27, 30, 38, 40, 42, 45, 47, 50, 62, 66 *L3*: 13 48 52 *L4*: 67

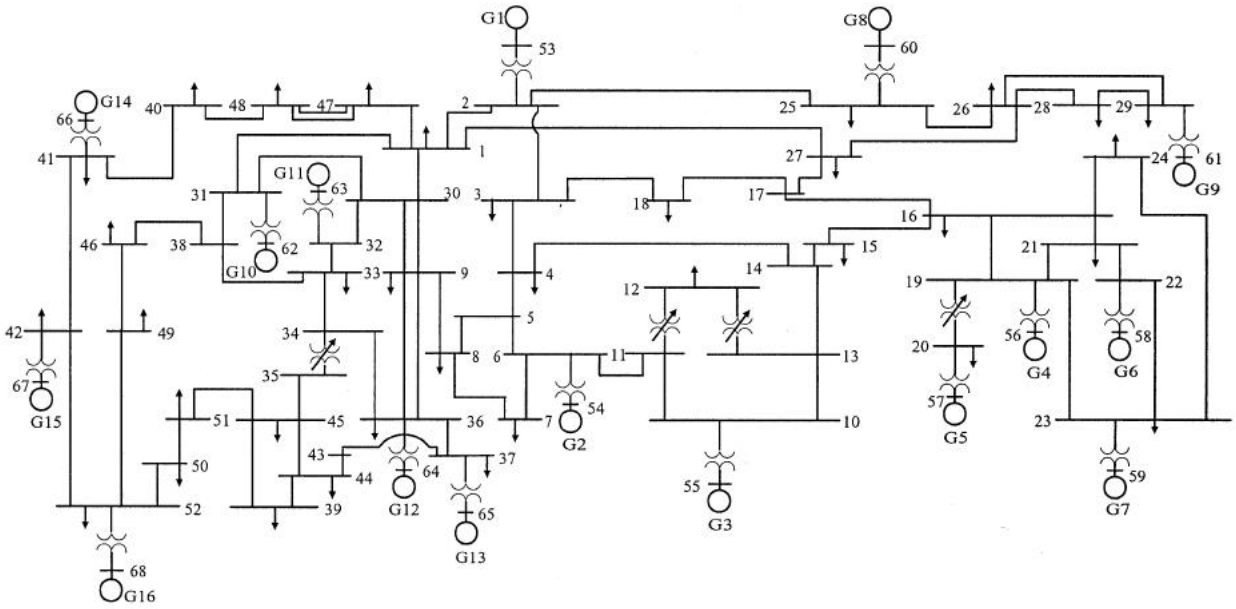


Figure 2.6: The 16-machine 68-bus system

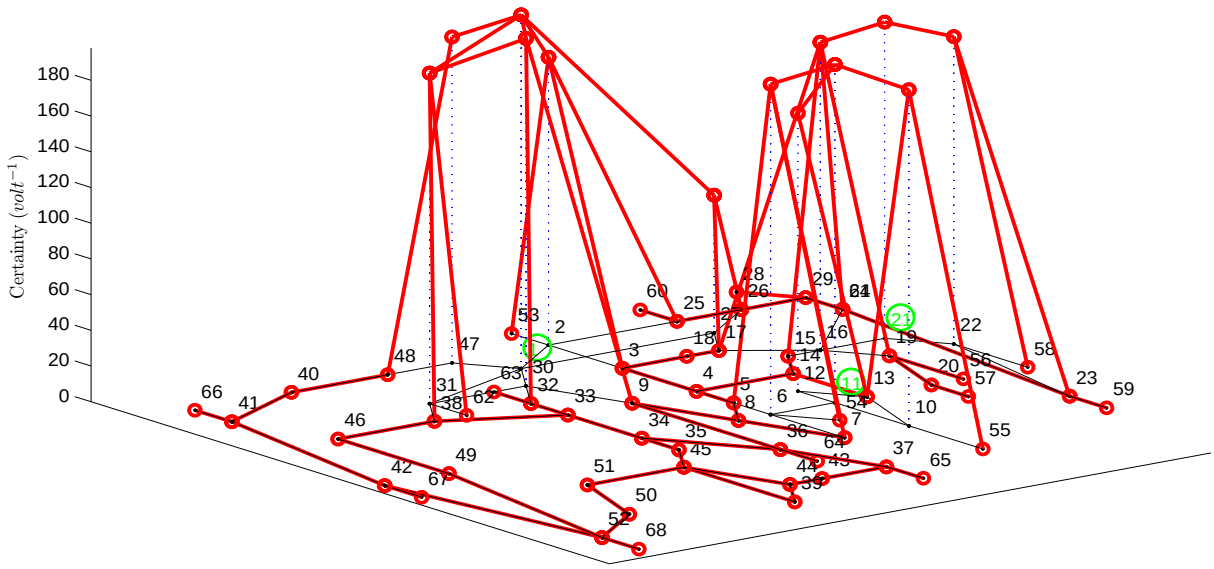


Figure 2.7: The certainty analysis plot of case 1 from Table 2.2, in the 16-machine 68-bus system

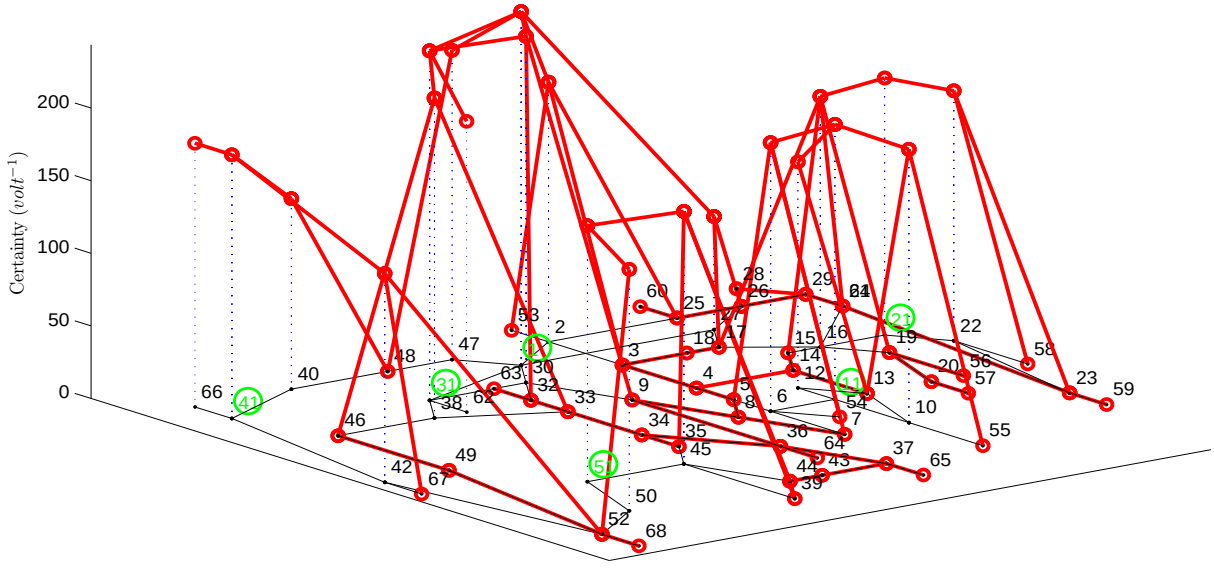


Figure 2.8: The certainty analysis plot of case 2 from Table 2.2, in the 16-machine 68-bus system

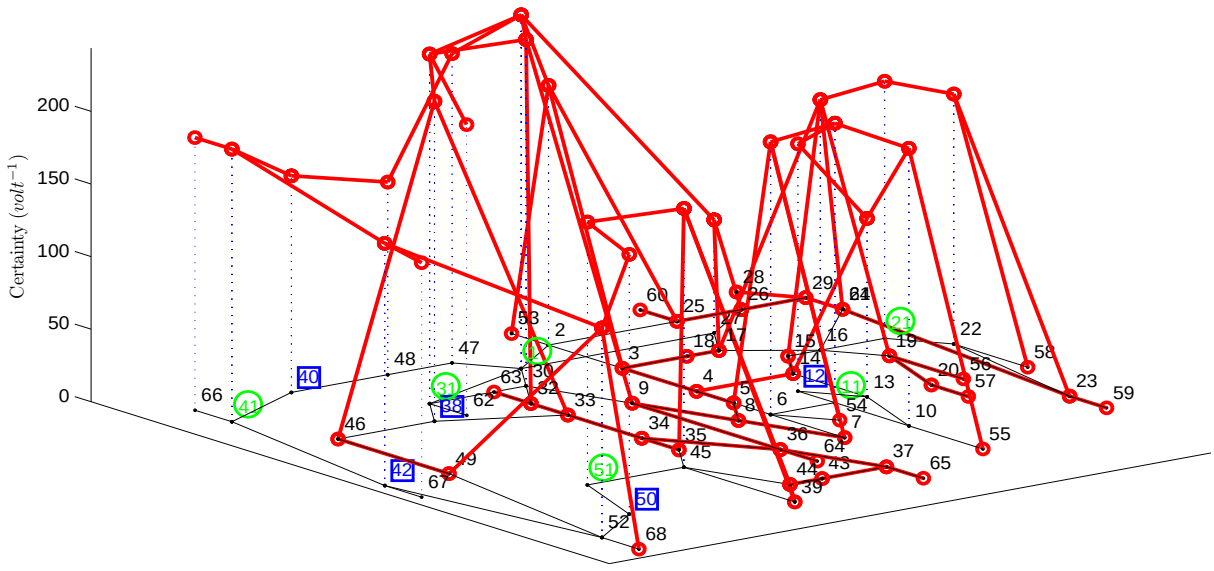


Figure 2.9: The certainty analysis plot of case 3 from Table 2.2, in the 16-machine 68-bus system

Comparison Among Multiple Optimal Solutions

In this subsection, we used the small system in Fig. 2.2 as an example and only considered PMU measurements. By solving the traditional optimization problem, we could in fact obtain four "optimal" solutions, with PMUs installed at buses $\{1, 6, 8\}$, $\{2, 4, 6\}$, $\{3, 4, 8\}$ and $\{4, 6, 8\}$ respectively. They are all "optimal" in the sense that they all make the entire system observable by using a minimum number of PMUs. To determine which PMU placement strategy is the best, we demonstrate the certainty plots of them in Fig. 2.10-2.13.

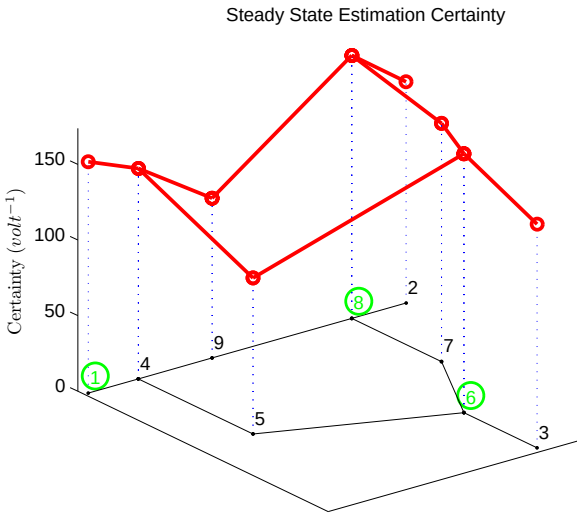


Figure 2.10: PMUs installed at buses 1, 6 and 8

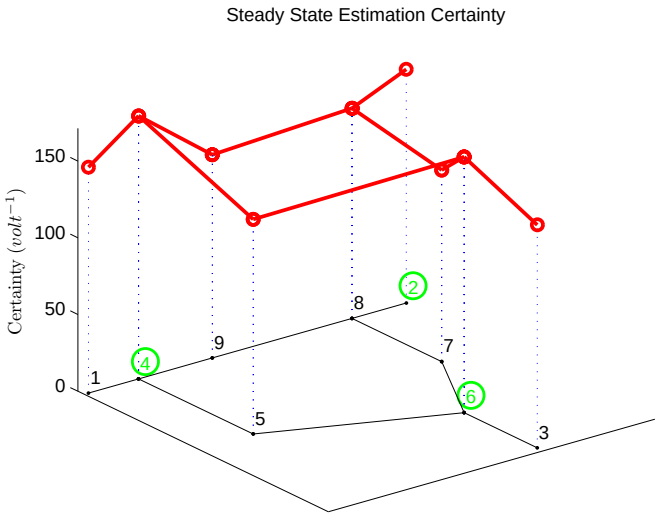


Figure 2.11: PMUs installed at buses 2, 4 and 6

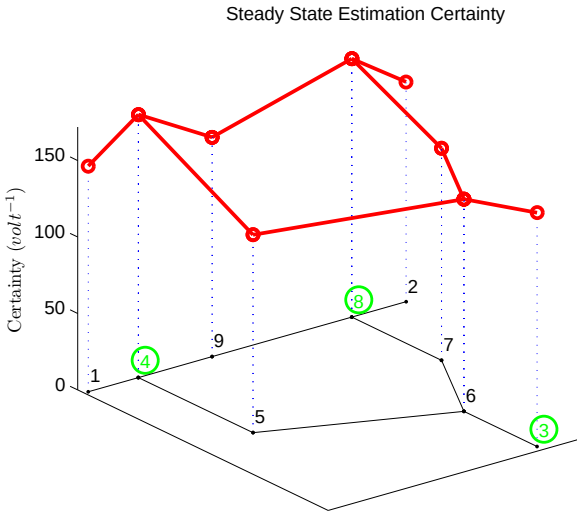


Figure 2.12: PMUs installed at buses 3, 4 and 8

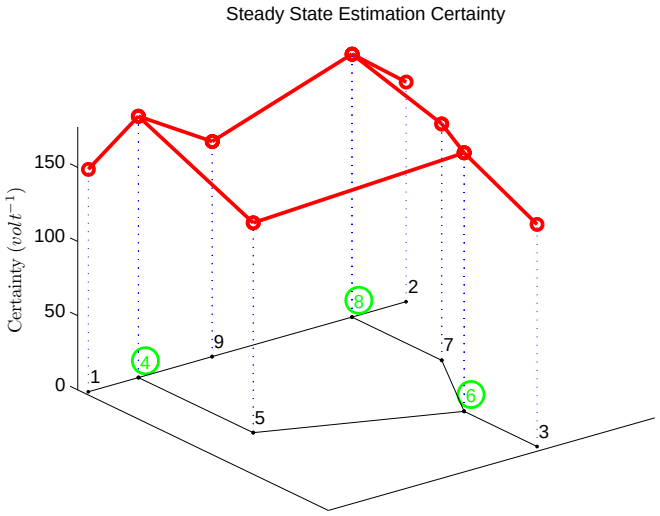


Figure 2.13: PMUs installed at buses 4, 6 and 8

Furthermore, it is up to the users to decide what criterion they prefer to use in making their specific choices. For instance, one can use the *maximum*, the *minimum*, the *mean* of these "certainties" (Fig. 2.14), or even some weighted function of them. From a practical point

of view, the users may want to choose the *minimum* as their criterion, because for the power system to function properly, even the worst estimates should exceed a certain standard. In our example, we can immediately tell that by all means, the PMU placement at buses $\{4, 6, 8\}$ outperforms other strategies.

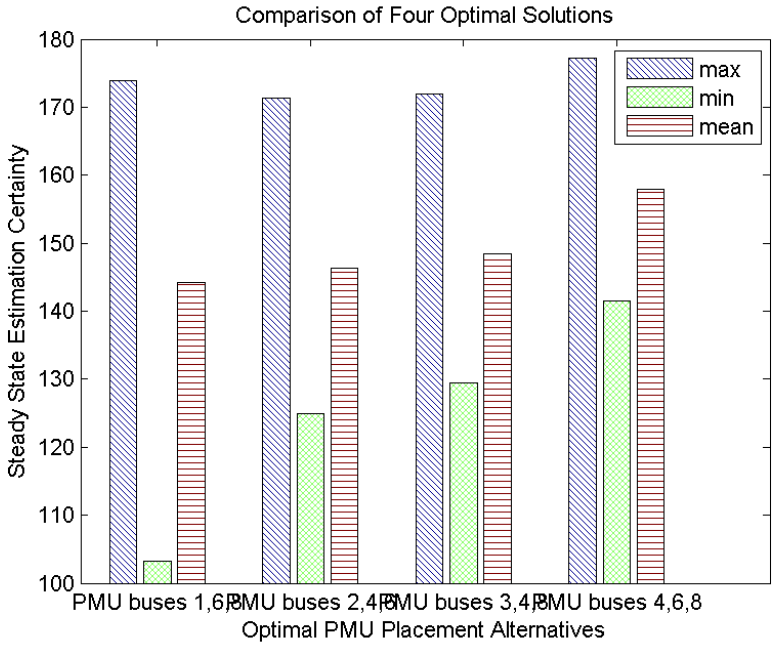


Figure 2.14: A comparison of steady-state certainties for the four “optimal" solutions, based upon three criteria: maximum (max), minimum (min) and average (mean).

2.9 My Method: PMU Placement for Power System Dynamic State Estimation

2.9.1 PMU Placement Evaluation

Previously we presented a stochastic approach to characterize the observability and corresponding uncertainty of the power system buses for any given configuration of PMUs, achieving full observability or not. Here we connect the optimal PMU placement with power system

dynamic state estimation: we restate the optimal PMU placement problem such that the solution is not only a minimum set of PMUs that can cover the entire power system, but also the one which benefits dynamic state estimation the most.

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}. As given in [25], the non-linear differential-algebraic equations regarding the generator i can be written as

$$\begin{cases} \frac{d\delta_i}{dt} = \omega_B(\omega_i - \omega_0) \\ \frac{d\omega_i}{dt} = \frac{\omega_0}{2H_i}(P_{m_i} - \frac{|E_i||V_i|}{X'_{d_i}}\sin(\delta_i - \theta_i) - D_i(\omega_i - \omega_0)) \end{cases} \tag{2.67}$$

where state variables δ and ω are the generator rotor angle and speed respectively, ω_B and ω_0 are the speed base and the synchronous speed in *per unit* [2], P_{m_i} is the mechanical input, H_i is the machine inertia[3], D_i is the generator damping coefficient and $|V_i|\angle\theta_i$ is the phasor voltage at the generator terminal bus i (which is a function of $\delta_1, \delta_2, ..., \delta_n$).

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

$$\frac{dx}{dt} = A_c x + w_c, \tag{2.68}$$

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,

[2] In the power systems analysis field of electrical engineering, a *per-unit system* is the expression of system quantities as fractions of a defined base unit quantity. Readers can refer to http://en.wikipedia.org/wiki/Per-unit_system

[3] 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.

whose entries for $i \in \{1, ..., n\}$ and $j \in \{1, ..., n\}$ ($i \neq j$) are the corresponding partial derivatives [25]

$$A_{c_{[2i-1,2i-1]}} = 0 \tag{2.69}$$

$$A_{c_{[2i-1,2i]}} = \omega_B \tag{2.70}$$

$$A_{c_{[2i,2i-1]}} = -\frac{\omega_0 |E_i|}{2H_i X'_{d_i}} \left[\frac{\partial |V_i|}{\partial \delta_i} \sin(\delta_i - \theta_i) + |V_i| \cos(\delta_i - \theta_i)(1 - \frac{\partial \theta_i}{\partial \delta_i}) \right] \tag{2.71}$$

$$A_{c_{[2i,2i]}} = -\frac{\omega_0}{2H_i} D_i \tag{2.72}$$

$$A_{c_{[2i-1,2j-1]}} = 0 \tag{2.73}$$

$$A_{c_{[2i-1,2j]}} = 0 \tag{2.74}$$

$$A_{c_{[2i,2j-1]}} = -\frac{\omega_0 |E_i|}{2H_i X'_{d_i}} \left[\frac{\partial |V_i|}{\partial \delta_j} \sin(\delta_i - \theta_i) + |V_i| \cos(\delta_i - \theta_i)(-\frac{\partial \theta_i}{\partial \delta_j}) \right] \tag{2.75}$$

$$A_{c_{[2i,2j]}} = 0. \tag{2.76}$$

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

$$A = I + A_c \cdot \Delta t. \tag{2.77}$$

The next issue is the discrete-time process noise covariance Q. As described by [20], if we assume the process noise "flows" through (is shaped by) the same system of integrators represented by A, we can integrate the continuous time process equation (3.48) over the time interval Δt to obtain a $2n \times 2n$ discrete-time (sampled) Q matrix:

$$Q = \int_0^{\Delta t} e^{A_c t} Q_c e^{A_c^T t} dt \tag{2.78}$$

Now written in the form of the process model from equation (2.1) we have

$$x_k = A x_{k-1} + w_{k-1} \tag{2.79}$$

$$= (I + A_c \cdot \Delta t) x_{k-1} + w_{k-1}, \tag{2.80}$$

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

The Measurement Model

For now we only consider the measurements provided by PMUs. According to [5, 6, 7], and our analysis of the bus observation "directness" in Section 2.8.1, all the observable buses are divided into two *levels*: Level 1 contains the directly observable buses with PMUs installed on them; while Level 2 contains the more "indirect" observable buses with PMUs installed on their neighbors instead of themselves.

As a result, the measurements are functions of the phasor voltages of all observable buses:

$$\begin{pmatrix} z_V \\ z_C \end{pmatrix} = \begin{pmatrix} I & 0 \\ Y_{CL_1} & Y_{CL_2} \end{pmatrix} \begin{pmatrix} V_{L_1} \\ V_{L_2} \end{pmatrix} + \begin{pmatrix} v_V \\ v_C \end{pmatrix} \tag{2.81}$$

We can simply denote the equation above as

$$z' = H''V + v' \tag{2.82}$$

However in this section, our state variables are no longer bus voltages, but dynamic generator state variables, so we need to introduce the expanded system nodal equation first [11]:

$$Y_{exp} \begin{pmatrix} E \\ V \end{pmatrix} = \begin{pmatrix} Y_{GG} & Y_{GL} \\ Y_{LG} & Y_{LL} \end{pmatrix} \begin{pmatrix} E \\ V \end{pmatrix} = \begin{pmatrix} I_G \\ 0 \end{pmatrix} \tag{2.83}$$

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.

Therefore the relationship of the bus voltages to the generator voltages can be expressed as [11]:

$$V = -Y_{LL}^{-1} Y_{LG} E = R_V E \tag{2.84}$$

where R_V is defined as the bus reconstruction matrix.

Combining equations (2.82) and (2.84), we have:

$$z' = H'' R_V E + v' = H' E + v' \tag{2.85}$$

Because the internal generator complex voltage vector $E = [|E_1|\angle\delta_1, |E_2|\angle\delta_2, ..., |E_n|\angle\delta_n]^T$ is a function of the state vector x, the measurements can be modeled as

$$z = h(x) + v \tag{2.86}$$

where $z = [|z_1'|, \arg(z_1'), |z_2'|, \arg(z_2'), ..., |z_m'|, \arg(z_m')]^T$ is the $2m \times 1$ measurement vector that consists of amplitude and phase angle of each PMU measurement.

Furthermore, the measurement model can be expressed in the form of equation (2.2) after linearization

$$z = Hx + v \tag{2.87}$$

where H is the corresponding Jacobian matrix [25]:

$$H = \begin{pmatrix} \frac{\partial h_1}{\partial \delta_1} & 0 & \frac{\partial h_1}{\partial \delta_2} & 0 & \cdots \\ \frac{\partial h_2}{\partial \delta_1} & 0 & \frac{\partial h_2}{\partial \delta_2} & 0 & \cdots \\ \frac{\partial h_3}{\partial \delta_1} & 0 & \frac{\partial h_3}{\partial \delta_2} & 0 & \cdots \\ \frac{\partial h_4}{\partial \delta_1} & 0 & \frac{\partial h_4}{\partial \delta_2} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{2.88}$$

and v is the measurement noise with normal probability distribution $p(v) \sim N(0, R)$. The measurement noise covariance R can be determined experimentally, by testing the PMUs over time.

2.9.2 Optimal PMU Placement

Here the PMU placement evaluation using steady-state uncertainty analysis is very similar to the method described in section 2.8.2, except that we are dealing with dynamic states now. In order to compute P^{∞} for all desired points in the state space, we need to identify the parameter matrices A, Q, H and R. Note that in these matrices, only the generator rotor angles $\{\delta_1, \delta_2, ..., \delta_n\}$ are variables, so one only need to decide at which *interest* points $\{\bar{x}_1, \bar{x}_2, ...\bar{x}_p\}$ where $\bar{x}_i \in [0, 2\pi]^n$, to evaluate P^{∞}. Finally, our approach is described by Algorithm 3.

Algorithm 3: Optimal PMU Placement for Dynamic State Estimation

```
Build the optimization problem as shown in (2.64) based on the network topology;
if there is more than one solution to this problem then
    for each solution set of PMUs do
        for each n-dimensional point x̄_i ∈ {x̄_1, x̄_2, ...x̄_p} do
            Determine Δt for the PMUs;
            Evaluate A with Δt using (2.77);
            Evaluate Q with Δt using (2.78);
            Evaluate H using (2.88);
            Evaluate R;
            Compute Ψ using (2.11);
            Compute the n eigenvectors of Ψ using (2.12);
            Compute P_i^∞ using (2.13);
        Compute an aggregated uncertainty value from P_i^∞
    Apply the aggregated uncertainty values to user-defined criterion;
    Determine the optimal PMU placement strategy;
else
    Return the only possible solution set of PMUs;
```

2.9.3 Simulation Results

In this section, we first simulate the uncertainty analysis on an ideal multi-machine system with PMU installed on every bus. Then we applied our method to evaluate different optimal PMU placement strategies and find the best solution.

Estimation Uncertainty of a Fully PMU-installed System

We carried out our steady-state uncertainty analysis on the same small test system consisting of three machines in a looped network of nine buses as shown in Fig. 2.2. The dynamic state vector is $x = \{\delta_1, \omega_1, \delta_2, \omega_2, \delta_3, \omega_3\}$. To better illustrate the results, we represent the entire set of interest points $\hat{x} = \{\delta_1, \delta_2, \delta_3\} \in [0, 2\pi]^3$ by a 3D grid. For each point $\hat{x}_i$, we obtain a 6×6 steady-state error covariance matrix P_i^∞, and then aggregate the information in P_i^∞ to the following value for visualization

$$f(\hat{x}_i) = \frac{\sum_{j=1,3,5} \sqrt{P^\infty_{i_{[j,j]}}} + w_B \sum_{j=2,4,6} \sqrt{P^\infty_{i_{[j,j]}}} \cdot \Delta t}{3}, \tag{2.89}$$

$$f(\hat{x}_i) = \frac{\sum_{k=1,2,3} (\sqrt{P^\infty_{i_{[2k-1,2k-1]}}} + w_B \sqrt{P^\infty_{i_{[2k,2k]}}} \cdot \Delta t)}{3} \tag{2.90}$$

where we use Δt to convert speeds to changes in angle over the inter-measurement period. Because

$$\sqrt{P^\infty_{i_{[2k-1,2k-1]}}} + w_B \sqrt{P^\infty_{i_{[2k,2k]}}} \cdot \Delta t \tag{2.91}$$

represents the uncertainty in estimating the rotor angle of generator k over time Δt, the value $f(\hat{x}_i)$ is the average uncertainty of the three generators over Δt at this particular point $\hat{x}_i$.

In this extreme case, we assume all buses are equipped with PMUs. Fig. 2.15 illustrate the $f(\hat{x})$ values throughout the $[0, 2\pi]^3$ space. One can clearly tell a periodic pattern in the plot, because the Jacobian parameter matrices involve trigonometric functions of $\hat{x}$. The darker areas reflect lower uncertainty, and imply better expected estimation. This ideal simulation indicates the best estimation uncertainty level we can reach with a PMU placed on every bus.

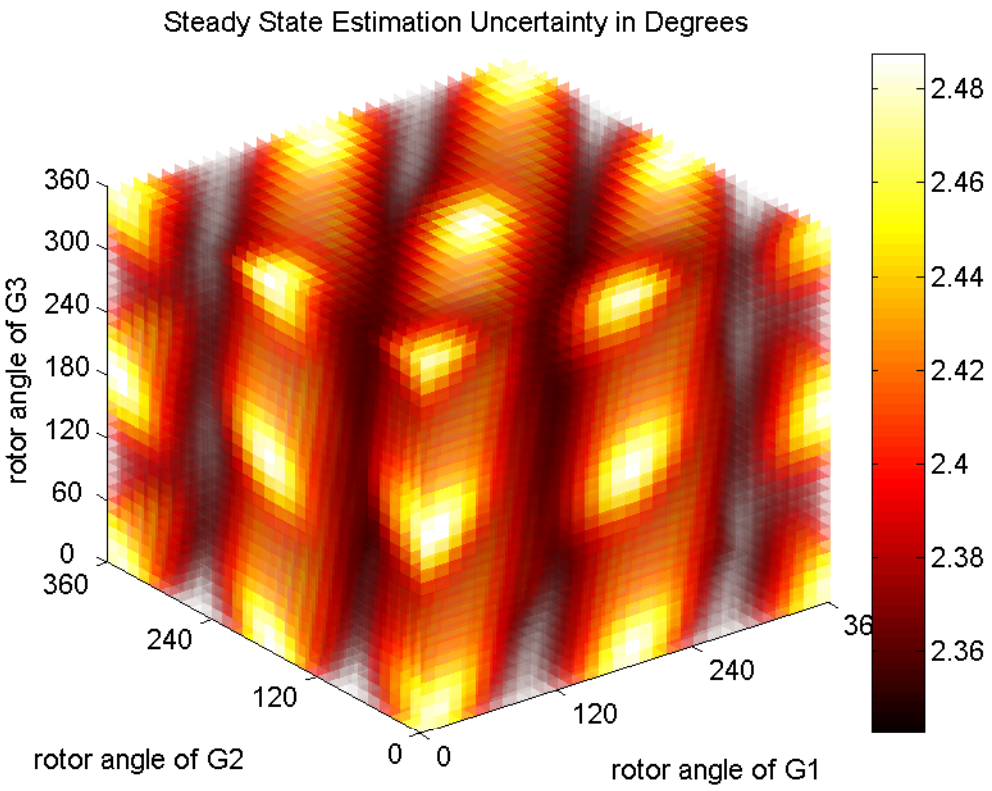


Figure 2.15: Steady-state estimation uncertainty versus rotor angles for the ideal case, i.e. when PMUs are installed on *all* buses.

Comparison Among Multiple Optimal Solutions

We used the same system in Fig. 2.2 as an example of finding the optimal PMU placement solution. After solving the optimization problem (2.64) with the linear integer programming method, we obtained four solutions, with PMUs installed at buses $\{1, 6, 8\}$, $\{2, 4, 6\}$, $\{3, 4, 8\}$ and $\{4, 6, 8\}$ respectively. According to previous research they can all be called "optimal" in the sense that they each make the entire system observable by using a minimum number of PMUs. To visualize the differences between these four PMU placement strategies via our approach, we provide the uncertainty plots for them in Fig. 2.16–2.19 using the same visualization method in Fig. 2.15, where darker color depicts lower uncertainty value.[4]

[4]Each of Fig. 2.16–2.19 uses the same colormap and plot range.

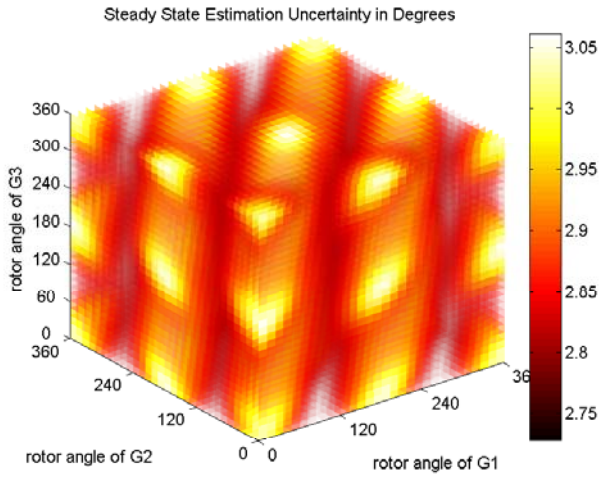


Figure 2.16: Steady-state estimation uncertainty versus rotor angles for PMUs installed at buses 1, 6 and 8.

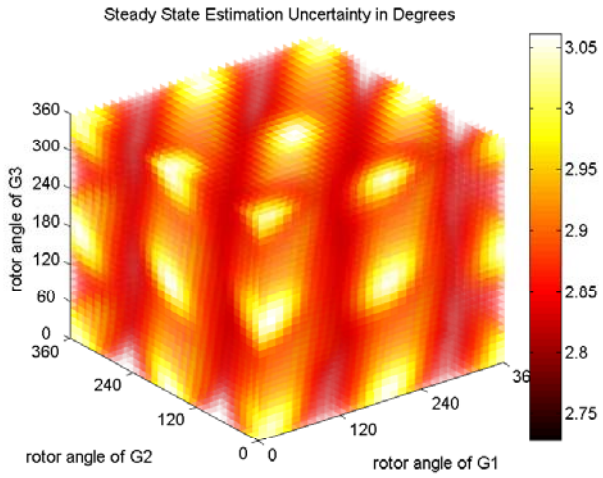


Figure 2.17: Steady-state estimation uncertainty versus rotor angles for PMUs installed at buses 2, 4 and 6.

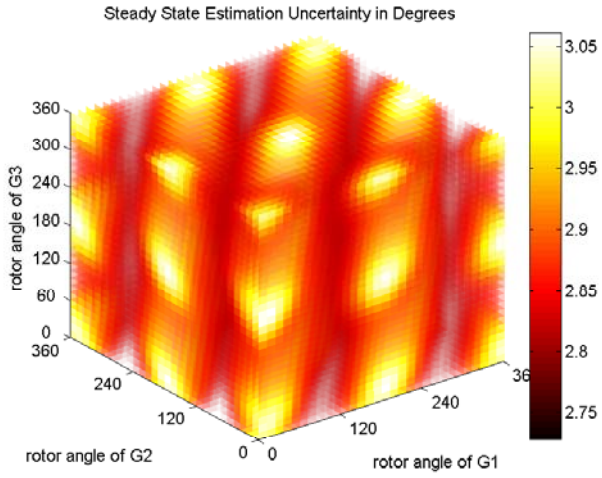


Figure 2.18: Steady-state estimation uncertainty versus rotor angles for PMUs installed at buses 3, 4 and 8.

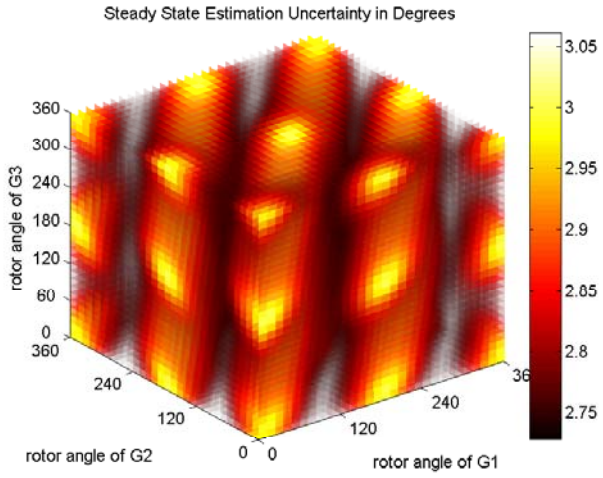


Figure 2.19: Steady-state estimation uncertainty versus rotor angles for PMUs installed at buses 4, 6 and 8.

Apparently, Fig. 2.19 has darker colors corresponding to better performance, as compared to Fig. 2.16–2.18. See also Fig. 2.20. One can choose different criteria depending on the circumstances. For instance, one could use the maximum, minimum, or mean of these aggregated uncertainties as depicted in Fig. 2.20, or even some weighted function of them. In our example, the scheme with PMUs placed at buses $\{4, 6, 8\}$ is more likely to be chosen by the designer,

because it has smaller max, min and mean uncertainties than the alternative scenarios. In practice, many factors may affect the final decision, *e.g.*, existing infrastructure, cost, *etc.* Our quantitative approach helps guide decision-making by providing concrete information about tradeoffs, which can not be offered by pure network topology analysis.

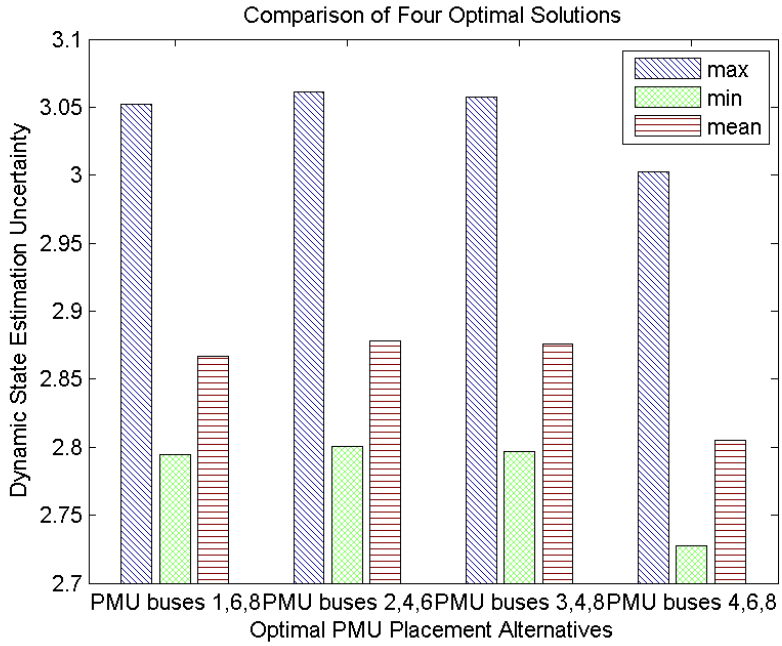


Figure 2.20: The comparison of the four “optimal" solutions based upon three criteria: *max*, *min* and *mean*

CHAPTER 3

FILTER COMPUTATION ADAPTATION

As measuring devices continue to develop towards higher accuracy and update frequency, the amount of measurement data associated with large complex systems becomes enormous. For example, in power systems that are usually large-scale and wide-area interconnected, the computational requirement for data processing has made real-time estimation a real challenge.

In this chapter we present a new method called Lower Dimensional Measurement-space (LoDiM) state estimation. It is an approach featuring a dynamic *measurement selection procedure*: a small measurement subset is dynamically and strategically chosen for each cycle to process. The smaller measurement-space yields smaller computational cost and hence lower reporting latency. In specific for power system dynamic state estimation, we present the Reduced Measurement-space Dynamic State Estimation (ReMeDySE) derived from LoDiM.

Although LoDiM is presented in the context of the Kalman filter, the associated *measurement selection procedure* is not filter-specific, i.e. it can be used with other state estimation methods such as particle and unscented filters. LoDiM, along with modern parallel computation techniques, can facilitate other large-scale, real-time and computation-intensive state tracking systems beyond the power systems.

3.1 Introduction

State estimation plays a basic and important role in modern industries. Particularly for power systems, it seeks to produce a reliable dynamic database for use in critical operational functions including real-time security monitoring, load-forecasting, economic despatch, and

load-frequency control.

There are typically four steps involved in power system state estimation procedures [21]:

1. Modeling. In this step, network topology is determined, and preliminary data is checked and calculated.
2. Observability determination. Various sensor measurements are classified as either *critical* or *redundant*. Critical measurements are necessary to achieve the system observability. Therefore, if some critical measurements are lost, pseudo measurements can be included. On the other hand, redundant measurements can be removed without affecting the system observability. A redundant measurement may belong to some *critical pairs* or *critical k-tuples*. A *critical pair* stands for two redundant measurements whose simultaneous removal from the measurement set will make the system unobservable. Similarly, a *critical k-tuple* contains a set of k redundant measurements, whose simultaneous removal will cause the system to become unobservable.
3. State estimation. Typically, a state estimator receives telemetered measurements from SCADA system in the time interval of several seconds. From the measurements and an assumed steady state system model, the state estimator generates a set of static state variables as a best estimate of the system conditions.
4. Detection and identification of gross measurement errors and/or errors in network structure.

Traditional SCADA systems collect RTU measurements. Updated every 2 to 4 seconds, they are too infrequent to capture the system dynamics, *e.g.* short circuits, line switching, load changes, *etc*. There has not been an issue historically, because traditional state estimators update the estimates every 3 to 5 minutes. Under the steady state system model assumption, none of the dynamic characteristics of a power system are reflected in the estimation. However, this changed with the introduction of modern phasor measurement technologies. High-speed and

high-accuracy synchrophasor measurements, combined with dynamic estimation techniques such as the Kalman filters, have motivated real-time power system state estimation.

Due to the tremendously increasing size and complexity of the interconnected power networks, tedious computations in real-time state estimation remain obstacles to overcome. Even with modern supercomputers, the massive data processing is still a very challenging task given the time and space (memory) complexities, despite significant work has been done aiming at accelerating the calculations via parallel implementations, such as OpenMP [26] and Global Arrays [27].

Here we introduce a different approach to reduce the computational burden. Previously we introduced an estimation approach called single-constraint-at-a-time (SCAAT) [14]. The idea behind the SCAAT approach is to use a Kalman filter to estimate a globally *observable* system using low dimensional measurements from potentially *unobservable* subspaces. In each filter cycle, SCAAT deals with much lower dimensional measurement data. The benefits include higher sampling rates, lower latencies, and improved accuracy. Although SCAAT was first developed for optoelectronic 3D tracking systems, it inspired us to design LoDiM [15, 16] and ReMedySE [17]. They are SCAAT-based algorithms for large-scale system state tracking, featuring dynamic measurement selection to reduce the computational burden.

In this Chapter, we first present a comprehensive study and proof of LoDiM, including the trade-offs and optimal number of measurements to be selected; and then we present ReMeDySE, which is derived and expanded from LoDiM to capture the *dynamic states* in power systems. Our method can significantly facilitate power system operations, and integrate with the powerful tool of *hierarchical, distributed* estimation as well as parallel computation techniques for more improvement.

3.2 Chapter Organization

The remainder of this chapter is organized as follows. Section 3.3 introduces the background. In Section 3.4, we review the traditional Kalman filter techniques, including Kalman filter (KF) and Extended Kalman filter (EKF). Section 3.5 gives a brief description of the SCAAT approach, which is an incremental tracking method with incomplete information. Section 3.6 presents our new approach, Lower Dimensional Measurement-space (LoDiM) state estimation, which performs Kalman filter upon dynamically selected measurements. It reduces the computation requirement dramatically, without sacrificing the tracking quality. We simulated the method on two power system models in different sizes to demonstrate and analyze the performance. Furthermore, we expand our approach to power system *dynamic state* estimation model, and developed Reduced Measurement-space Dynamic State Estimation (ReMeDySE) in Section 3.7, particularly for power system *dynamic state* estimation.

3.3 Background and Related Work

In modern power systems, measurement devices with higher accuracy and update rate are increasingly deployed, setting new requirements for real-time computations. Significant previous work has attempted to alleviate the computational load problem. One approach is to explore the potential computational power: the authors of [28] used Petri net (PN) theory to achieve the optimum utilization of processors, in a state estimator based on Kalman filter.

Another attempt is to reduce the computational complexity. For example, in [29] the authors proposed a method for systems that have more measurements than states. They described an equivalent state-space system in which the number of measurements equals the number of states. The system model is ideally assumed static and linear, and the preparation overhead was ignored in the complexity analysis.

[30] illustrated a measurement selection procedure using an Extended Kalman filter. As stated by the authors, an inherent limitation of the proposed method is that measurement selection is based entirely on the steady-state sensitivity matrix. According to this approach: (1) the actual information content of the candidate measurements under typical operating conditions is not considered; (2) the measurement rankings obtained are local and dependent on the steady state chosen as the base case; and (3) dynamic and non-linear effects are not considered.

[14] presented SCAAT, a Kalman-filter-based incremental tracking algorithm using incomplete information. It estimates a globally observable system using only measurements from locally unobservable systems. The underlying principle is that the single observations provide some information about the user's state, and thus can be used to incrementally improve a previous estimate. Based on SCAAT, another category called MCAAT, which processes multiple constraints at a time, was further derived.

This inspired us to propose LoDiM with sophisticated measurement selection procedure [15, 17]. It is also a Kalman-filter-based approach, with higher reporting rates and lower latency. Then we take the dynamic states of power systems into consideration and present ReMeDySE [17], an estimator that is suitable for reflecting power system dynamic characteristics.

In [31] we considered another approach to reduce the computational load of an ensemble Kalman filter. The main idea is to identify and process the most informative measurement subspace based on their capability to reduce the trace of the *a posteriori* error covariance. We construct a low-dimensional measurement set from linear combinations of the original measurement set, by computing a transformation matrix during each cycle. When the number of measurements is large and number of non-zero covariance eigenvalues is small, this algorithm can make an effective trade-off between computational complexity and estimation efficacy.

3.4 The Kalman Filter Techniques

Because of their efficiency, robustness and typical accuracy, Kalman filters [13, 19] have been employed in a wide range of applications from economic analysis to space navigation, and are especially popular in the areas of motion tracking systems. In power system applications, they have been extensively used to improve the computational performance of traditional state estimation process since the 1970's [32, 33]. Recent study [11, 34] also applied them to power system dynamic state models.

Since here we focus on discrete Kalman filters, we will omit the word "discrete" for the rest of this thesis. The Kalman Filter is a stochastic estimator for the instantaneous state of a dynamic system. The filter is very powerful in several aspects: it supports estimations of past, present, and even future states, and it can do so even when the precise nature of the modeled system is unknown. The Extended Kalman Filter was proposed later for non-linear models. In this section, we will give a brief overview of the discrete Kalman Filter (KF) and the Extended Kalman Filter (EKF). Readers could refer to [35] for a more detailed introduction.

3.4.1 The Kalman Filter

The Kalman filter, or KF, is a set of mathematical equations that provides an efficient computational (recursive) means to estimate the state of a process, in a way that minimizes the mean of the squared error. In order to apply the Kalman filter, we must first have a hypothetically true model of the system in hand.

Recall that in Section 2.4, we have introduced a generalized *linear* system model:

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1} \tag{3.1}$$

$$z_k = Hx_k + v_k \tag{3.2}$$

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:

Prediction Step

1. Project the state ahead:

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1} \tag{3.3}$$

   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, in this case, the *a posteriori* state estimate $\hat{x}_{k-1}$ at time step $k-1$, the state transition matrix A, and the control input (optional).

2. Project the error covariance ahead:

$$P_k^- = AP_{k-1}A^T + Q \tag{3.4}$$

   We define $e_k^- \equiv x_k - \hat{x}_k^-$ to be the *a priori* estimate error, so $P_k^- \equiv E[e_k^- e_k^{-T}]$ is the *a priori* estimate error covariance, where E denotes statistical expectation.

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:

Correction Step

1. Compute the Kalman gain:

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \tag{3.5}$$

   Where K is a $n \times m$ matrix called the Kalman gain matrix,

2. Update the error covariance:

$$P_k = (I - K_k H)P_k^- \tag{3.6}$$

$e_k \equiv x_k - \hat{x}_k$ is called the *a posteriori* estimate error, and $P_k \equiv E[e_k e_k^T]$ becomes the *a posteriori* estimate error covariance, where E denotes statistical expectation. P_k contains important information reflecting the filter's uncertainty in the estimation.

Under certain conditions when the system is controllable and observable, as $k \to \infty$, the limiting behavior of the Kalman filter can result in a bounded steady-state error covariance P [36], if we have

$$\lim_{k\to\infty} P_k^- = \lim_{k\to\infty} P_k = P < \infty \tag{3.7}$$

which means if the time is long enough, more iterations of the Kalmn filter will no longer affect the value of error covariance P, whether *a priori* or *a posteriori*. Thus by omitting the subscript "k" and superscript "$+$" [10], we have:

$$\begin{aligned}
P &= (I - KH)P \ (According\ to\ equation\ (3.6)) \\
&= P - KHP \\
&= P - PH^T(HPH^T + R)^{-1}HP \ (According\ to\ equation\ (3.5)) \\
&= A(P - PH^T(HPH^T + R)^{-1}HP)A^T + Q \ (According\ to\ equation\ (3.4)) \\
&= APA^T - APH^T(HPH^T + R)^{-1}HPA^T + Q
\end{aligned}$$

This error covariance expression is the Discrete Algebraic Riccati Equation (DARE), as seen in equation (2.10) from Section 2.5.

3. Update state estimate with measurement z_k:

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \tag{3.8}$$

We define $\hat{x}_k \in \mathcal{R}^n$ to be the *a posteriori* state estimate at time step k.

z_k is the actual measurement at time step k, and $H\hat{x}_k^-$ is the predicted measurement at time step k, The difference $(z_k - H\hat{x}_k^-)$ is conventionally called the *innovation*.

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 value while smaller values of R place more weight on the measured values.

There is an alternative formulation for equations (3.5) and (3.6) in correction steps 1 and 2, they are:

1′ Update the error covariance:

$$P_k = [(P_k^-)^{-1} + H^T R^{-1} H]^{-1} \tag{3.9}$$

Equation (3.9) is equivalent to (3.6) [36], because according to equation (3.6),

$$\begin{aligned}
(P_k)^{-1} &= ((I - K_k H)P_k^-)^{-1} \\
&= (P_k^-)^{-1}(I - K_k H)^{-1} \\
&= (P_k^-)^{-1}(I - P_k^- H^T (H P_k^- H^T + R)^{-1} H)^{-1} \\
&= (P_k^-)^{-1}(H^{-1}(H P_k^- H^T + R)(H P_k^- H^T + R)^{-1} H - P_k^- H^T (H P_k^- H^T + R)^{-1} H)^{-1} \\
&= (P_k^-)^{-1}((P_k^- H^T + H^{-1} R - P_k^- H^T)(H P_k^- H^T + R)^{-1} H)^{-1} \\
&= (P_k^-)^{-1}((H^{-1} R)(H P_k^- H^T + R)^{-1} H)^{-1} \\
&= (P_k^-)^{-1}(H^{-1}(H P_k^- H^T + R) R^{-1} H) \\
&= (P_k^-)^{-1}(P_k^- H^T R^{-1} H + I) \\
&= (P_k^-)^{-1} + H^T R^{-1} H
\end{aligned}$$

By applying the inverse operations on both sides, we have

$$P_k = [(P_k^-)^{-1} + H^T R^{-1} H]^{-1}$$

$2'$ Compute the Kalman gain:

$$K_k = P_k H^T R^{-1} \tag{3.10}$$

The equivalence of two Kalman gain formulations in equations (3.10) and (3.5) can be proved as follows [36]:

$$\begin{aligned}
K_k &= P_k^- H^T (H P_k^- H^T + R)^{-1} \\
&= P_k P_k^{-1} P_k^- H^T (H P_k^- H^T + R)^{-1} \\
&= P_k ((P_k^-)^{-1} + H^T R^{-1} H) P_k^- H^T (H P_k^- H^T + R)^{-1} \ (According\ to\ equation\ (3.9)) \\
&= P_k (H^T + H^T R^{-1} H P_k^- H^T)(H P_k^- H^T + R)^{-1} \\
&= P_k H^T R^{-1} (R + H P_k^- H^T)(R + H P_k^- H^T)^{-1} \\
&= P_k H^T R^{-1}
\end{aligned}$$

One advantage of equations (3.5) and (3.6) over equations (3.9) and (3.10) is that only one matrix inversion is required, and we do not need to worry about the singularity of matrix P_k^-, that is, whether $|P_k^-| = 0$.

The measurement update equations are responsible for the feedback, i.e. for incorporating a new measurement into the *a priori* estimate to obtain an improved *a posteriori* estimate.

To sum up, the time update (“predictor") projects the current state estimate ahead in time, and the measurement update (“corrector") adjusts the projected estimate by an actual measurement at that time. We can visually express the final estimation algorithm by a Kalman filter cycle as shown in Fig. 3.1: once the initial estimates for $\hat{x}_0$ and P_0 are fed into the filter, it evolves the state estimates in a recursive predictor-corrector fashion.

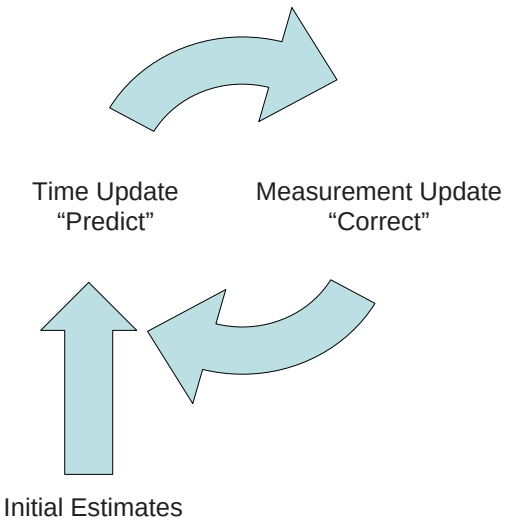


Figure 3.1: The operation of a discrete Kalman filter cycle

3.4.2 The Extended Kalman Filter

In practice, the process model and (or) the measurement relationship to the process is usually *non-linear*. A *non-linear* system model can be generalized as:

$$x_k = a(x_{k-1}, u_{k-1}, w_{k-1}) \tag{3.11}$$

$$z_k = h(x_k, v_k). \tag{3.12}$$

For instance, we can only express the power system dynamic states variables in non-linear differential-algebraic process and measurement equations, representing the power system dynamic behavior:

$$\frac{dx}{dt} = f(x, y) \tag{3.13}$$

$$z_k = h(x_k, v_k)$$

In generalized non-linear system form, the equations become:

$$x_k = x_{k-1} + f(\hat{x}_{k-1}, y_{k-1})\Delta t \tag{3.14}$$

$$z_k = h(x_k, v_k)$$

Our objective is to estimate these dynamic states, and EKF was proposed to deal with such problems.

We could approximate the state and measurement vectors in (3.11) and (3.12) by:

$$\tilde{x}_k = a(\hat{x}_{k-1}, u_{k-1}, 0) \tag{3.15}$$

$$\tilde{z}_k = h(\tilde{x}_k, 0) \tag{3.16}$$

where $\hat{x}_{k-1}$ is an a *posteriori* estimate of the state at step $k-1$.

Then by linearizing the estimation about $\tilde{x}_k$ and $\tilde{x}_k$, we obtain a new set of *linear* difference equations and measurement equation for the *error process*:

$$(x_k - \tilde{x}_k) \approx A_k(x_{k-1} - \hat{x}_{k-1}) + W_k w_{k-1} \tag{3.17}$$

$$(z_k - \tilde{z}_k) \approx H_k(x_k - \tilde{x}_k) + V_k v_k \tag{3.18}$$

where

x_k and z_k are the true state and measurement vectors at step k,

$\tilde{x}_k$ and $\tilde{z}_k$ are the approximated state and measurement vectors at step k.

$\hat{x}_{k-1}$ is an a *posteriori* estimate of the state vector at step $k-1$.

$A_k = \frac{\partial a}{\partial x}|_{\hat{x}_{k-1}, u_{k-1}, 0}$ is the Jacobian matrix of partial derivatives of a with respect to x,

$W_k = \frac{\partial a}{\partial w}|_{\hat{x}_{k-1}, u_{k-1}, 0}$ is the Jacobian matrix of partial derivatives of a with respect to w,

$H_k = \frac{\partial h}{\partial x}|_{\tilde{x}_k, 0}$ is the Jacobian matrix of partial derivatives of h with respect to x,

$V_k = \frac{\partial h}{\partial v}|_{\tilde{x}_k,0}$ is the Jacobian matrix of partial derivatives of h with respect to v.

The Jacobian matrices A, W, H and V have time step subscript k because they may vary at each time step.

For simplicity, we denote the *prediction error* as

$$\tilde{e}_{x_k} = x_k - \tilde{x}_k, \tag{3.19}$$

and *measurement residual* as

$$\tilde{e}_{z_k} = z_k - \tilde{z}_k. \tag{3.20}$$

We define

$$\varepsilon_k = W_k w_{k-1} \tag{3.21}$$

$$\eta_k = V_k v_k \tag{3.22}$$

as two new mutually independent random variables with normal probability distributions:

$$p(\varepsilon_k) \sim N(0, W_k Q_{k-1} W_k^T) \tag{3.23}$$

$$p(\eta_k) \sim N(0, V_k R_k V_k^T) \tag{3.24}$$

Thus we write the error process 3.17 as

$$\tilde{e}_{x_k} \approx A_k(x_{k-1} - \hat{x}_{k-1}) + \varepsilon_k \tag{3.25}$$

$$\tilde{e}_{z_k} \approx H_k(x_k - \tilde{x}_k) + \eta_k \tag{3.26}$$

for short. Note that this is a linearized system for us to estimate conveniently by applying Kalman Filter.

An Extended Kalman filter, or EKF, is nothing but a Kalman filter that linearized about the current mean and covariance. With some slight modifications to the KF equations, the EKF

has the following steps:

$$Prediction: \begin{cases} \hat{x}_k^- = a(\hat{x}_{k-1}, u_{k-1}, 0) \\ P_k^- = A_k P_{k-1} A_k^T + W_k Q_{k-1} W_k^T \end{cases} \tag{3.27}$$

$$Correction: \begin{cases} K_k = P_k^- H_k^T (H_k P_k^- H_k^T + V_k R_k V_k^T)^{-1} \\ \hat{x}_k = \hat{x}_k^- + K_k(z_k - h(\hat{x}_k^-, 0)) \\ P_k = (I - K_k H_k) P_k^- \end{cases} \tag{3.28}$$

3.5 The SCAAT Method

[37] has argued that there is a direct relationship between the *complexity* of the estimation algorithm, the corresponding *speed* (execution time per estimation cycle), and the *change* in system state between estimation cycles.

This relationship can be clearly visualized with the *Complexity-Speed-Change* cycle in Fig. 3.2 [1]. The notes on the outside of the circle represent a vicious cycle: as the algorithmic complexity increases, the corresponding execution time increases, which allows for significant (and usually non-linear) system state changes between estimation cycles. This in turn implies the need for a more complex dynamic process model and estimation algorithm.

The *single-constraint-at-a-time* (SCAAT) approach, first proposed and described in [14], is a Kalman-filter-based incremental tracking algorithm with an attempt to reverse this cycle. Instead of waiting to form and process a complete collection of measurements (constraints) like the traditional Kalman filter would do, the SCAAT method intentionally fuses a single measurement per estimate cycle. Indeed, single measurements underconstrain the mathematical solution, however, they still can provide *some* information about the system state, and thus *incrementally* improve a previous estimate. By applying SCAAT, the algorithmic complexity is drastically reduced, which reduces the corresponding execution time, and hence reduces the amount of state changes between estimation cycles. To deal with limited state changes, we

only need to use a simple dynamic process model in the estimation algorithm, which further simplifies the computations. The SCAAT has formed a cycle as denoted on the inside of the circle.

Figure 3.2: Complexity-Speed-Change Cycle [1]

3.6 My Method: Lower Dimensional Measurement-space (LoDiM) State Estimation

In a common scenario where the processor running the estimation algorithm needs to execute some additional tasks (*e.g.*, simulation of an urgent contingency), we will have to release some computational resources for the tasks. Ideally, we want the estimation process to continue with reduced computational cost, but without sacrificing the reporting rate. LoDiM is thus proposed for this purpose.

The LoDiM method employs a Kalman filter that incorporates lower dimensional (subspace) measurements in each cycle, to adapt to limited computational resources. It employs a special measurement selection procedure to strategically reduce the measurement-space di-

mension. As a result, it yields a much smaller computational load, lower latency, and most importantly, reliable performance.

In this section, we will first give a brief description of the design principle behind LoDiM, and then discuss our dynamic measurement selection method and the structure of LoDiM.

3.6.1 Principle of Design

As one would expect, performing dynamic state estimation with KF/EKF (or any filter) is a rather computationally intensive process. For small systems, the computation may be fast enough for real time control applications; However the growth of three factors will significantly affect the computational effort: the size of the system, the complexity of model components, and the number of measurements to be processed. Table 3.1 [38] has listed the upper bound on computational requirements of the Kalman filter, measured by *FLoating-point OPeration* (FLOP) counts. FLOP takes into account matrix multiplications and additions. Computational cost in the table is considered upper bound because the implementation steps are not yet subject to various optimizations, such as storage saving, data reuse, *etc*. More detailed discussions can be found in [38, 39].

Table 3.1: Kalman Filter Computational Cost Upper Bound

Step	Variable	Computation Breakdown	FLOP Breakdown	FLOP counts
1	$\hat{x}_k^-$	$A\hat{x}_{k-1}$	$2n^2 - n$	$2n^2 - n$
2	P_k^-	$P_{k-1}A^T$	$2n^3 - n^2$	$4n^3 - n^2$
		$AP_{k-1}A^T$	$2n^3 - n^2$	
		$AP_{k-1}A^T + Q$	n^2	
3	K_k	HP_k^-	$2n^2m - nm$	$2m^3 + 4nm^2 + 2n^2m - 2nm$
		$HP_k^-H^T$	$2nm^2 - m^2$	
		$HP_k^-H^T + R$	m^2	
		$(HP_k^-H^T + R)^{-1}$	$2m^3$	
		$(P_k^-H^T)(HP_k^-H^T + R)^{-1}$	$2nm^2 - nm$	
4	$\hat{x}_k$	$H\hat{x}_k^-$	$2nm - m$	$4nm$
		$z_k - H\hat{x}_k^-$	m	
		$K_k(z_k - H\hat{x}_k^-)$	$2nm - n$	
		$\hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$	n	
5	P_k	$K_k(HP_k^-)$	$2nm^2 - m^2$	$n^2 + 2nm^2 - m^2$
		$P_k^- - K_k(HP_k^-)$	n^2	

Adding up the FLOP counts at each step in Table 3.1 yields an estimate of the computational cost for a Kalman filter cycle, which is at most:

$$\begin{aligned}&(2n^2 - n) + (4n^3 - n^2) + (2m^3 + 4nm^2 + 2n^2m - 2nm) + (4nm) + (n^2 + 2nm^2 - m^2)\\ &= 2m^3 + 6nm^2 + 2n^2m + 4n^3 + 2n^2 + 2nm - m^2 - n \end{aligned} \tag{3.29}$$

We have noticed the expensive cost of calculating the Kalman gain K_k in step 3, because it involves the inversion of a $m \times m$ matrix $(HP_k^-H^T + R)$, with complexity of $O(m^3)$. This makes the computation intractable when the number of measurements m is too large for the available computational resource to handle promptly, which unfortunately, can happen at any moment.

On the other hand, if we can reduce the measurement-space dimension to adapt to limited

computational resource, *i.e.*, use only a subset of the available measurements to update (perhaps a subset of) the states during each Kalman filter cycle, the computation cost could be reduced dramatically while still maintaining observability over time *and* improving accuracy per the SCAAT approach [14]. Specifically we propose using a subset of m_σ measurements, where $1 \leqslant m_\sigma \ll m$. The question is *which* subset of m_σ measurements to be used in each cycle. Previous work suggests pre-determined measurement subsets, however the dynamic nature of today's power systems requires more flexibility. By performing a principal component analysis (PCA) on the error covariance P, we can determine the state subspace to be updated most urgently (*e.g.* with larger estimation uncertainty than others).

Because covariance matrices are always symmetric and positive semidefinite, they have several important properties. Before the measurement update ("Correction" step) begins, let us consider the PCA of the *a priori* error covariance matrix $P^- = U \cdot D \cdot U^T$:

1. There exists an orthonormal basis U ($UU^T = U^TU = I$, where I is the identity matrix), whose columns are eigenvectors of P^-, such that the error covariance matrix expressed in this basis is diagonal. The axes of this new basis are called *Principal Components* of P^-.
2. As the off-diagonal elements of this new diagonal covariance matrix D are zeros, the new variables defined by this new basis (the projections of the *a priori* estimate error $e^- = x - \hat{x}^-$ on the Principal Components) are uncorrelated.
3. The diagonal elements of this new matrix D are the eigenvalues of P^-. So the variances of the projections of error e^- on the Principal Components are equal to the corresponding eigenvalues of P^-.
4. The eigenvalues in D are decreasingly ordered. The mth eigenvalue corresponds to the mth eigenvector.

The principal components that correspond to the largest elements of D, indicate the axes

in the state space that have the largest estimation uncertainties. Our strategy is to target these uncertainties first, thus we need to find the set of measurements that can reduce these uncertainties most efficiently. For example, in a 3D tracking application example, if we use several cameras to estimate the location of a certain object, and we noticed the uncertainty is growing rapidly in one direction, then in the next cycle we would ideally use a camera which is looking in an orthogonal direction.

3.6.2 Measurement Selection Procedure

Similar to SCAAT, LoDiM constrains the unknowns *over time* and refines the estimation continually. Nonetheless, it is the measurement selection that differentiates LoDiM from SCAAT. In LoDiM, we select the measurements that help our estimate the most (i.e. reduce estimation uncertainty most effectively) during each iteration.

After the time update ("Prediction" step) of each Kalman filter iteration, we have the $n \times n$ *a priori* error covariance matrix

$$P^{-} = U \cdot D \cdot U^{T} \tag{3.30}$$

where D is the diagonal matrix consisting of the eigenvalues of P^{-} in decreasing order, and U is the orthonormal basis whose columns are the corresponding eigenvectors. Note that the full PCA can also be a time-consuming process, especially if the state space is large. For this reason, for now we only investigate the first eigenvector (dominant eigenvector) u_1 in U, which represents the directions that we are most uncertain about in the state space.

The dominant eigenpair (the largest eigenvalue and the corresponding eigenvector u_1) can be conveniently computed by existing algorithms such as the well-known *power method* [40]. Start with a unit vector $q^{(0)} \in \mathcal{R}^n$, which is an approximation to the dominant eigenvector, the

power method iteratively produces a vector sequence $q^{(k)}$ as follows:

$$p^{(k)} = P^{-}q^{(k-1)} \tag{3.31}$$

$$q^{(k)} = p^{(k)}/\|p^{(k)}\|_2 \tag{3.32}$$

$$u^k = [q^{(k)}]^H P^{-} q^{(k)}, k = 1, 2, ... \tag{3.33}$$

Except for special starting points, the iterations converge to the dominant eigenvector u_1 quite efficiently. Table 3.2 has listed the upper bound on computational requirements of the power method at every iteration, measured by FLOP counts.

Table 3.2: Power Method Computational Cost Upper Bound

Step	Variable	Computation Breakdown	FLOP Breakdown	FLOP counts
1	$p^{(k)}$	$P^{-}q^{(k-1)}$	$2n^2 - n$	$2n^2 - n$
2	$q^{(k)}$	$[p^{(k)}]^H p^{(k)}$ $\sqrt{[p^{(k)}]^H p^{(k)}}$ $p^{(k)}/\sqrt{[p^{(k)}]^H p^{(k)}}$	$2n - 1$ 1 n	$3n$
3	u^k	$[q^{(k)}]^H P^{-}$ $[q^{(k)}]^H P^{-} q^{(k)}$	$2n^2 - n$ $2n - 1$	$2n^2 + n - 1$

Adding up the FLOP counts at each step in Table 3.2 yields an estimate of the computational cost for a power method iteration, which is at most:

$$(2n^2 - n) + (3n) + (2n^2 + n - 1) = 4n^2 + 3n - 1 \tag{3.34}$$

Because the power method is an iterative method, one has to stop at some point. Typically one carries on the process until a convergence criterion of the eigenvalue is satisfied or, alternatively, until a given maximum number of iterations is reached. In our experiments, we follow the latter approach for each measurement selection procedure (*e.g.*, $k_{max} = 10$), such that we can save the computational cost and include the procedure with each Kalman cycle.

Now consider the measurements. We can rewrite the measurement equation as

$$\begin{aligned} z &= Hx + v \\ &= HUU^T x + v \\ &= (HU)x' + v \\ &= H'x' + v \end{aligned} \tag{3.35}$$

Where $x' = U^T x$ is the new state vector defined by the new basis U, H' is the corresponding $m \times n$ new measurement Jacobian matrix, and v is the unchanged measurement noise vector with $p(v) \sim N(0, R)$. Here we assume R to be a $m \times m$ diagonal covariance matrix, ***i.e.***, the measurement noise is uncorrelated.

Because the basis U is composed of unit vectors, H'_{ij} can also be considered as the magnitude of the projection (i.e. the scalar projection) of the ith measurement direction vector in the direction of the jth basis in U. We now have

$$Hu_1 = [H'_{11} \quad H'_{21} \quad H'_{31} \ldots H'_{m1}]^T \tag{3.36}$$

Intuitively, for the same basis, say u_1, the larger H'_{i1} is, the better the corresponding ith measurement could reduce the uncertainty in this basis direction. However we should also keep in mind that different measurements have different amounts of noise. Thus for u_1 we create a "ranking" vector r_1 from H' and R using the following adjustment:

$$r_1 = [\frac{H'^2_{11}}{R_{11}} \quad \frac{H'^2_{21}}{R_{22}} \quad \frac{H'^2_{31}}{R_{33}} \ldots \frac{H'^2_{m1}}{R_{mm}}]^T \tag{3.37}$$

where R is the measurement noise covariance matrix. The $m \times 1$ vector r_1 evaluates the uncertainty calibration abilities of each measurement regarding the most significant uncertainty component, scaled by the corresponding measurement noise level. Now we are going to prove the following:

Lemma 1 *The larger $\frac{H'^2_{i1}}{R_{ii}}$ is, the more effectively the i^{th} measurement can reduce the uncertainty along direction u_1.*

Proof. In [36], it has been shown that the error covariance update in the measurement update ("Correction" step)

$$\begin{aligned} P &= (I - KH)P^- \\ &= (I - P^-H^T(HP^-H^T + R)^{-1}H)P^- \\ &= P^- - P^-H^T(HP^-H^T + R)^{-1}HP^- \end{aligned} \tag{3.38}$$

is equivalent to

$$P^{-1} = (P^-)^{-1} + H^TR^{-1}H \tag{3.39}$$

The inverse of the error covariance, P^{-1}, is often called the *information matrix*. According to (3.30), we have

$$\begin{aligned} P^{-1} &= (UDU^T)^{-1} + H^TR^{-1}H \\ &= U^{-T}D^{-1}U^{-1} + U^{-T}U^TH^TR^{-1}HUU^{-1} \\ &= U^{-T}[D^{-1} + (HU)^TR^{-1}(HU)]U^{-1} \\ &= U^{-T}(D^{-1} + H'^TR^{-1}H')U^{-1} \end{aligned} \tag{3.40}$$

Let us denote the matrix $(D^{-1} + H'^TR^{-1}H')$ in equation (3.40) by Σ, then Σ_{11} is the "information" of the new state variable in direction u_1, which we were most uncertain about. This information value is increased/improved from D_{11}^{-1} to Σ_{11} by

$$(Hu_1)^TR^{-1}(Hu_1) = \sum_{i=1}^{m} \frac{H_{i1}'^2}{R_{ii}} \tag{3.41}$$

So if we are only willing to incorporate m_σ measurements (instead of the full m measurements) into the measurement update equation, the ones with the largest $\frac{H_{i1}'^2}{R_{ii}}$ values would be our choice.

■

We can easily locate the m_σ measurements with the largest values among these m elements in vector r_1 (3.37). Thus when our budget of computation time and memory space is tight, we could use only m_σ measurements in the next step but still achieve stable estimation results.

Using this approach to reduce state estimation uncertainty is similar to fighting the Hydra in Greek mythology: if we are not able to destroy all the “heads" at once, at least we can aim at and cut off the most threatening “head" during each round.

3.6.3 LoDiM Architecture

For computational efficiency in large-scale Kalman-filter-based state estimation, we propose a better realization for our new method, LoDiM, with a parallel task structure.

To generalize our approach, let us consider a non-linear system described by equations (2.5) and (2.6), with a large measurement-space. LoDiM has its main state estimation cycle, which is similar to SCAAT algorithm, running on the foreground:

1. Compute the time Δ_t since the previous estimate.
2. Predict the state and error covariance. Share the predicted error covariance P^- with the background process.

$$\begin{cases} \hat{x}^- = a_{\Delta_t}(\hat{x}_{t-\Delta_t}, 0) \\ P^- = A_{\Delta_t} P_{t-\Delta_t} A_{\Delta_t}^T + Q_{\Delta_t} \end{cases} \tag{3.42}$$

3. If this is the first cycle, choose m_σ measurements (which is a much smaller measurement subset) randomly; otherwise, choose the m_σ measurements nominated by the background process. Predict the measurement and compute the corresponding Jacobian.

$$\begin{cases} \hat{z} = h_\sigma(\hat{x}_t^-, 0) \\ H = H_\sigma(\hat{x}_t^-, 0) \end{cases} \tag{3.43}$$

4. Compute the *Kalman gain*.

$$K = P^- H^T (H P_k^- H^T + R_{\sigma,t})^{-1} \tag{3.44}$$

5. Correct the predicted state estimate and error covariance from (3.42) using the actual sensor measurement $z_{\sigma,t}$.

$$\begin{cases} \hat{x}_t = \hat{x}^- + K(z_{\sigma,t} - \hat{z}) \\ P_t = (I - KH)P^- \end{cases} \tag{3.45}$$

Meanwhile, LoDiM has its auxiliary measurement selection process as described in subsection 3.6.2, running on the background:

1. Compute the principal component u_1 of the error covariance P^- predicted on the foreground.
2. Compute Hu_1 according to (3.36) and the *ranking* vector r_1 according to (3.37).

$$r_1 = (Hu_1) .* (Hu_1)./diag(R) \tag{3.46}$$

   where $.*$ and $./$ denote the element-by-element operations.

3. Select the m_σ measurements corresponding to the m_σ largest elements in r_1, to be used by the foreground process.

Overall, the LoDiM algorithm process can be expressed by the flow chart in Figure 3.3.

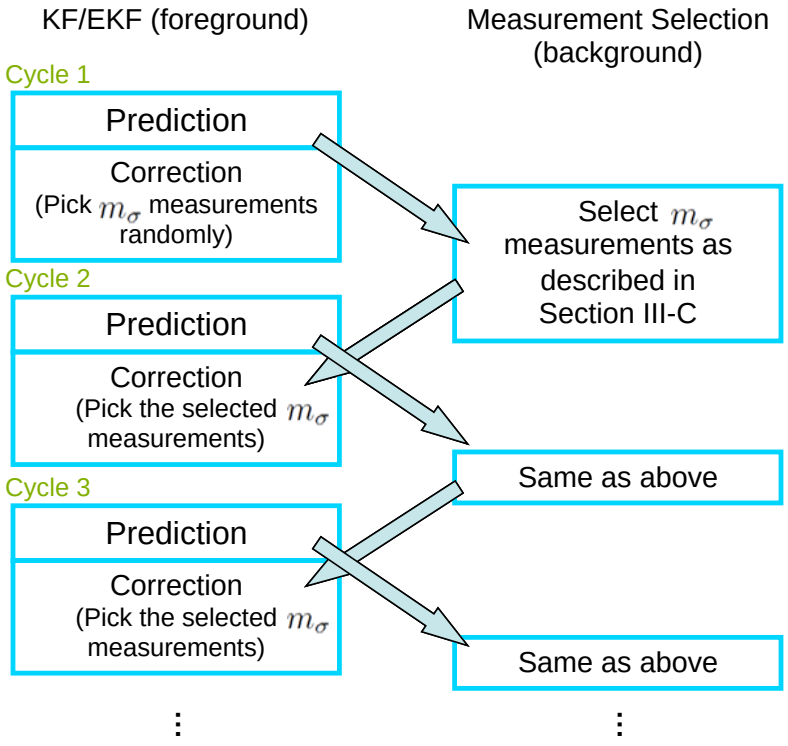


Figure 3.3: The LoDiM algorithm flow chart.

In the future, this algorithm is subject to further improvement by exploiting more task parallelism and data parallelism using modern parallel computation techniques and hardware.

Another future research topic is that for large systems with exploited sparsity, it makes sense to integrate LoDiM into *hierarchical* estimation methods [41, 42, 43, 44]. Take large-scale wide-area power system state estimation as an example, a rational approach could be a hybrid method which includes the following three major steps:

1. *Local State Estimation*. A Large-scale wide-area power system can be divided into smaller non-overlapping subsystems. Local estimation can be done simultaneously and independently for all subsystems using the LoDiM algorithm.
2. *Coordination Vector Estimation*. Each subsystem is connected to its neighbors through tie-lines, which end at boundary buses belonging to these subsystems. This step requires

input of tie-line measurements from the synchronized PMUs and part of the voltage measurements from each subsystem (only those of the boundary buses), while the output of estimated coordination phase angle vector and the local state estimation from all subsystems are accessible for the next step.

3. *Global Coordination.* All local estimates are coordinated into a unified system-wide state estimation.

3.6.4 Simulation Results

LoDiM Performance on a Smaller System

In this section, we investigate the proposed LoDiM state estimation performance using the 3-machine 9-bus system as shown in Fig. 2.2 from Section 2.8.3. We simulate an emergency event: a three-phase fault at bus 4 that happens at $t = 1$, and then cleared in 0.15 seconds. This event represents a large disturbance emergency.

For this nine-bus system, we simulate the 17 state variables, which are the bus voltage magnitudes and phase angles, recorded as the “true" system states. There are 54 simulated Phasor Measurement Unit (PMU) measurements, including both voltage phasor and current phasor measurements. They are updated every $1/30$ second, and combined with 5% random noise. The algorithm is simulated to run on a processor with 10 megaFLOPS (10^7 *FLoating-point Operation Per Second*) of computing power available.

Without loss of generality, we visualize the state estimation results for bus 2. Fig. 3.4 depicts the voltage magnitude tracking result of bus 2 during the 10 seconds from $t = 0$ to $t = 10$, using the conventional Kalman filter. The black solid line plots the true state, while the red dot-dot line plots the estimated state. The regular EKF dynamic state estimation uses the entire measurement set as input at each update step, resulting in lower estimation rate.

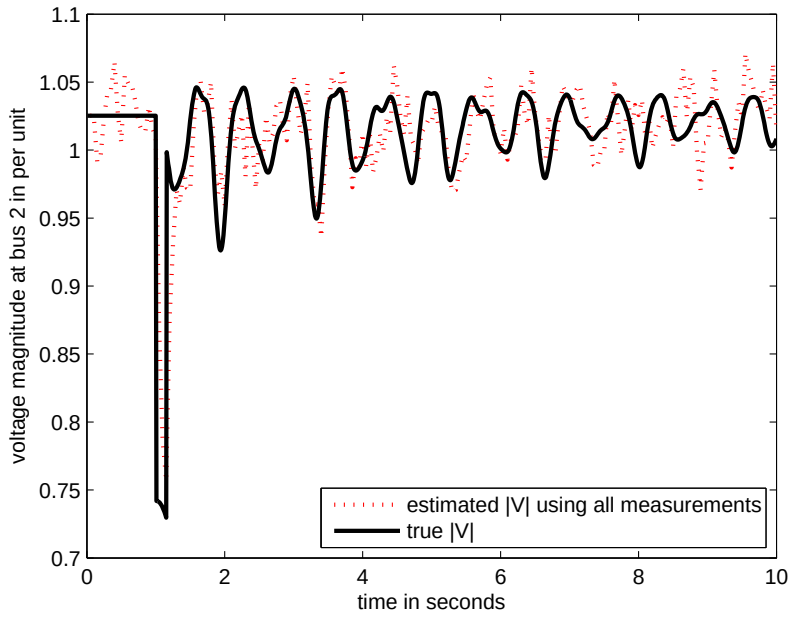


Figure 3.4: Bus 2 Voltage magnitude estimation using conventional Kalman filter

Next, we implement a reduced measurement-space state estimation with a naive approach: a small subset of the measurements (6 measurements in our experiment) are chosen randomly during each shorter estimation cycle. Fig. 3.5 shows the tracking results using a green dash-dot line.

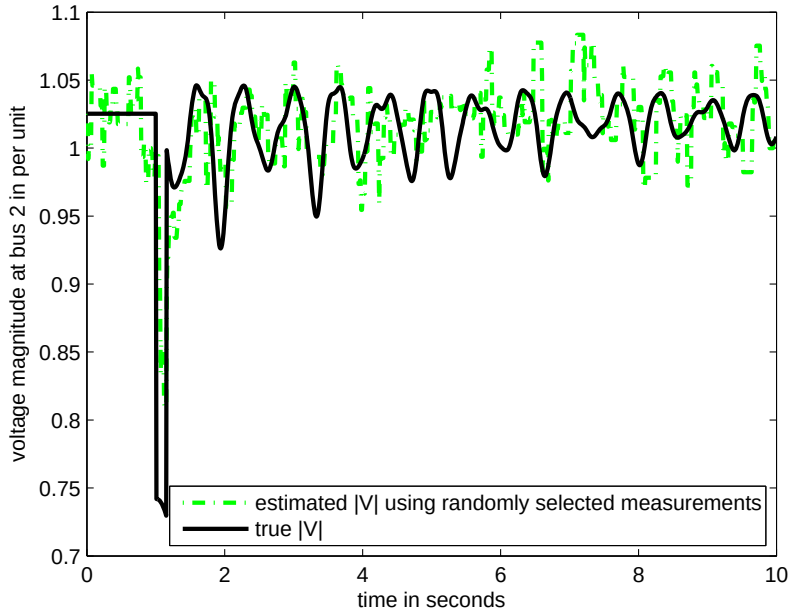


Figure 3.5: Bus 2 Voltage magnitude estimation using randomly chosen measurements

Our LoDiM state estimation method use the same number of dynamically selected measurements (6 measurements also) as input during each short estimation cycle. However, the small measurement subset is selected by the *measurement selection procedure* described in section 3.6.2. Fig. 3.6 shows the voltage magnitude tracking result of the same bus during the same period, using LoDiM method. The black solid line still represents the true state, while the blue dash-dash line plots the estimated state.

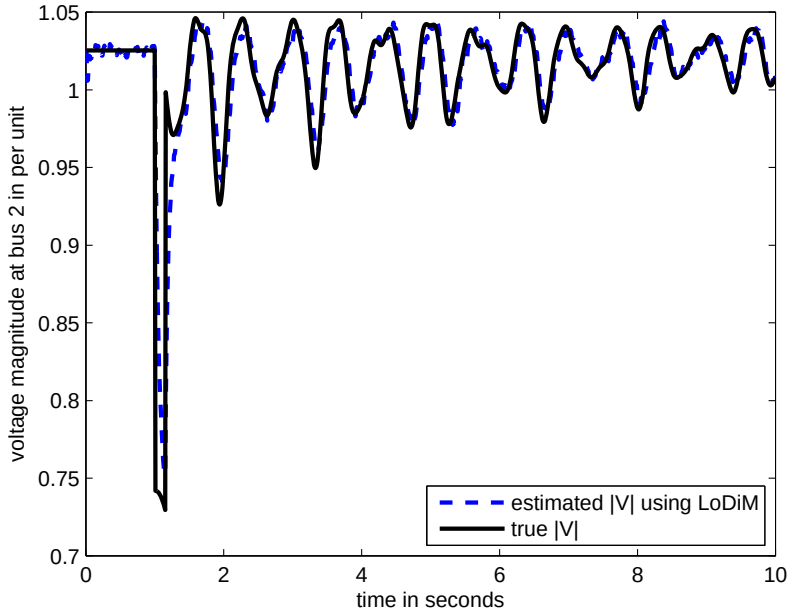


Figure 3.6: Bus 2 Voltage magnitude estimation using LoDiM

Finally, let us take a closer look at the performances of different state estimation methods described above. Within a short time period from $t = 4$ to $t = 6$, the true state and the estimated states using these three approaches are plotted in Fig. 3.7 for a better comparison. This figure illustrates how our proposed LoDiM method, appearing both smooth and accurate, outperforms the other two methods in this simulation.

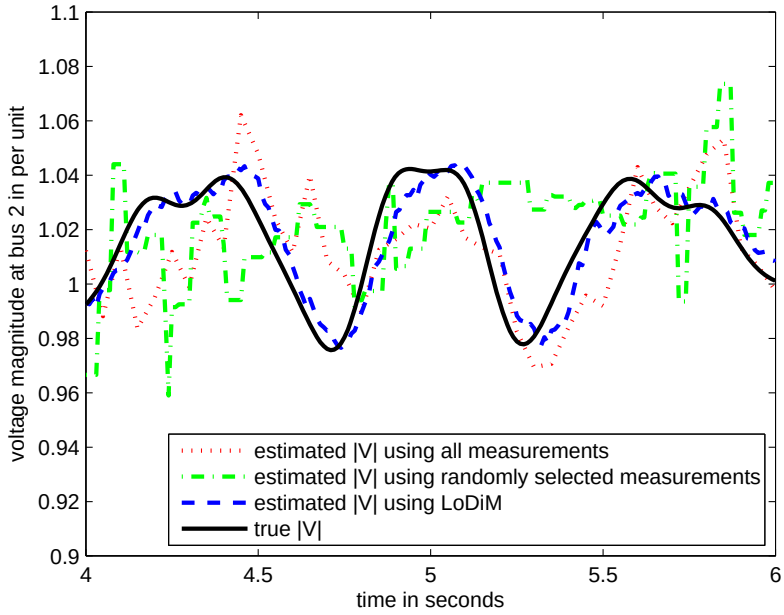


Figure 3.7: The performance comparison of state estimation with full measurement-space (in red), with randomly selected measurements (in green) and with LoDiM (in blue), from $t = 4$ to $t = 6$

LoDiM Performance on a Larger System

We have also simulated the proposed LoDiM state estimation on the 16-generator 68-bus system as shown in Fig. 2.6 from Section 2.8.3, which represents the New England/New York interconnected system. A three-phase fault occurring at time $t = 1.1$ is simulated on bus 29, and then cleared in 0.05 seconds.

There are 135 state variables to be estimated in this 68-bus system. We have simulated 468 PMU measurements combined with 5% random noise, including both voltage phasor and current phasor measurements, at a sampling rate of 30 samples/second. The algorithm is simulated to run on a processor with 1 gigaFLOPS (10^9 *FLoating-point Operation Per Second*) available computing power.

First, we visualize the state estimate results for bus 60, a generator bus. Fig. 3.8, Fig. 3.9

and Fig. 3.10 have illustrated the voltage magnitude tracking results of bus 60 during the 10 seconds from $t = 0$ to $t = 10$, using the following different methods:

- a conventional Kalman filter (which takes the entire measurement set as input, resulting in lower estimation rate);
- a Kalman filter with randomly chosen measurements (which takes a smaller random set of 70 measurements as input, during each shorter estimation cycle);
- the LoDiM state estimation method (which also takes 70 measurements as input during each short estimation cycle; nevertheless, the small measurement subset is dynamically selected by the *measurement selection procedure* described in section 3.6.2).

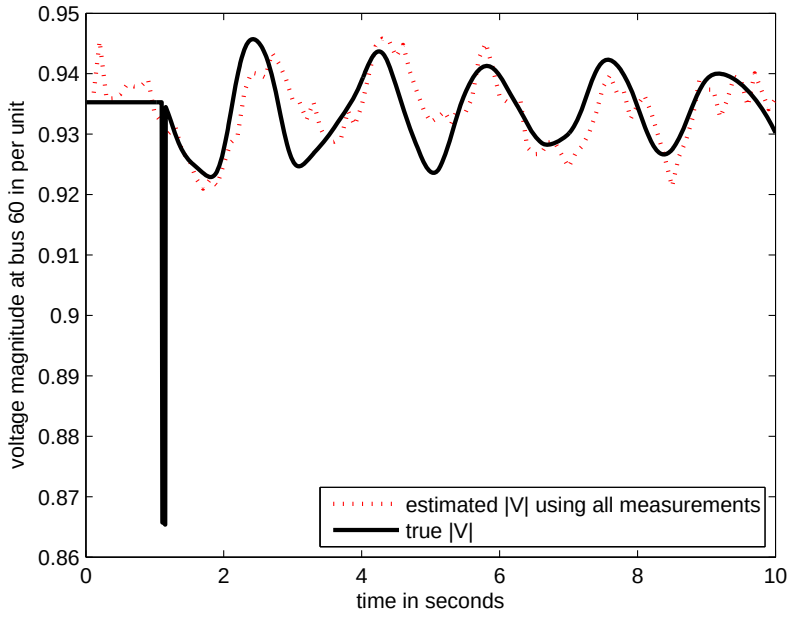


Figure 3.8: Bus 60 Voltage magnitude estimation using conventional Kalman filter

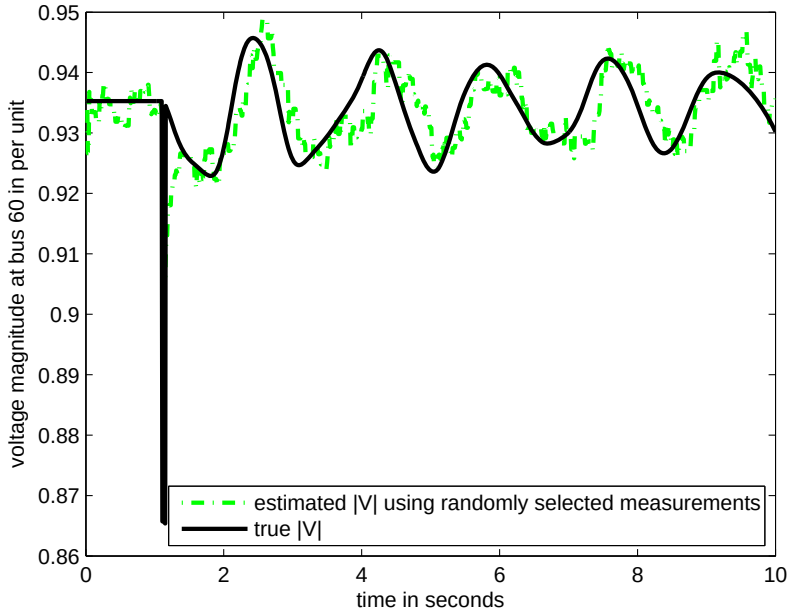


Figure 3.9: Bus 60 Voltage magnitude estimation using randomly chosen measurements

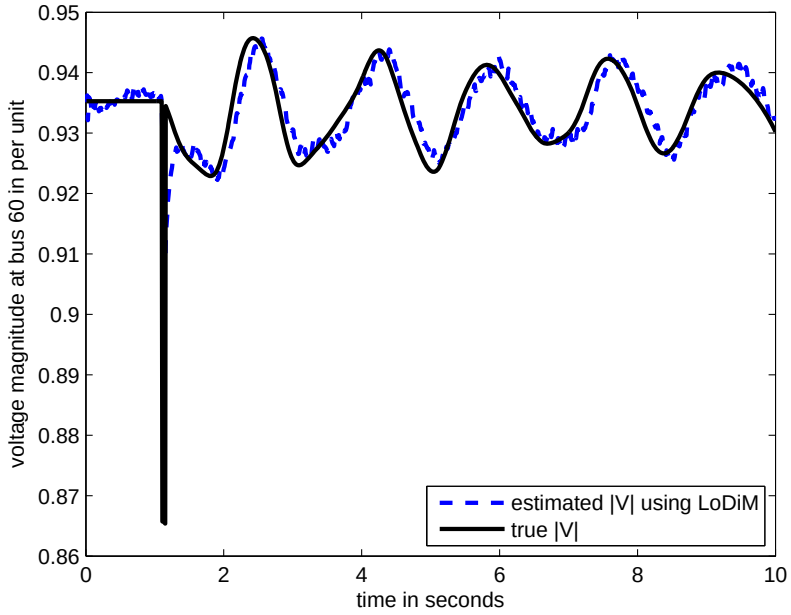


Figure 3.10: Bus 60 Voltage magnitude estimation using our LoDiM method for measurement selection.

Next, we examine a close-up shot for a better comparison. At bus 60, during the short time

window from $t = 4$ to $t = 6$ seconds, the true state and the estimated states using these three approaches are plotted in Fig. 3.11. Another similar comparison is performed at bus 26, a load bus. The zoom-in plot shown in Fig. 3.12 illustrates how the proposed LoDiM algorithm has improved the state tracking quality.

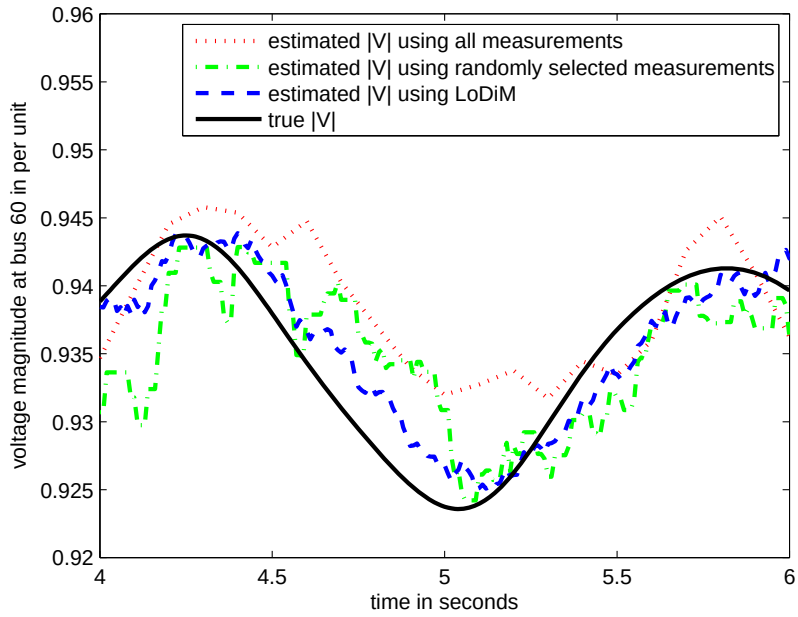


Figure 3.11: The performance comparison of different methods at generator bus 60 in the 16-generator 68-bus system, with full measurement-space (in red), with randomly selected measurements (in green) and with LoDiM (in blue), from $t = 4$ to $t = 6$

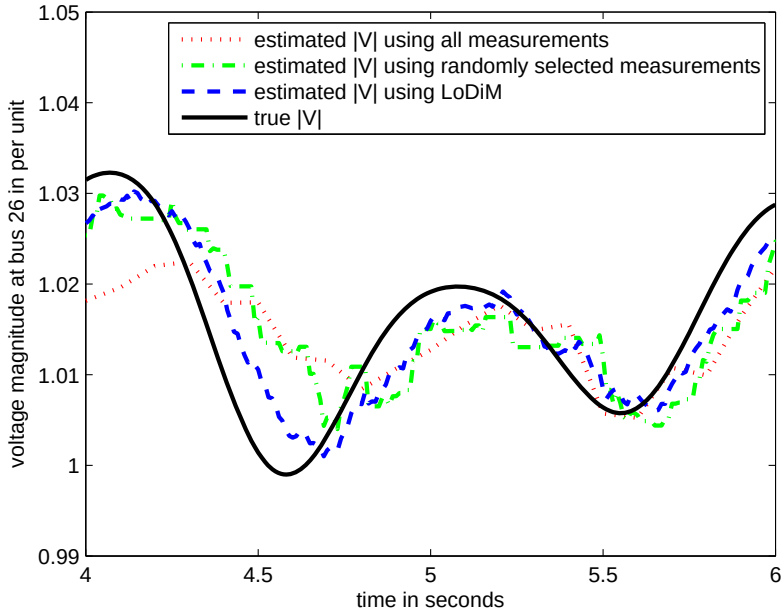


Figure 3.12: The performance comparison of different methods at generator bus 26 in the 16-generator 68-bus system, with full measurement-space (in red), with randomly selected measurements (in green) and with LoDiM (in blue), from $t = 4$ to $t = 6$

Optimal Number of Selected Measurements

The number of measurements to be selected in each cycle, m_σ, affects the estimation results. Fig. 3.13 presents the zoom-in bus 2 voltage magnitude tracking results for the smaller 3-generator 9-bus system, from $t = 2$ to $t = 3$. Similar estimation behavior is observed in the larger 16-generator 68-bus system (Fig. 3.14). The true state is plotted using a black solid line. Our LoDiM state estimation method uses dynamically selected measurements during each cycle as described in section 3.6.2. The blue dash-dash line plots the state when $m_\sigma = 6$. If m_σ is too small, *e.g.* $m_\sigma = 2$, the estimated state resembles a step function as shown in green dash-dot line. This is because very little information is integrated in each cycle, while the speed-up of the cycle period is not significant enough to capture the state variation. On the other hand, the red dotted line plots the estimated state using a conventional Kalman filter, where the entire measurement set is used as input, resulting in a lower estimation rate.

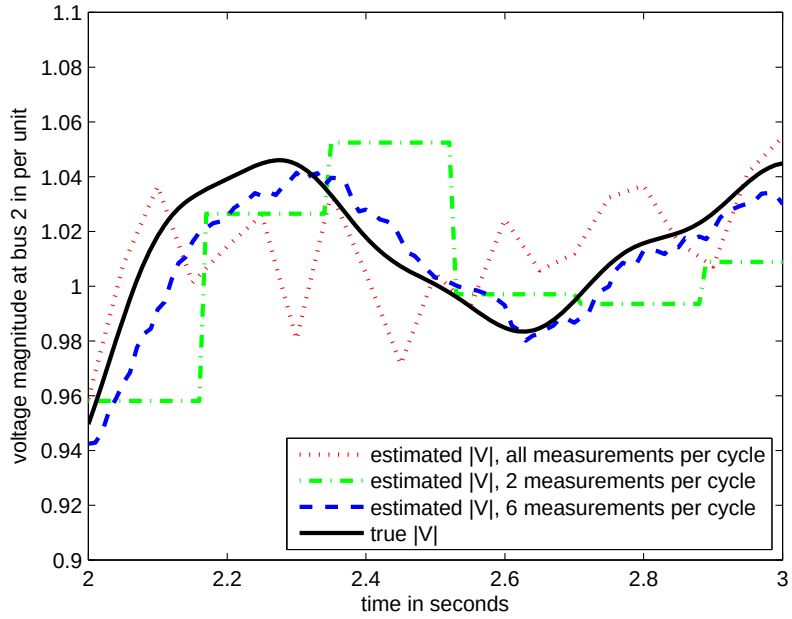


Figure 3.13: Performance comparison of LoDiM with different m_σ values, at bus 2 in the 3-generator 9-bus system

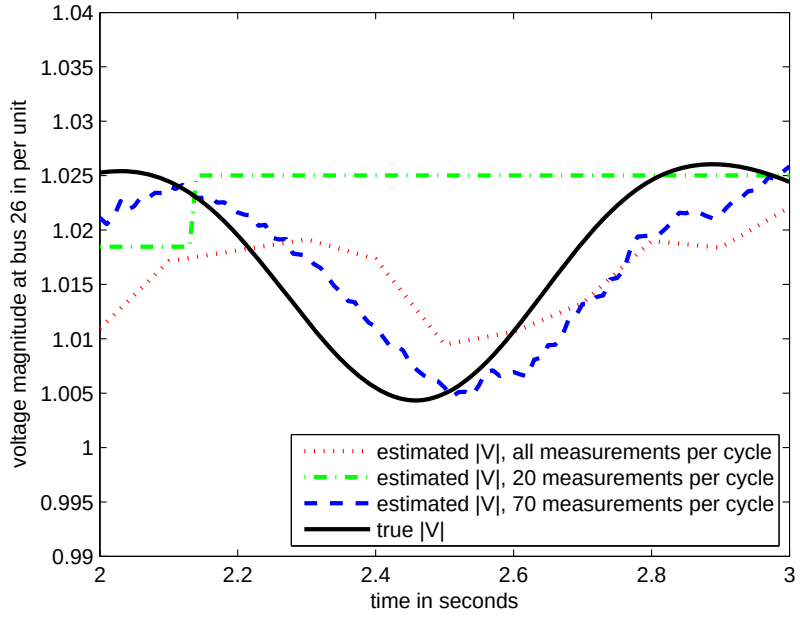


Figure 3.14: Performance comparison of LoDiM with different m_σ values, at bus 26 in the 16-generator 68-bus system

As depicted above, the LoDiM performance comparison with different m_σ values reflects an interesting trade-off between gained information and processing time. This observation leads us to believe that there exists a “sweet spot" to achieve optimal performance. To confirm our supposition, we simulated LoDiM with all possible values of m_σ, *i.e.* from $m_\sigma = 1$ to full measurement-space on both power system models.

The average state estimation absolute error (per unit, of all buses, within 10 seconds) of LoDiM, using different amount of measurements (in percentage), are plotted in blue dash-dash lines for these two system models respectively in Fig. 3.15 and 3.16. Similarly, the computational requirements (FLOP counts) for each filter cycle, using different quantities of measurements, are plotted in green dot-dot lines. Optimally adapting to the available computing power, the percentage of measurements to be selected during each cycle is 76% for the small system, and 27% for the large system.

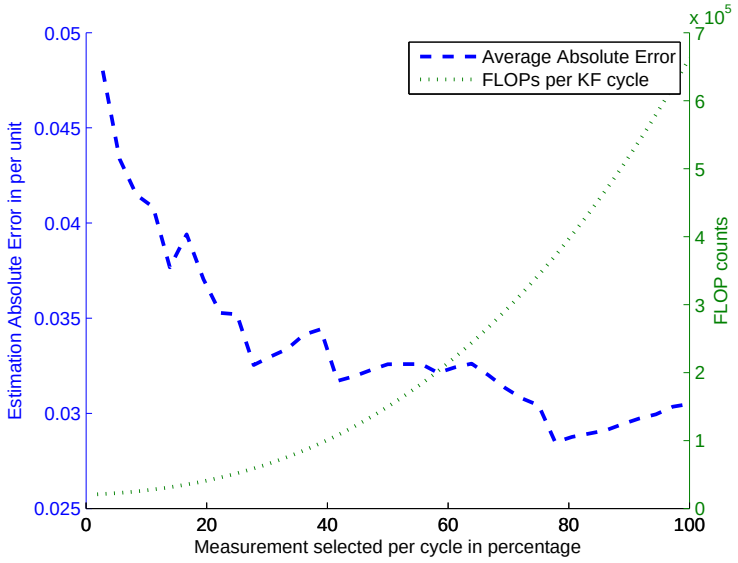


Figure 3.15: LoDiM performance versus different amount of selected measurements in the 3-generator 9-bus system, with 10 megaFLOPS computing power available

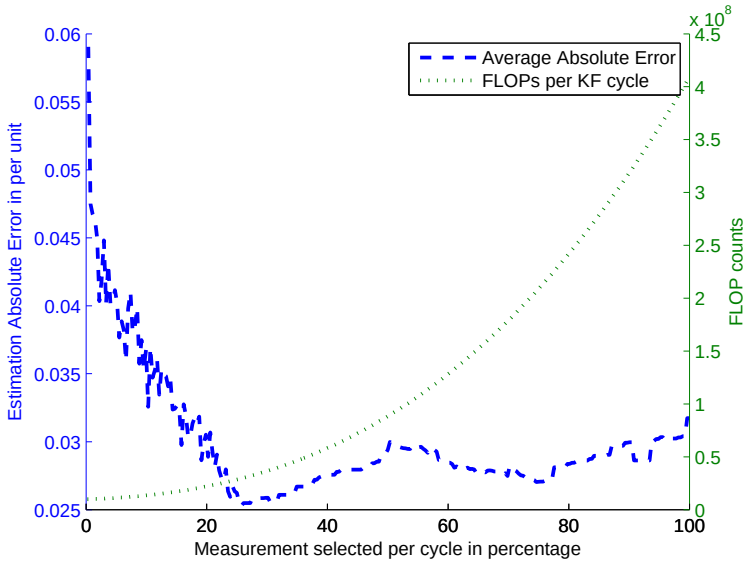


Figure 3.16: LoDiM performance versus different amount of selected measurements in the 16-generator 68-bus system. with 1 gigaFLOPS computing power available

3.7 My Method: Reduced Measurement-space Dynamic State Estimation (ReMeDySE) in Power Systems

Modern phasor measurement technologies provide high-speed sensor data with precise time synchronization for power system analysis, motivating us to consider including dynamic states in the state estimation process, *e.g.*, generator rotor angles and generator speeds, along with the *static state elements* of bus voltage magnitudes and phase angles.

Applying Kalman filter techniques to dynamic state estimation is a developing research area in modern power systems. For example, [11] investigated the feasibility of applying Kalman filter techniques to include *dynamic state elements* in the state estimation process. Compared to traditional steady state estimators, the Kalman filter is able to track dynamic state variables efficiently and accurately. However, in large-scale and wide-area interconnected power systems, the combination of high computational complexity (primarily due to the very large number of measurements) and slow processing speed (primarily due to limited computational resources)

presents a significant challenge. To help address this challenge, we have derived the Reduced Measurement-space Dynamic State Estimation (ReMeDySE) from the LoDiM algorithm in Section 3.6, for power system *dynamic state* estimation. Together with the fast advancing phasor measurement technologies, the Kalman-filter-based power system dynamic state estimation should be plausible to implement in on-line systems.

3.7.1 The Power System Dynamic State Space

The specific state elements we consider here are generator rotor angles and generator speeds. Without loss of generality, in a power system that consists of n generators and m buses, 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 for generator i can be written as

$$\begin{cases} \frac{d\delta_i}{dt} = \omega_B(\omega_i - \omega_0) \\ \frac{d\omega_i}{dt} = \frac{\omega_0}{2H_i}(P_{m_i} - \frac{|E_i||V_i|}{X'_{d_i}}\sin(\delta_i - \theta_i) - D_i(\omega_i - \omega_0)) \end{cases} \tag{3.47}$$

where state variables δ and ω are the generator rotor angle and speed respectively, ω_B and ω_0 are the speed base and the synchronous speed in per unit, P_{m_i} is the mechanical input, H_i is the machine inertia, D_i is the generator damping coefficient and $|V_i|\angle\theta_i$ is the phasor voltage at the generator terminal bus i.

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

$$\frac{dx}{dt} = A_c x + w_c, \tag{3.48}$$

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 a $2n \times 2n$ continuous time state transition Jacobian matrix.

Section 2.9 describes the detailed derivations of the corresponding *discrete-time* state transition matrix A and *discrete-time* process noise covariance Q (also see [11] and [8] for additional information). According to the process model from equation (3.1) we have

$$x_k = Ax_{k-1} + w_{k-1} \tag{3.49}$$

$$= (I + A_c \cdot \Delta t)x_{k-1} + w_{k-1}, \tag{3.50}$$

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

Power system measurements include bus voltage magnitudes, bus voltage angles and line flows. The measurement equations can be derived using the nodal admittance matrix and bus voltage reconstruction matrix, according to the expanded system nodal equation:

$$Y_{exp} \begin{pmatrix} E \\ V \end{pmatrix} = \begin{pmatrix} Y_{GG} & Y_{GL} \\ Y_{LG} & Y_{LL} \end{pmatrix} \begin{pmatrix} E \\ V \end{pmatrix} = \begin{pmatrix} I_G \\ 0 \end{pmatrix} \tag{3.51}$$

where $E = [|E_1|\angle\delta_1, |E_2|\angle\delta_2, ..., |E_n|\angle\delta_n]^T$ is the vector of internal generator complex voltages, $V = [|V_1|\angle\theta_1, |V_2|\angle\theta_2, ..., |V_n|\angle\theta_n]^T$ 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. Therefore the relationship of bus voltages V to generator voltages E can be expressed as:

$$V = -Y_{LL}^{-1} Y_{LG} E = R_V E \tag{3.52}$$

where R_V is defined as the bus reconstruction matrix. In a power flow model, the line flows are functions of bus voltages V, and hence can be easily constructed as functions of generator voltages E using (3.52).

The measurement model, as discussed in Section 2.9 (also see [11] and [8] for additional information), can be expressed in the form of equation (3.2) after linearization

$$z = Hx + v \tag{3.53}$$

where H is the corresponding Jacobian matrix

$$H = \begin{pmatrix} \frac{\partial h_1}{\partial \delta_1} & 0 & \frac{\partial h_1}{\partial \delta_2} & 0 & \cdots \\ \frac{\partial h_2}{\partial \delta_1} & 0 & \frac{\partial h_2}{\partial \delta_2} & 0 & \cdots \\ \frac{\partial h_3}{\partial \delta_1} & 0 & \frac{\partial h_3}{\partial \delta_2} & 0 & \cdots \\ \frac{\partial h_4}{\partial \delta_1} & 0 & \frac{\partial h_4}{\partial \delta_2} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{3.54}$$

and v is the measurement noise with normal probability distribution $p(v) \sim N(0, R)$. The measurement noise covariance R can be determined experimentally, by off-line testing of the measurement devices over time and (if desired) under different conditions.

Note that the measurement Jacobian H has all zero entries in even-numbered columns. Thus in the measurement selection procedure (previously described in Section 3.6.2), we can save the computational cost by performing PCA on an *a priori* error covariance submatrix P_s^-, consisting of elements of the odd-numbered rows and odd-numbered columns of P^-. The computational cost of this procedure is further reduced, because the corresponding measurement Jacobian submatrix H_s is then formed by only odd-numbered columns of H.

3.7.2 Simulation Results

We used the ReMeDySE approach on the standard 16-generator 68-bus model (Fig. 2.6) representing the New England/New York interconnected system. We simulated a three-phase fault at bus 29, starting from $t = 1.1$ and lasting for 0.05 seconds, as a dynamic event representing a large disturbance emergency.

We simulated the dynamic state variables (the rotor angles and speeds of each generator) and recorded them as the actual dynamic states. Here we made the strong assumption that all bus voltages and phase angles are measured by PMUs. In the future we will relax this assumption. These simulated PMU measurements contain 1% random noise this time.

Fig. 3.17 depicts the actual and the estimated speed (using ReMeDySE with only 15% measurements per cycle, and regular EKF with all PMU measurements per cycle) of generator 2 within 10 seconds from $t = 0$ to $t = 10$, to capture its reaction to this large disturbance.

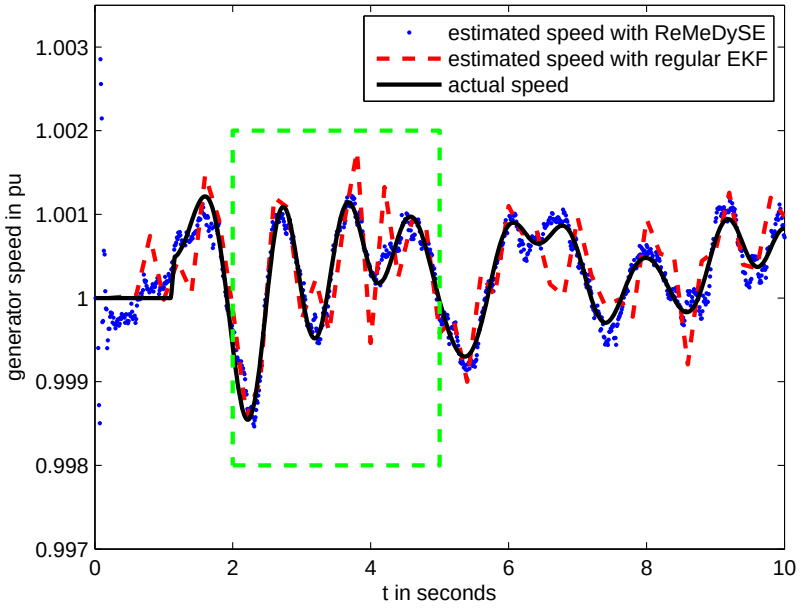


Figure 3.17: Estimation of rotor speed at generator 2 using ReMeDySE and regular EKF, with a 3-phase fault at bus 29

For a better comparison, Fig. 3.18 zooms in the green dashed-line window, and gives us a close-up look of the short window from $t = 2$ to $t = 5$ of our simulation. Similarly, Fig. 3.19 shows the rotor angle estimation results at this same generator during the same window.

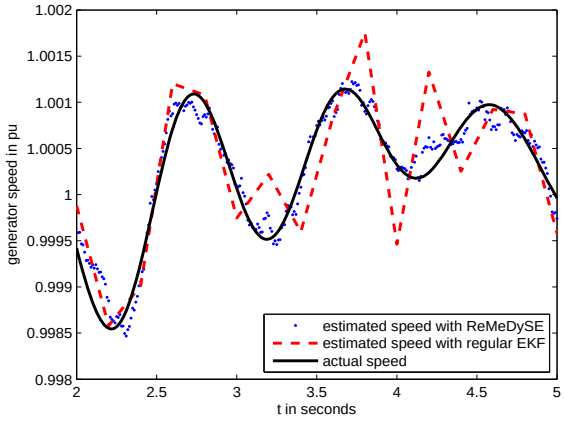


Figure 3.18: Performance comparison of ReMeDySE and regular EKF estimating rotor speed at generator 2, from $t = 2$ to $t = 5$

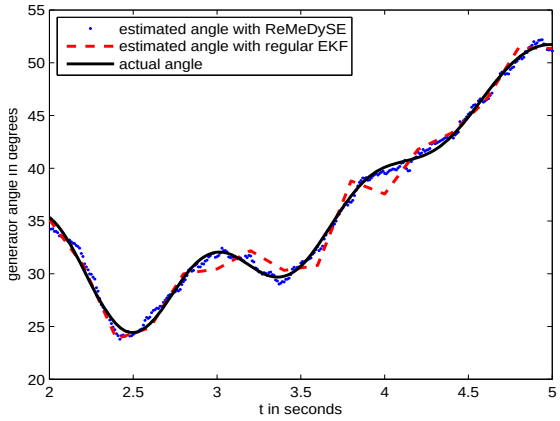


Figure 3.19: Performance comparison of ReMeDySE and regular EKF estimating rotor angle at generator 2, from $t = 2$ to $t = 5$

Take generator 10 as another example, Fig. 3.20 and Fig. 3.21 illustrate the rotor speed and angle estimation results at generator 10 respectively, during the same period of time.

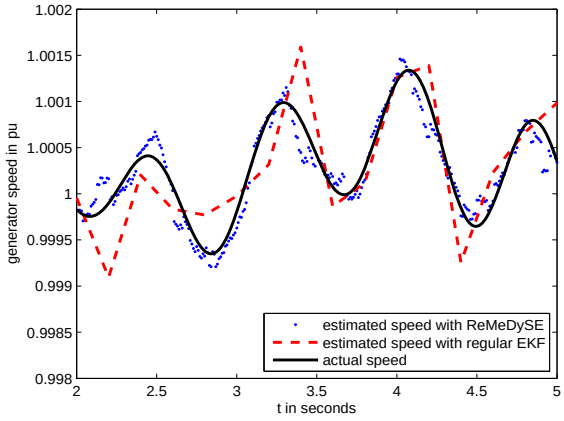


Figure 3.20: Performance comparison of ReMeDySE and regular EKF estimating rotor speed at generator 10, from $t = 2$ to $t = 5$

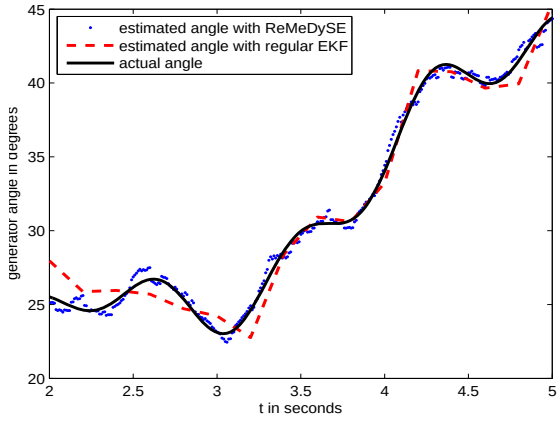


Figure 3.21: Performance comparison of ReMeDySE and regular EKF estimating rotor angle at generator 10, from $t = 2$ to $t = 5$

These plots demonstrate that under the experimental conditions, our proposed ReMeDySE approach outperforms the traditional EKF approach with higher accuracy. The primary reason is that the standard batch approach for EKF-based dynamic state estimation uses the entire

measurement set as input at each step, resulting in lower estimation rate and higher latency. The dynamic nature of the state variables introduces additional noise (uncertainty) into the process, and this noise increases as the latency increases. Typically when latency is small, the measurement device noise dominates; when latency is large, the process noise dominates. In ReMeDySE, only a small and dynamically selected subset of PMU measurements is used, to adapt to the limited computational resource. Our method, combined with developing PMU technologies, makes high-speed high-accuracy and high-flexibility dynamic state estimation realizable in future power systems.

CHAPTER 4

ADAPTIVE AND ROBUST ESTIMATION

Kalman filters have been successfully applied in many different research areas over the last five decades. In modern power systems, especially with the emergence of real-time digital meters, they are commonly seen in system *state tracking* methods to enable real-time monitoring and control.

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 we propose an efficient adaptive Kalman filter algorithm, that can identify and reduce the impact of incorrect system modeling and/or bad PMU measurements. We call this novel approach Adaptive Kalman Filter with Inflatable Noise Variances (AKF with InNoVa). Simulations demonstrate its robustness to sudden changes of system dynamics and erroneous measurements.

Utilizing AKF with InNoVa as a fundamental building block, later we develop a real-time two-stage Kalman filter approach [45]: at every time step, stage one takes the raw PMU data into the AKF with InNoVa, to estimate the *traditional* states (bus voltage phasors); then stage two feeds the results into an Extended Kalman filter (EKF) to estimate the truly dynamic states (generator rotor angles and speed). Our method is efficient, effective and robust under various simulated scenarios.

4.1 Introduction

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. Nowadays, the high-speed time-synchronized data provided by PMUs is making real-time power system state estimation possible.

The Kalman filter is emerging as a popular mechanism for estimating such power system states dynamically, in conjunction with real-time phasor measurement technologies. Ideally, the employed process and measurement models would represent the “true" systems, with accurate statistical noise characteristics. In practice one has to develop models for the systems, whereas all such models are more or less sub-optimal.

Fortunately, as statistician George Box once said, “All models are wrong, but some are useful" [46]. With these “useful" models, Kalman filter are able to estimate states even when the precise nature of the modeled system is unknown. Nevertheless, when the system experiences unpredictable process changes, and/or the measurements contain significant errors, the estimated states could deviate from the true states rapidly.

A lot of efforts have been made to improve Kalman filters under different circumstances [47, 48, 49, 50]. Here we propose a more general approach: Adaptive Kalman filter with Inflatable Noise Variances (*AKF with InNoVa*), to achieve more robust state estimation. By simultaneously adjusting noise parameters Q *and* R on-the-fly, this novel AKF approach is remarkably efficient in dealing with inaccurate system models and un-modeled measurement errors. Taking one step further, we propose a real-time two-stage Kalman filter approach, for estimating the *traditional states* as well as the *dynamic states* in a power system. At every time step, stage one takes the raw PMU measurements into the AKF with InNoVa, to estimate the *traditional states*; then stage two takes the results into an Extended Kalman filter (EKF) to estimate the truly *dynamic states*. Both stages produce state estimates that are critical in

various power system applications.

4.2 Chapter Organization

In the rest of this chapter, Section 4.3 introduces the background. Section 4.4 briefly describes the traditional methods for bad data detection and identification. Section 4.5 presents the principles and implementation details of our AKF with InNoVa. Built on this adaptive Kalman filter algorithm, the two-stage Kalman filter approach is proposed in Section 4.6.

4.3 Background and Related Work

Traditionally, a weighted least square (WLS) estimator is the classic method for power system state estimation. In WLS state estimation, methods such as *Largest Normalized Residual Test* and *Hypothesis Testing Identification* are proposed and extensively used for bad data detection and identification [21]. For example, in [51] the authors study the usage of phasor measurements in WLS state estimation, and how to identify bad data with normalized residual vectors. In [50], although the authors proposed a robust extended Kalman filter (REKF) for dynamic harmonic state estimation, which can detect the presence of sudden load change or suspicious measurements with normalized EKF innovation vectors, they still use normalized WLS residual vectors to confirm bad measurements.

WLS estimators are considered “static" because at any one time instant, the state is calculated only from the measurement set at that time. On the other hand, if an estimator can predict the state ahead of time, it is considered “dynamic". [33] gives an overview on different methodologies and developments in power system tracking and dynamic state estimation based on previous research.

Although the term *dynamic state estimation* (DSE) can be traced back to the 1970s [32], it was the development of phasor measurement technologies since the 1980s [2] that made it possible to capture the dynamic behavior of the system. Most DSE approaches are Kalman-filter-based, despite the fact that a reliable system state evolution (process) model is usually not available. To overcome this situation, the approach proposed in [52] predicts the load flow data at the next time instant using a short-term nodal load forecasting technique, and combines the PMU data at the next time instant for DSE. Another approach is presented in [53], where the authors use historical system state data and the Holt's double exponential smoothing method to estimate process model parameters such as the *state transition matrix* and the trend component. [54] presents more detailed descriptions of these approaches and a comprehensive survey of other state estimator formulations in the presence of PMUs.

A major challenge with DSE using PMUs is that large-scale measurement data needs to be processed in a very short period of time (short system update cycle). This is one reason why dynamic adaptation of noise modeling parameters seems attractive. Previous studies have demonstrated satisfying tracking results by adjusting noise parameters, under different conditions. For example, to deal with unpredictable process changes, [47] has proposed a reverse prediction adaptive Kalman filter algorithm, which adjusts the process noise covariance Q to improve filter precision, assuming the measurement model is correct. As a comparison, [48] adjusts the measurement noise covariance R instead of Q, to increase the robustness against erroneous measurements. Assuming the *a priori* noise statistics Q and R are both unknown, the authors of [49] constructed the autocorrelation functions of innovation sequence, and incorporated information about the quality of the autocorrelation parameters through a WLS method. The latter permits to generate a convergent sequence to the steady state filter, which allows to determine Q and R after some manipulations.

In our approach, we have in mind the existence of unexpected process changes as well as measurement errors. Our proposed method, Adaptive Kalman Filter with Inflatable Noise

Variances (AKF with InNoVa), is able to adjust noise statistics Q and R on-the-fly by analyzing the *innovation* and *residual* vectors. These inflatable (and deflatable) noise variances are the keys to our algorithm's robustness.

Furthermore, as mentioned above, [11] and [34] utilized an extended Kalman filter (EKF) based approach to estimate the *dynamic states* of power systems. 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, and the hypothetical models are the "true" models, with accurate noise statistical characteristics. Now with AKF with InNoVa, we can relax this assumption: our two-stage Kalman filter approach is able to take in inaccurate system models and erroneous measurements, and produce robust estimation of the *traditional states* as well as the *dynamic states* in a power system, given sufficient data redundancy.

4.4 Traditional Bad Data Detection and Identification

In practice, even with accurately modeled systems, unexpected bad measurements can occur due to various reasons. Thus people have developed many techniques to handle these bad data. Once detected, they need to be identified and eliminated or corrected, to support unbiased state estimation.

Considering the state space models described in Section 2.4 and the WLS state estimator described in Section 2.6, we will briefly introduce some traditional bad data detection and identification methods in the following subsections.

4.4.1 Chi-squares χ^2-test

One of the most common methods for detecting bad data is *Chi-squares (χ^2) test*. It is comprised of the following steps:

1. After solving the WLS estimation problem, compute the objective function $J(\hat{x})$ from equation (2.15):

$$J(\hat{x}) = \sum_{i=1}^{m} \frac{(z_i - h_i(\hat{x}))^2}{R_{ii}} = \sum_{i=1}^{m} (e_{z_i}^N)^2 \tag{4.1}$$

   where $\hat{x} \in \mathcal{R}^n$ is the estimated state vector, $h(\hat{x}) \in \mathcal{R}^m$ is the estimated measurement vector, $z \in \mathcal{R}^m$ is the measurement vector, $(z - h(\hat{x}))$ is the measurement residual vector, $R_{m \times m}$ is the measurement error covariance matrix.

   Note that here we are approximating measurement error $(z - h(x))$ by measurement residual $(z - h(\hat{x}))$, and at least n measurements need to satisfy power balance equations, so at most $(m - n)$ of the normalized measurement "errors" $e_{z_i}^N$ will be independent random variables with Standard Normal Distribution $\mathcal{N}(0, 1)$. Thus $J(\hat{x})$ is assumed to have a χ^2 distribution with $(m - n)$ degrees of freedom.

2. Test the value of $J(\hat{x})$:

$$J(\hat{x}) \begin{cases} \geq \chi^2_{(m-n,p)}, & bad\ data\ detected \\ < \chi^2_{(m-n,p)}, & no\ bad\ data \end{cases} \tag{4.2}$$

   The threshold $\chi^2_{(m-n,p)}$ can be looked up from the χ^2 distribution table, corresponding to a pre-specified detection confidence level with probability p.

χ^2 test is computationally inexpensive, however the trade-off is its inaccuracy. It may fail to detect bad data due to the approximation of measurement errors by measurement residuals in equation (4.1).

4.4.2 Largest Normalized Residual Test

To promote accuracy, one of the most popular bad data identification (and inherently bad data detection) method, the *Largest Normalized Residual* (r^N_{max}) *Test* method, was proposed with the following steps:

1. After solving the WLS estimation problem, compute the normalized residual vector r^N_z with each element

$$r^N_{z_i} = \frac{|z_i - h_i(\hat{x})|}{\sqrt{\Omega_{ii}}}, \quad i = 1, 2, ..., m \tag{4.3}$$

   where $\Omega_{m\times m}$ is the measurement residual covariance matrix calculated as follows

$$\Omega = R - H(H^T R^{-1} H)^{-1} H^T \tag{4.4}$$

   The detailed proof of $Cov(r_z) = \Omega$ can be found in many technical literatures such as [21].

2. If the largest element in the normalized residual r^N_z is larger than the pre-specified identification threshold (*e.g.*, 3.0), the corresponding measurement is suspected as bad data, the procedure continues to step 3. Else, the procedure concludes with no bad data suspected.
3. Eliminate the suspected measurement from the measurement set and go to step 1.

The (r^N_{max}) test can perform the detection and identification processes simultaneously using normalized residuals. However, the measurement residual covariance matrix Ω is more computationally expensive, thus χ^2 test is often used for preliminary bad measurement detection.

4.5 My Method: Adaptive Kalman Filter with Inflatable Noise Variances (AKF with InNoVa)

This section first briefly discusses the system modeling of the conventional power system state estimation problem, with PMU measurements. Then we give a detailed description of the

Adaptive Kalman filter with Inflatable Noise Variances algorithm.

4.5.1 System modeling

With the future power systems in mind, we assume the measurements are high-speed synchrophasors provided by PMUs [55]. Because the relationship between the measurements and the states can be effectively linear, the rectangular coordinate formulation will be preferable [54], where the real and imaginary parts of bus voltages are considered state variables. This formulation can prevent numerical problems encountered during flat start when using current phasors. For the process model, the power system is assumed stable, hence a quasi-static power system model is employed in our simulations (note that this could be a wrong assumption).

Due to resource limitations, engineers typically have to place PMUs selectively at scattered locations to cover the entire system [8, 56, 57], resulting in low data redundancy level. Nonetheless, the redundancy can be increased if we treat the state predictions in our dynamic estimation method as another input source. Here we assume that a PMU installed at a specific bus can measure the bus voltage phasor, as well as the current phasors along all the lines incident to the bus. The modeling details can be found in our previous research [8] (also see Section 2.8).

4.5.2 Adaptive Kalman Filter Algorithm

Recall that we have introduced Kalman filters in Section 3.4. The traditional Kalman filter is actually a recursive solution to the discrete-data linear filtering problem. It estimates a process by using a form of feedback control: the filter estimates the process state at some time and then obtains feedback in the form of noisy measurements.

The equations for the Kalman filter fall into two groups—time update equations ("Prediction") and measurement update equations ("Correction"):

$$Prediction: \begin{cases} \hat{x}_k^- = A\hat{x}_{k-1} \\ P_k^- = AP_{k-1}A^T + Q \end{cases} \tag{4.5}$$

$$Correction: \begin{cases} K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \\ \hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \\ P_k = (I - K_k H)P_k^- \end{cases} \tag{4.6}$$

We define the *innovation vector*, $\mathcal{I}_k^- = (z_k - H\hat{x}_k^-)$, to be the difference between the true measurement vector and that computed from the *a priori* state estimate. The corresponding *innovation covariance* is $S_k = HP_k^- H^T + R$ [36].

Although some authors use the terms "innovation" and "residual" interchangeably, they should not be confused here. Throughout we refer to $\mathcal{I}_k^- = (z_k - H\hat{x}_k^-)$ as the *innovation*, to distinguish it from the *residual* $\mathcal{I}_k = (z_k - H\hat{x}_k)$, which is the difference between the true measurement vector and that computed from the *a posteriori* state estimate [58].

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

Proof. According to the *a posteriori* state estimate in (4.6) by incorporating the measurement, we have

$$\begin{aligned} \hat{x}_k &= \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \\ H\hat{x}_k &= H\hat{x}_k^- + HK_k(z_k - H\hat{x}_k^-) \\ z_k - H\hat{x}_k &= (z_k - H\hat{x}_k^-) - HK_k(z_k - H\hat{x}_k^-) \\ z_k - H\hat{x}_k &= (I - HK_k)(z_k - H\hat{x}_k^-), \end{aligned} \tag{4.7}$$

where I is the identity matrix. Then the mean of the *residual* is

$$E(z_k - H\hat{x}_k) = (I - HK_k)E(z_k - H\hat{x}_k^-) = 0, \tag{4.8}$$

and the covariance of the *residual* is

$$\begin{aligned} cov(z_k - H\hat{x}_k) &= (I - HK_k)cov(z_k - H\hat{x}_k^-)(I - HK_k)^T \\ &= (I - HK_k)S_k(I - HK_k)^T. \end{aligned} \tag{4.9}$$

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

$$\begin{aligned} I - HK_k &= I - HP_k^- H^T(HP_k^- H^T + R)^{-1} \\ &= R(HP_k^- H^T + R)^{-1} \\ &= RS_k^{-1}. \end{aligned} \tag{4.10}$$

Combining (4.9) and (4.10), and we can write

$$\begin{aligned} cov(z_k - H\hat{x}_k) &= RS_k^{-1}S_k(RS_k^{-1})^T \\ &= R(S_k^{-1})^T R^T \\ &= RS_k^{-1}R, \end{aligned} \tag{4.11}$$

because the *innovation covariance* S_k and measurement noise covariance R are both symmetric matrices. We define $T_k = RS_k^{-1}R$ as the *residual covariance*. ■ Based on the lemma proved above, we propose our new algorithm: Adaptive Kalman Filter with Inflatable Noise Variances (AKF with InNoVa).

In the AKF we treat all un-modeled 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 also treat deviations from the true process model as un-modeled noise that is not reflected in the process noise covariance Q. For now we do not consider cases with multiple/switching process models.

To assess the performance of a filter, traditionally people examine the *innovation* $\mathcal{I}_k^- = (z_k - H\hat{x}_k^-)$, which should have a normal distribution with zero mean and covariance S_k. When the implemented system model does not match reality, the mean of the innovation can shift, and the magnitude cam grow, so that eventually some element(s) in the *normalized innovation vector* (normalized by its covariance S_k) will exceed a predetermined threshold. Usually, the corresponding measurement(s) are identified as the *outliers*, and are excluded from further computation. However, due to the nature of Kalman filter (Figure 4.1), it is usually impossible to determine whether the shift/growth is caused by a system model mismatch, a measurement error, or both. If it is the first case, we certainly do want to preserve those apparent “outlier" measurements.

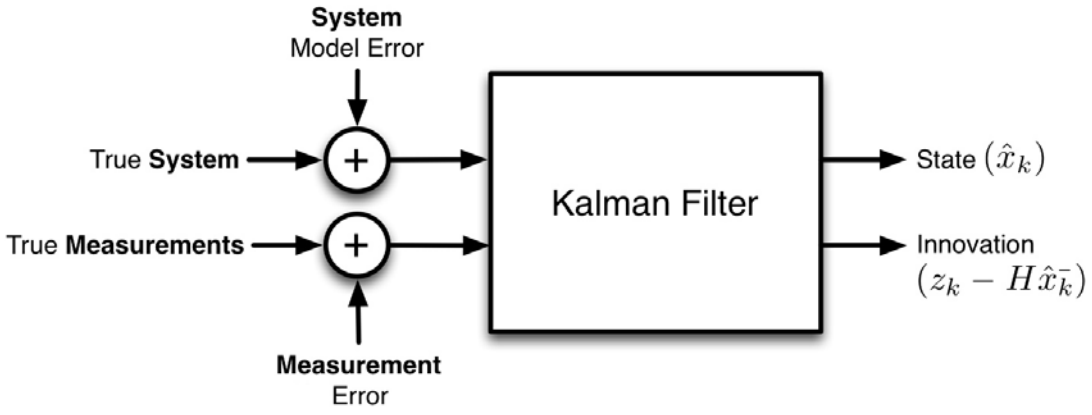


Figure 4.1: High-level diagram of Kalman filter, courtesy of Greg F. Welch. System model error, and/or measurement error, can all lead to significant state estimation error when passing through a Kalman filter

In other words, it is insufficient to determine the “outliers" by solely examining the *innovations*, because Q and R are already blended (indistinguishable) in S_k. This is why we want to investigate the *residual* $\mathcal{I}_k = (z_k - H\hat{x}_k)$ as well: its ideal covariance T_k can be used to identify the un-modeled measurements errors. Thus instead of simply discard suspicious measurements, this pair of *innovation-residual tests* allow us to give them a second chance.

The basic idea of our approach is as follows. To simplify the notation we omit the time

step count k. Within each cycle, after the prediction step, we compute the ideal *innovation covariance* S and the *normalized innovation vector* $\tilde{\mathcal{I}}^-$ where

$$\tilde{\mathcal{I}}_i^- = |\mathcal{I}_i^-|/\sqrt{S_{ii}}, \quad i = 1, 2, ...n. \tag{4.12}$$

If $\tilde{\mathcal{I}}_i^- > \tau$ for some threshold τ, then $i \in Out$, where Out holds the outlier indices. In our experiments, we used $\tau = 3$ (measurement units).

First we assume the anomaly are caused by unknown process noise, so we want to "inflate" Q by a diagonal matrix ΔQ, so that P^- is also inflated by ΔQ. Thus S is consequently inflated by $\Delta S = H(\Delta Q)H^T$ and

$$\tilde{\mathcal{I}}_i^- = |\mathcal{I}_i^-|/\sqrt{S_{ii} + \Delta S_{ii}} \leq \tau, \quad i = 1, 2, ...n. \tag{4.13}$$

It is easy to show that

$$\Delta S_{ii} = \sum_{j=1}^{n} H_{ij}^2 \Delta Q_j = (H(i,:) \cdot H(i,:))([\Delta Q_1, ...\Delta Q_n]^T), \tag{4.14}$$

where "$\cdot$" denotes dot product. We could simply use linear programming method to solve an optimization problem:

$$\begin{aligned} &\min \sum_{j=1}^{n} \Delta Q_j \qquad (4.15)\\ s.t. \quad &\Delta S_{ii} = (H(i,:) \cdot H(i,:))([\Delta Q_1, ...\Delta Q_n]^T) \\ &\geq (|\mathcal{I}_i^-|/\tau)^2 - S_{ii}, \quad \forall i \in Out \\ &\Delta Q_1 \geq 0, \Delta Q_2 \geq 0, ...\Delta Q_n \geq 0. \end{aligned}$$

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

$$\tilde{\mathcal{I}}_i = |\mathcal{I}_i|/\sqrt{T_{ii}}, \quad i = 1, 2, ...n. \tag{4.16}$$

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

Now we can separate the measurement “noise" from the process “noise". Let us denote $ProcOut = Out \backslash MeasOut = \{i : i \in Out \quad and \quad i \notin MeasOut\}$ as the set difference between Out and $MeasOut$. If $MeasOut$ is not empty, we will recalculate the ΔQ as in optimization problem (4.15), except only for $\forall i \in ProcOut$, and update Q to $Q + \Delta Q$. As for the measurements, we “inflate" R in this way: for $\forall i \in MeasOut$, R_{ii} is inflated to $\lambda_i R_{ii}$. As a result, $T = RS^{-1}R$ is also inflated so that the ith diagonal element is now $\lambda_i^2 T_{ii}$ and

$$\tilde{\mathcal{I}}_i = |\mathcal{I}_i|/\lambda_i\sqrt{T_{ii}} \leq \tau, \quad i = 1, 2, ...n. \tag{4.17}$$

It is very easy to compute λ_is by

$$\lambda_i = (|\mathcal{I}_i|/\sqrt{T_{ii}})/\tau, \quad i \in MeasOut. \tag{4.18}$$

Finally, with the inflated Q and R, we recompute the correction step (4.6) 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. Thus we adopt an exponential decay in the process model to enable automatic deflation of the parameters. The decay speed can be customized by users, according to their expectation and the specific circumstances. For instance, the user can set a decay constant r, such that

$$Q_k = Q_{k-1} * e^{-r}, R_k = R_{k-1} * e^{-r} \tag{4.19}$$

To sum up, the AKF with InNoVa has the following steps within each cycle:

1. Predict the state and error covariance using time update equations (4.5).
2. Compute the *normalized innovation vector* $\tilde{\mathcal{I}}^-$ as in (4.12). Determine the set of outliers Out by collecting any i such that $\tilde{\mathcal{I}}_i^-$ exceeds the threshold τ. If $Out = \emptyset$, deflate Q and R as in (4.19) and go to step 8. Else, continute.
3. Construct the optimization problem (4.15) for $\forall i \in Out$ and solve for ΔQ_i.

4. Compute the *normalized residual vector* $\tilde{\mathcal{I}}$ as in (4.16). Determine the set of measurement outliers $MeasOut$ by collecting any i such that $\tilde{\mathcal{I}}_i$ exceeds the threshold τ, and the set of process outliers $ProcOut = Out \backslash MeasOut$ by taking the set difference. If $MeasOut = \emptyset$, go to step 7. Else, continue.
5. Reconstruct the optimization problem (4.15) for $\forall i \in ProcOut$ and solve for ΔQ_i.
6. Inflate measurement noise variances R_{ii} to $\lambda_i R_{ii}$, where λ_i is computed in (4.18).
7. Inflate process noise variances Q_{ii} to $Q_{ii} + \Delta Q_i$.
8. With updated Q and R, correct the state and error covariance using measurement update equations (4.6) and go to step 1.

4.5.3 Algorithm Performance Evaluation

In order to compare our proposed algorithm against other Kalman filter algorithms, we have considered two measures of accuracy that are most commonly used in quantitative methods of forecasting:

- Mean Absolute Error, denoted by *MAE*. At each time step, it is calculated in the following way:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{x}_i - x_i| = \frac{1}{n}\sum_{i=1}^{n}|e_i| \tag{4.20}$$

  where $\hat{x}_i$ is the estimated state at bus i, x_i is the actual state at bus i, and $e_i = |\hat{x}_i - x_i|$ is the absolute error.

- Mean Absolute Percentage Error, denoted by *MAPE*. At each time step, it is calculated as follows:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}|\frac{\hat{x}_i - x_i}{x_i}| \times 100\% = \frac{1}{n}\sum_{i=1}^{n}|\frac{e_i}{x_i}| \times 100\% \tag{4.21}$$

where $\hat{x}_i$ is the estimated state at bus i, x_i is the actual state at bus i, and $e_i = |\hat{x}_i - x_i|$ is the absolute error.

Mean Absolute Percentage Error is the mean of the sum of all percentage errors for a given data set taken without regard to sign. In other words, the absolute relative errors are summed and the average computed. Comparing to *MAE*, *MAPE* does not depend on the scaling of the variable, which may be inconvenient if the criteria are used for comparing predictive accuracy across different variables or different time ranges. However, *MAPE* achieves scale independence by a simple division by x_i. In practical applications, this entails two major problems:

- If there are zero values, there will be a singularity problem. The contribution of bus i at that time point and hence the *MAPE* are undefined.
- Even if x_i is only approximately zero, the relative contribution of bus i will be enormous. Usually, there is no justification for preferring a high precision for small values of x_i. When there is a perfect fit, *MAPE* is zero. However *MAPE* has no restriction on its upper level.

In our application of power system analysis, we have adopted a per-unit system, in which the system quantities are expressed as fractions of a defined base unit quantity. Besides, when a three-phase fault happens, the voltage at that bus can actually drop to a value that is very close to zero. Taking all aspects into consideration, we have decided to use *MAE* to evaluate the state estimation accuracy of each algorithm.

4.5.4 PMU Measurement Evaluation

According to IEEE C37.118.1-2011 — IEEE Standard for Synchrophasor Measurements for Power Systems [55], the measurement of a PMU must maintain certain phasor accuracy

over a range of operating conditions. The theoretical phasor value of a signal being measured and the estimated values obtained from a PMU may include differences in both magnitude and phase angle. These differences have been combined into a single error quantity called the “Total Vector Error" or *TVE*. The standard has defined this quantity as the measure of accuracy for PMUs.

The criterion *TVE* is an aggregated difference between a “perfect" sample of a theoretical synchrophasor and the estimate given by the PMU under test at the same time instant. It is defined as the square root of the difference squared between the real and imaginary parts of the theoretically “true" phasor and the estimated phasor, ratioed to the magnitude of the theoretical phasor and presented as a percentage as seen in Equation (4.22). In this way, the value is normalized and expressed as per unit of the theoretical phasor.

$$TVE_n = \sqrt{\frac{(Re(\hat{x}_n) - Re(x_n))^2 + (Im(\hat{x}_n) - Im(x_n))^2}{(Re(x_n))^2 + (Im(x_n))^2}} * 100\% \tag{4.22}$$

where $Re(\hat{x}_n)$ and $Im(\hat{x}_n)$ are the estimated phasor values given by the unit under test at time instant n, $Re(x)_n$ and $Im(x)_n$ are the true phasor values of the input signal at time instant n. Note that as with most measurements, the “true" value can never be precisely known, thus we rely on calibration to establish the bounds within which the measurement (the vector) has a high probability of ocurring.

A 1% criterion established by setting $TVE = 1\%$ in Equation (4.22) can be visualized as a small circle drawn on the end of the phasor vector. The maximum magnitude error is 1% when the error in phase angle is zero, and the maximum error in angle is about 0.573 degrees. Fig. 4.2 shows the circle, whose size is greatly exaggerated for clarity.

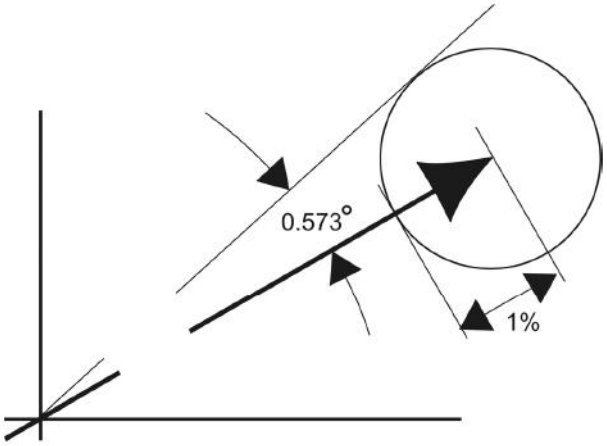


Figure 4.2: The circular boundary of a 1% *TVE* criterion

4.5.5 Simulation Results

We have simulated the performance of the proposed adaptive Kalman filter algorithm (AKF with InNoVa) on several multi-machine systems with PMU measurements, under various conditions. For better illustration, we will show the simulated results from the 3-generator-9-bus system, Fig. 2.2, as seen in Section 2.8.3. In total 5 seconds are simulated in time steps of 0.01 seconds. The bus voltages are stable until an unexpected emergency event occurs: a three-phase fault at bus 6 happens at $t = 1$ second, and is then cleared in 0.15 seconds. This event represents a large disturbance emergency, causing the oscillations in all bus voltages. The simulated PMU data contains random noise where $TVE = 5\%$. Under ideal conditions, where system models and noise statistics match the reality and no measurement errors, our AKF with InNoVa performs exactly the same as the traditional KF. The normal cases are omitted here, because we are more interested in the filter performance under adverse circumstances.

We have simulated different abnormal test cases to illustrate the robustness of our proposed algorithm. In each case, we executed three Kalman filter algorithms for a more informative comparison: one is the traditional Kalman filter; another is a naive robust Kalman filter (RKF)

which does not adjust model parameters, but uses the largest normalized innovation test to identify and exclude bad measurements; the last one is our proposed AKF with InNoVa algorithm. We visualize the voltage tracking results with three-dimensional, real-part and imaginary-part plots separately.

Then we illustrate the *MAE* plots and their histograms for a more comprehensive comparison. In this particular application, the states are complex voltages. Thus the *MAE* can also be written as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \|\hat{x}_i - x_i\| = \frac{1}{n}\sum_{i=1}^{n} \|e_i\| \tag{4.23}$$

where $\| \cdot \|$ denotes the norm (magnitude) of the complex values. To maintain consistency of our experiments in this subsection, the data range of each histogram is set to be $[0, 1]$, and the bin (*i.e.* interval) size is set to be 0.01. These test cases are stated below.

Case 1: under-valued initial Q and R

In this case, the process noise variance of each state variable in per unit is initialized to be $1e^{-6}$, resulting in a quite small process noise covariance Q, which means the modeler is very confident that the system will be stable. Also, the PMU measurement noise is initialized to be 3% only, while it is actually simulated to be 5% *Total Vector Error*.

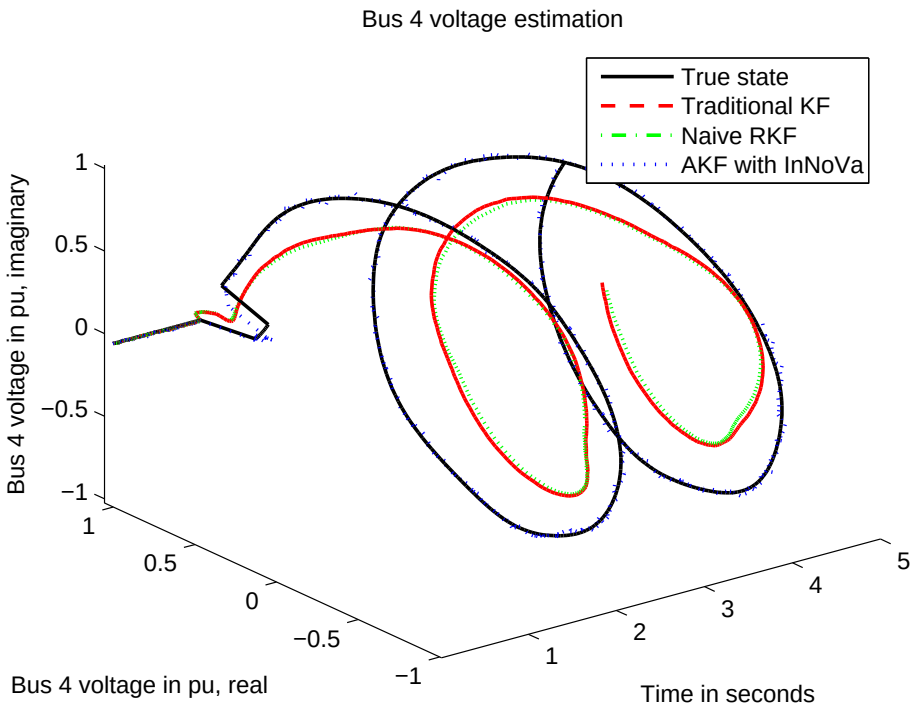
Without loss of generality, we will show the voltage tracking results of the three filters for bus 4 in Fig. 4.3. In this and the following simulations, we always use a black solid line to plot the true state, a red dash-dash line to plot the traditional KF estimated state, a green dash-dot line to plot the naive RKF estimated state, and a blue dot-dot line to plot the state estimated by our AKF with InNoVa.

From Fig. 4.3 we can tell that the KF and the naive RKF are both strongly affected by the small Q: the estimated states have slow and small oscillations. On the other hand, it is obvious

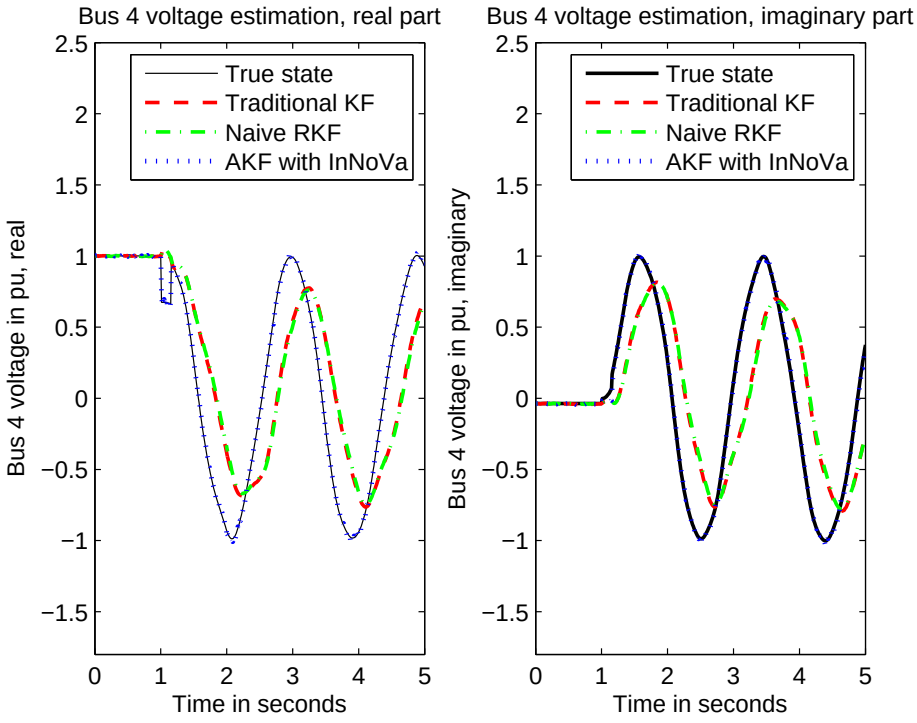
that the AKF with InNoVa tracks closely to the true state. We have observed that in as little as 5 seconds, the process noise variance at bus 6 has grown to 0.1128, which is the largest among all noise variances. This is no coincidence because the fault occurs at bus 6 exactly. Also, the average measurement noise becomes 4.8%, which is much closer to the ideal 5%.

Next, the state estimation accuracy of the three filters is evaluated using the *MAE* plots and their histograms, and plotted in Fig. 4.4(a) and Fig. 4.4(b).

Fig. 4.4 illustrated how our AKF with InNoVa has achieved the highest accuracy. Not as robust as its name might suggest, the naive RKF has even larger *MAE* values than the traditional KF after the fault occurs. The reason is that the naive RKF, being totally unaware of its poor estimating performance, is in fact largely caused by a very small Q initialization, and has dropped some measurements it falsely suspected. Hence its estimation results have drifted even further from the true states.

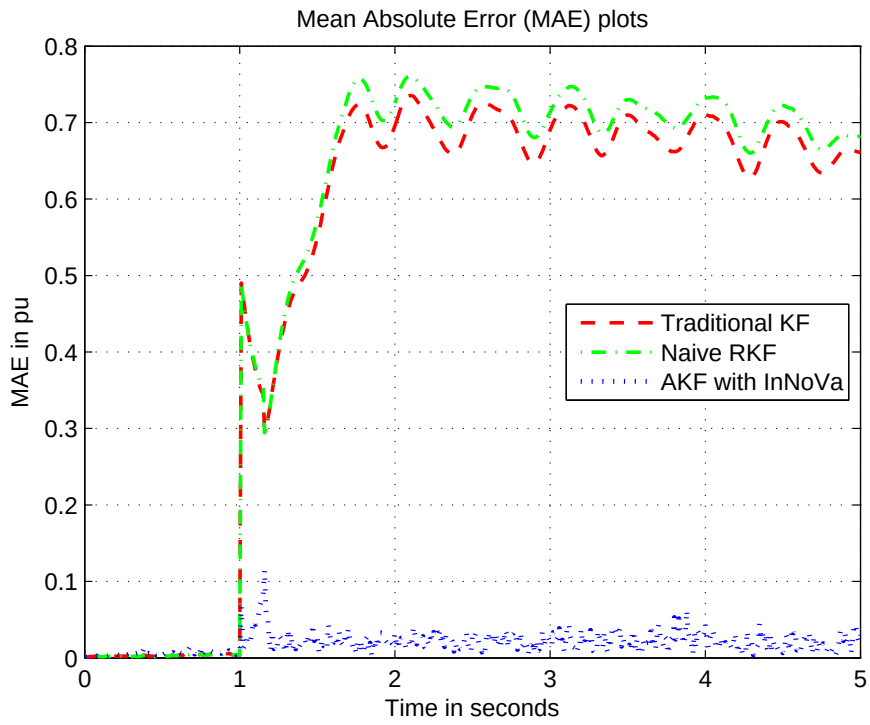


(a) Bus 4 voltage estimated by three filter in case 1

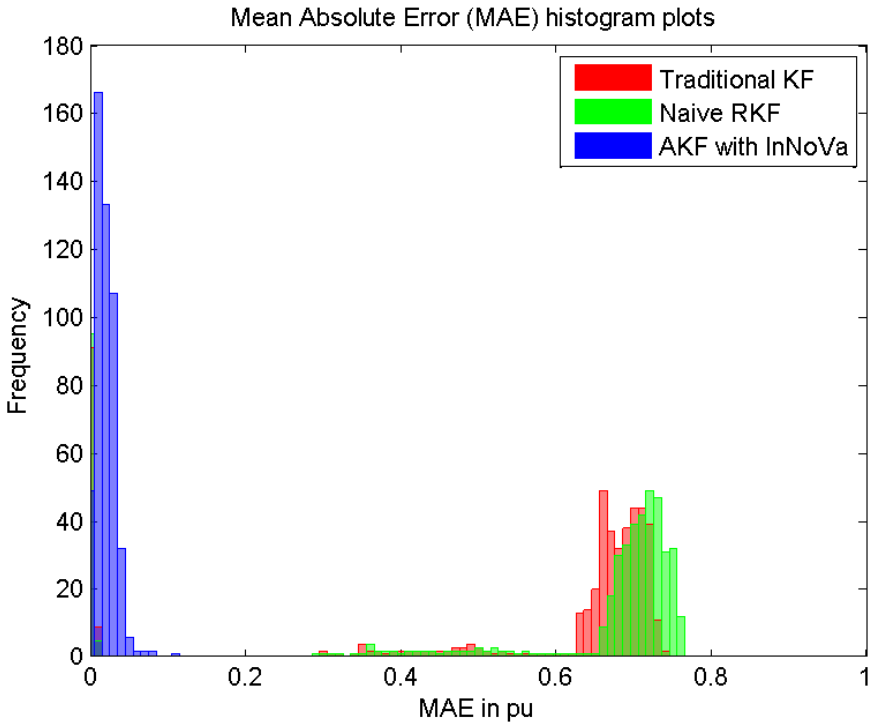


(b) Bus 4 voltage real and imaginary parts estimated by three filters in case 1

Figure 4.3: Voltage estimation results of three filters in case 1



(a) MAE comparison of three filters in case 1



(b) MAE histograms of three filters in case 1

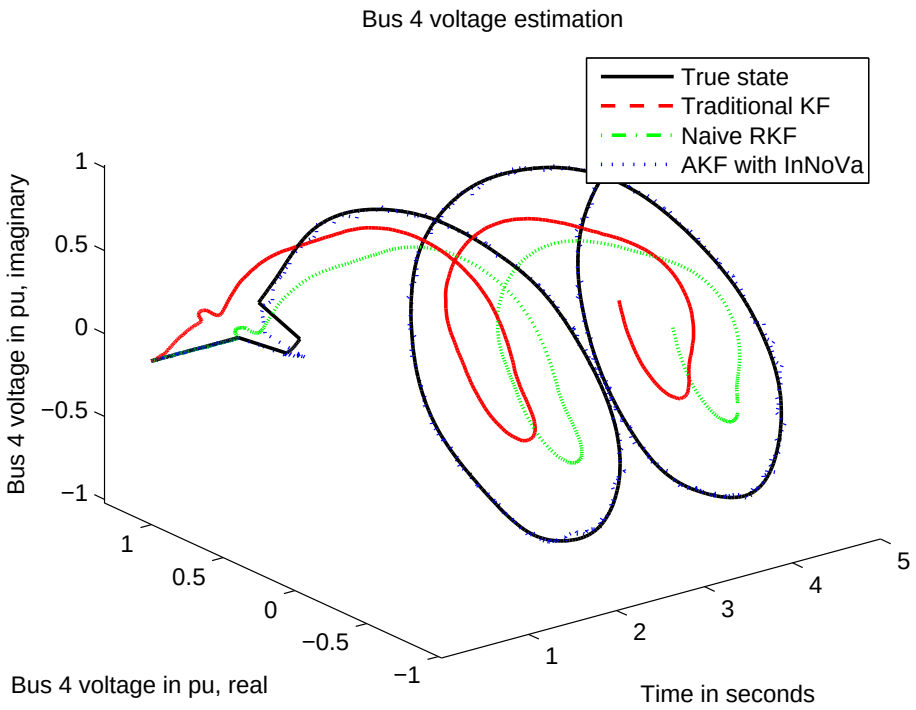
Figure 4.4: Voltage estimation accuracy of three filters in case 1

Case 2: under-valued initial Q and a malfunctioning device

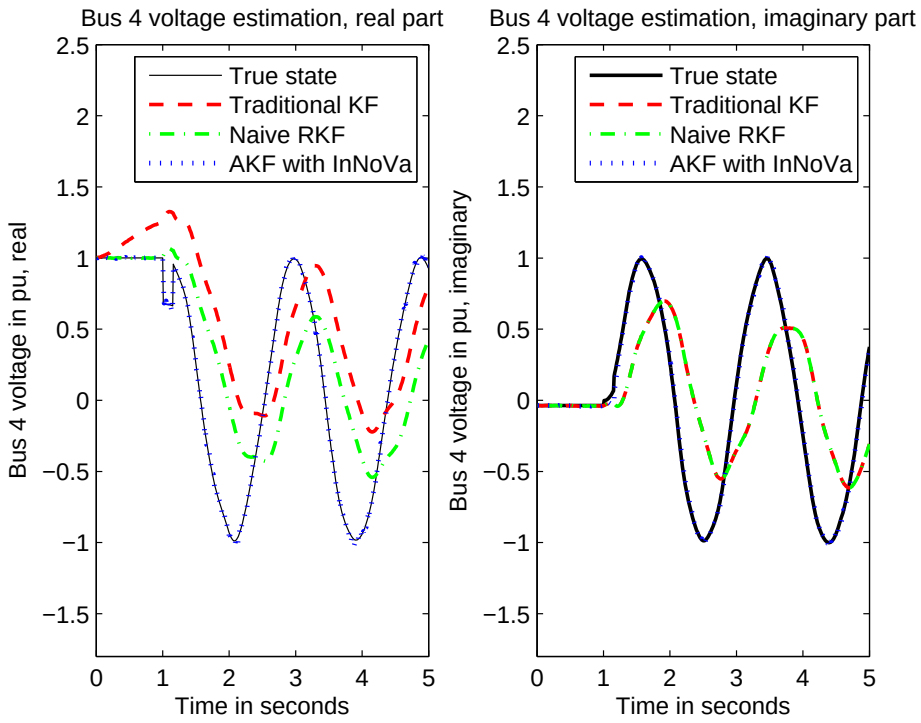
With the same small Q initialized as above, this time R is set correctly to 5%, expect one problem: a malfunctioning PMU is installed at bus 4. The bus voltages provided by this PMU always contain a systematic error, with a normal distribution $\mathcal{N}(3, 0.1^2)$. Fig. 4.5 depicts the tracking results for bus 4.

In this case, the naive RKF has identified and dropped the bad measurement. However, it was still inevitably affected by the small Q. Our proposed algorithm not only adjusts Q (the process noise variance at bus 6 has grown to 0.1126, which is still the largest process noise variance), but also identifies the bad measurement. Although our algorithm does not exclude the measurement—because we intend to avoid producing unobservable states, the noise level of this erroneous measurement becomes 110.18%, which is significantly higher than others at 5%. Hence this measurement is not trusted and has negligible impact on the state in our algorithm.

Fig. 4.6 shows that our AKF with InNoVa still offers the best performance. As one may notice, after the three-phase fault, the naive RKF has a relatively stable *MAE* level than the traditional KF, *i.e.*, it has definitely improved the $max(MAE)$ value. Yet still, the under-valued Q setting has restricted and compromised its performance.

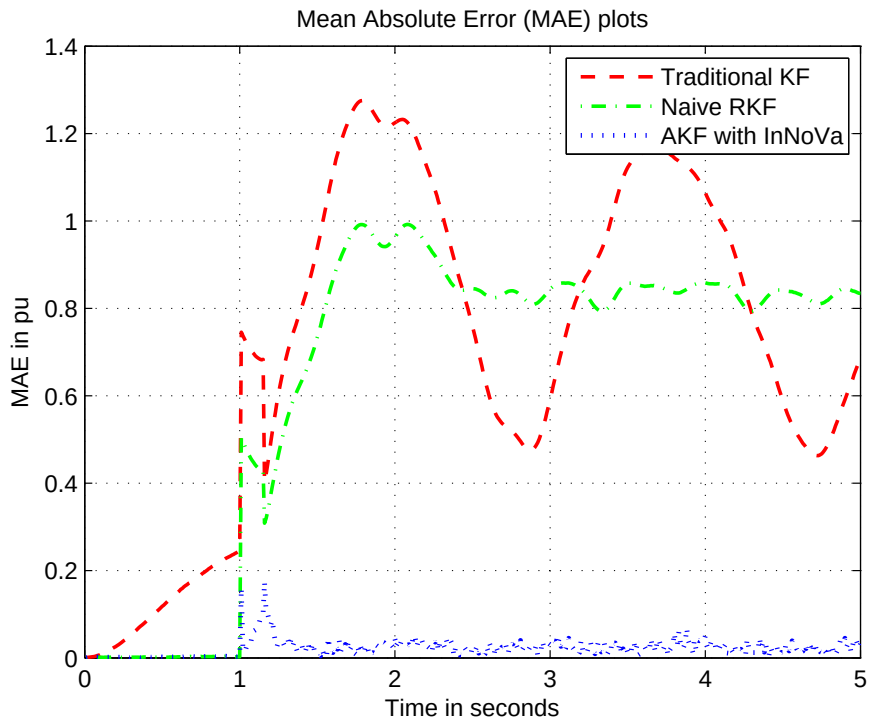


(a) Bus 4 voltage estimated by three filter in case 2

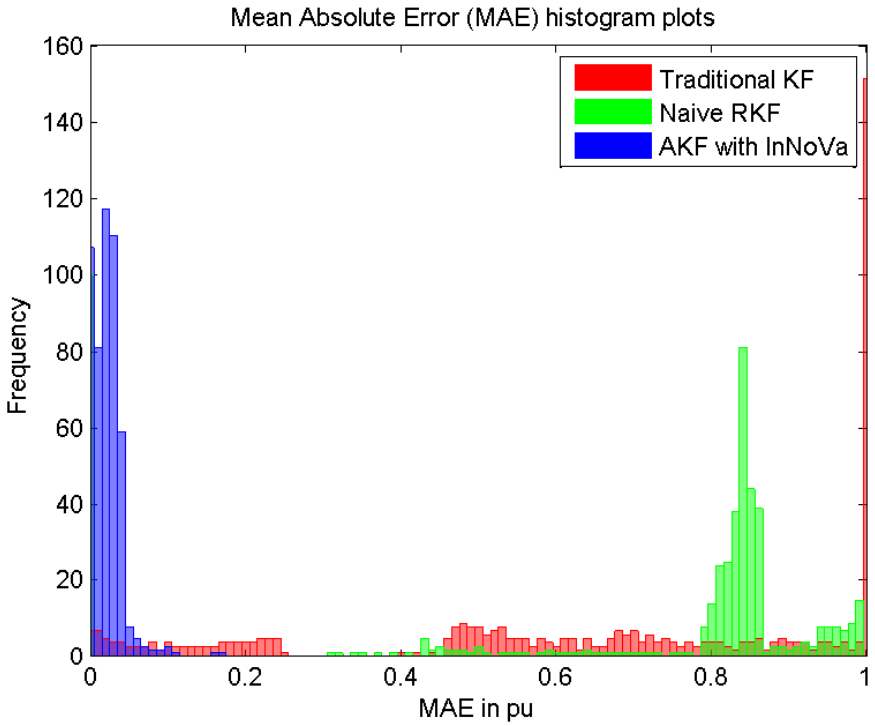


(b) Bus 4 voltage real and imaginary parts estimated by three filters in case 2

Figure 4.5: Voltage estimation results of three filters in case 1



(a) MAE comparison of three filters in case 2



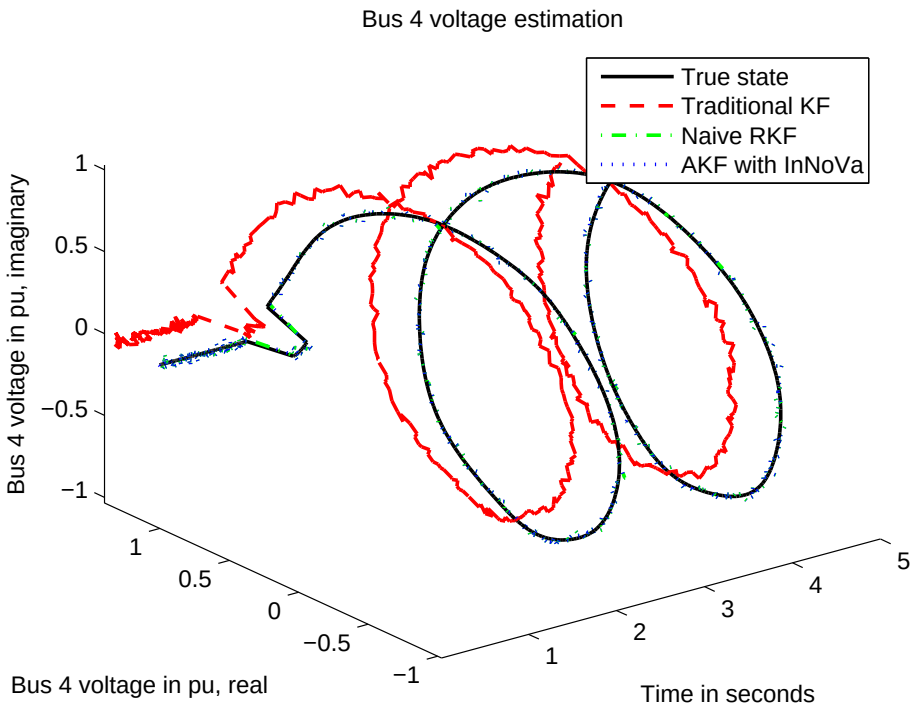
(b) MAE histograms of three filters in case 2

Figure 4.6: Voltage estimation accuracy of three filters in case 2

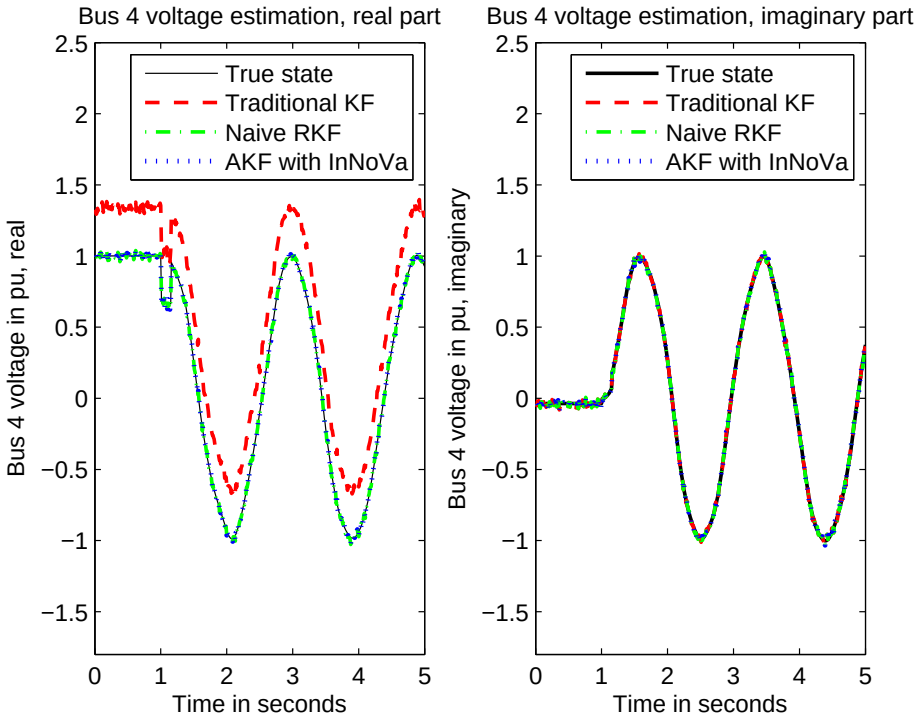
Case 3: fairly-valued initial Q and a malfunctioning device

With the same malfunctioning PMU installed at bus 4, we now have a more appropriate Q set-up: the process noise variance of each state variable is initialized to be 0.1. Fig. 4.7 displays the tracking results for bus 4.

We can tell from Fig. 4.7 that when the initial Q is appropriately scaled, the naive RKF functions properly with respect to bad data detection and correction—it can estimate almost as accurately as the AKF with InNoVa. From the histogram plot Fig. 4.8(b) we observe more clearly that the naive RKF performs quite closely to the AKF with InNoVa in this case, while the traditional KF has suffered from a “shift" caused by the “malfunctioning" PMU data.

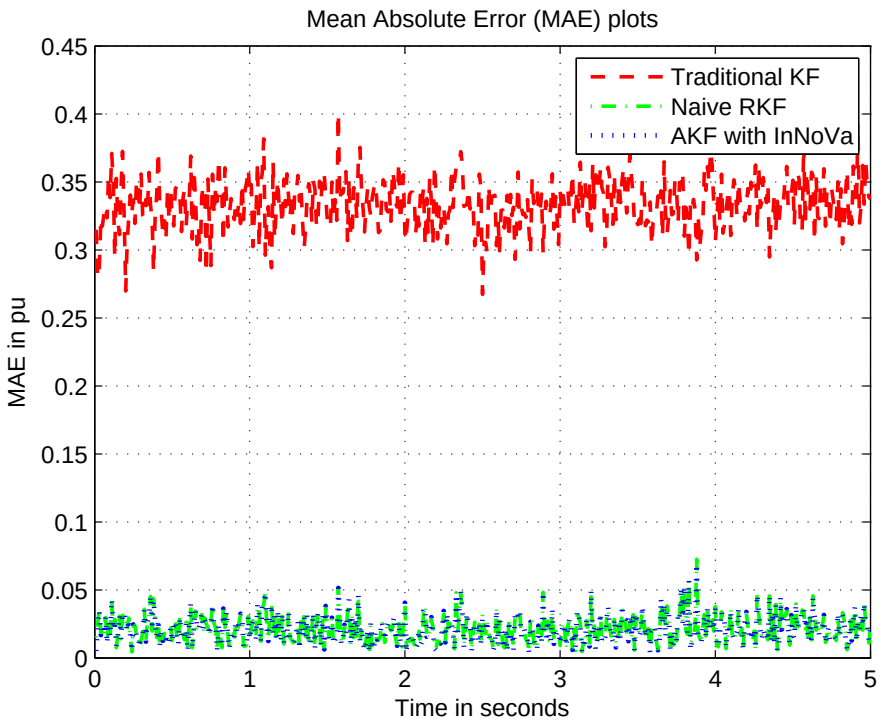


(a) Bus 4 voltage estimated by three filters in case 3

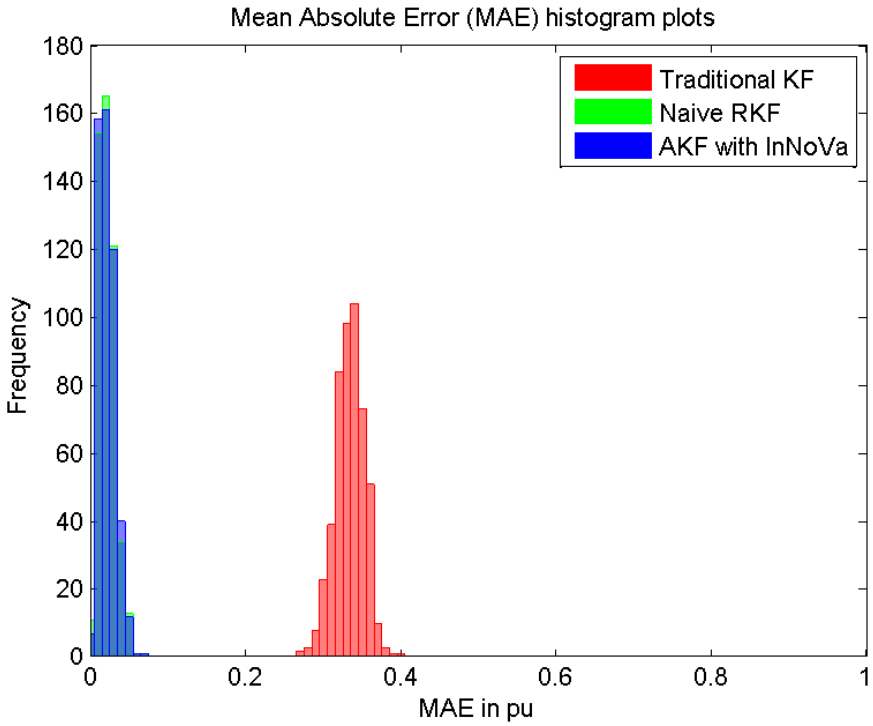


(b) Bus 4 voltage real and imaginary parts estimated by three filters in case 3

Figure 4.7: Voltage estimation results of three filters in case 3



(a) MAE comparison of three filters in case 3



(b) MAE histograms of three filters in case 3

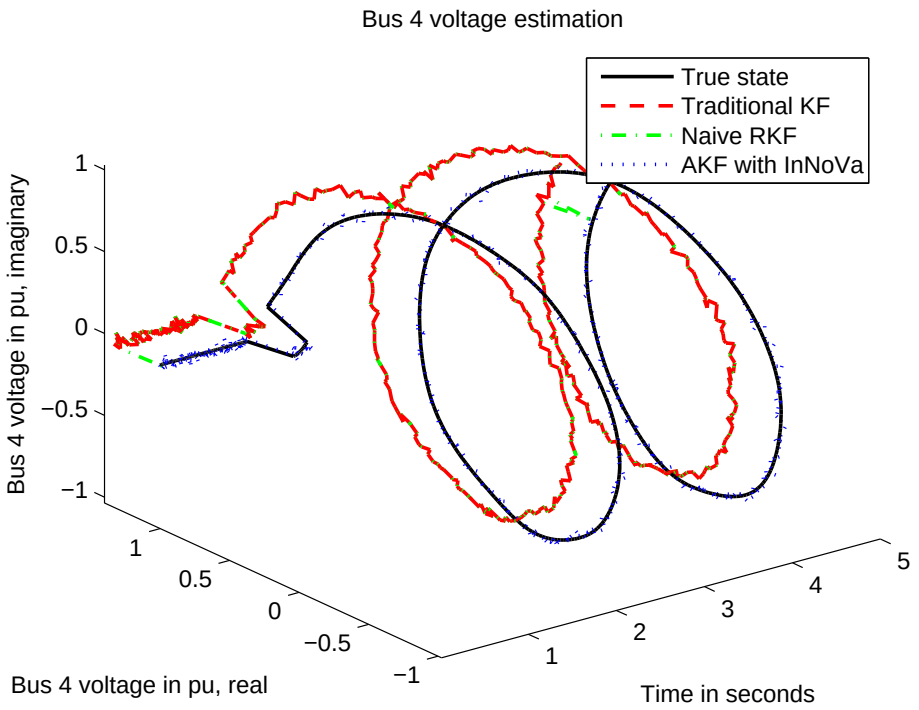
Figure 4.8: Voltage estimation accuracy of three filters in case 3

Case 4: over-valued initial Q and a malfunctioning device

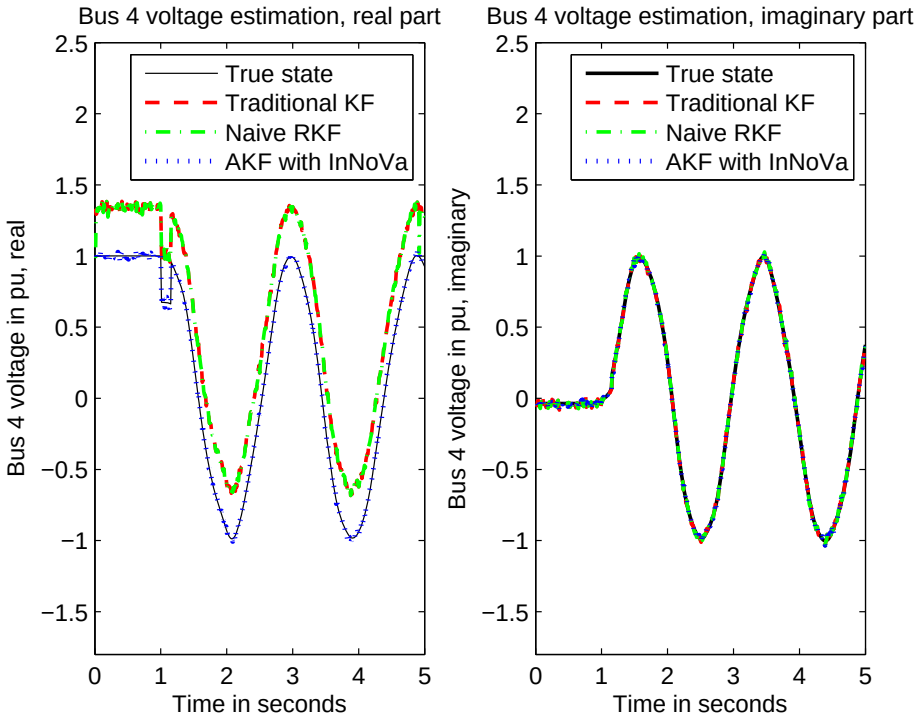
Now we increase Q significantly: the process noise variance of each state variable is initialized to be 1. The malfunctioning PMU behaves the same as in case 2 and 3. The tracking results for bus 4 are illustrated in Fig. 4.9.

Interestingly in this scenario, the naive RKF almost never detects the systematic error. Because Q is set so large (hence S is so large) that the *normalized innovations* always fall under the threshold. Nevertheless, the systematic error can not escape from the “radar" of our algorithm: the *normalized residual test*. In fact, comparing to other measurement noise at 5%, the noise of this erroneous measurement has been inflated to 110.14%. We can tell from Fig. 4.9 that the AKF with InNoVa still performs the best.

Fig. 4.10(a) and Fig. 4.10(b) further illustrated this benefit: the estimation results produced by the traditional KF and the naive RKF almost overlap with each other, while the AKF with InNoVa remains robust and reliable.

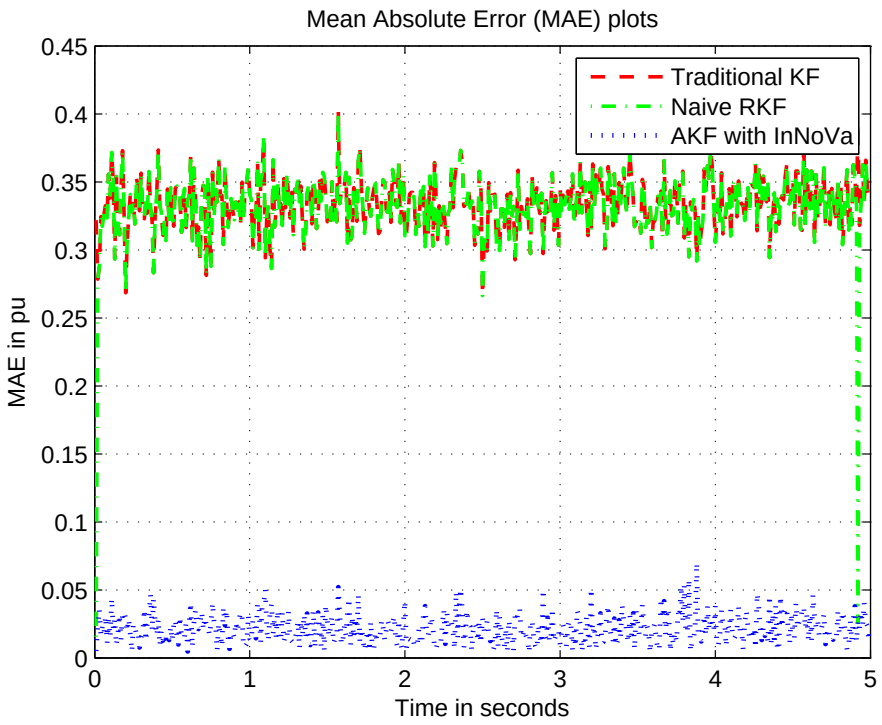


(a) Bus 4 voltage estimated by three filters in case 4

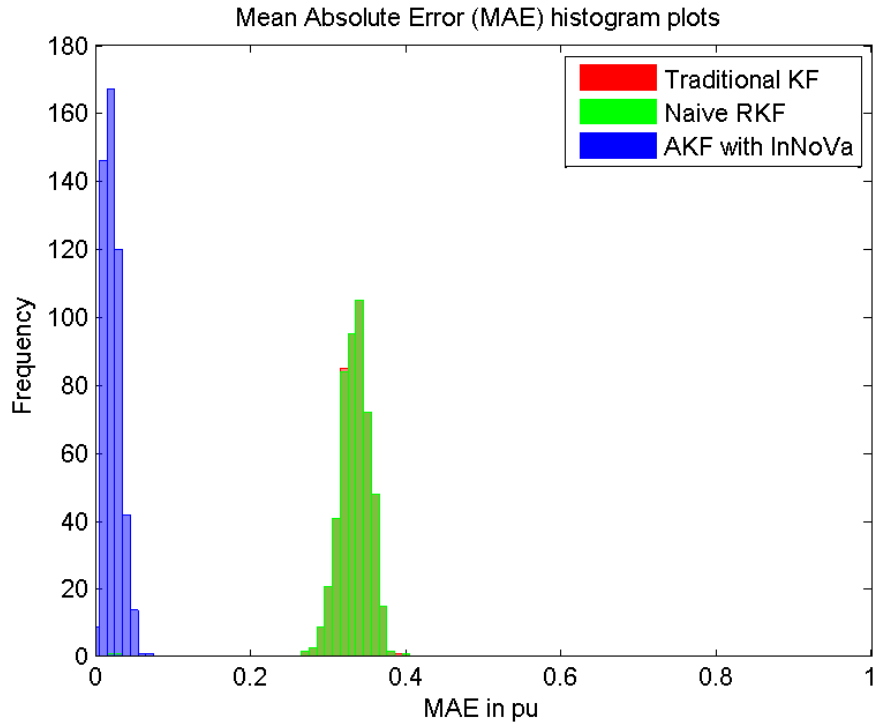


(b) Bus 4 voltage real and imaginary parts estimated by three filters in case 4

Figure 4.9: Voltage estimation results of three filters in case 4



(a) MAE comparison of three filters in case 4



(b) MAE histograms of three filters in case 4

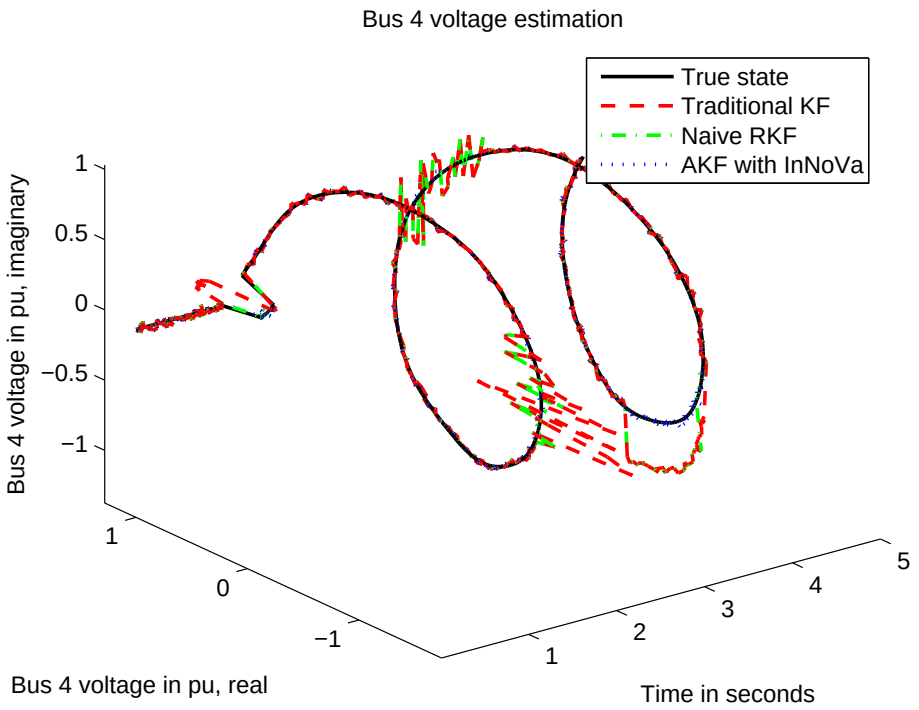
Figure 4.10: Voltage estimation accuracy of three filters in case 4

Case 5: over-valued initial Q and complex noise interference

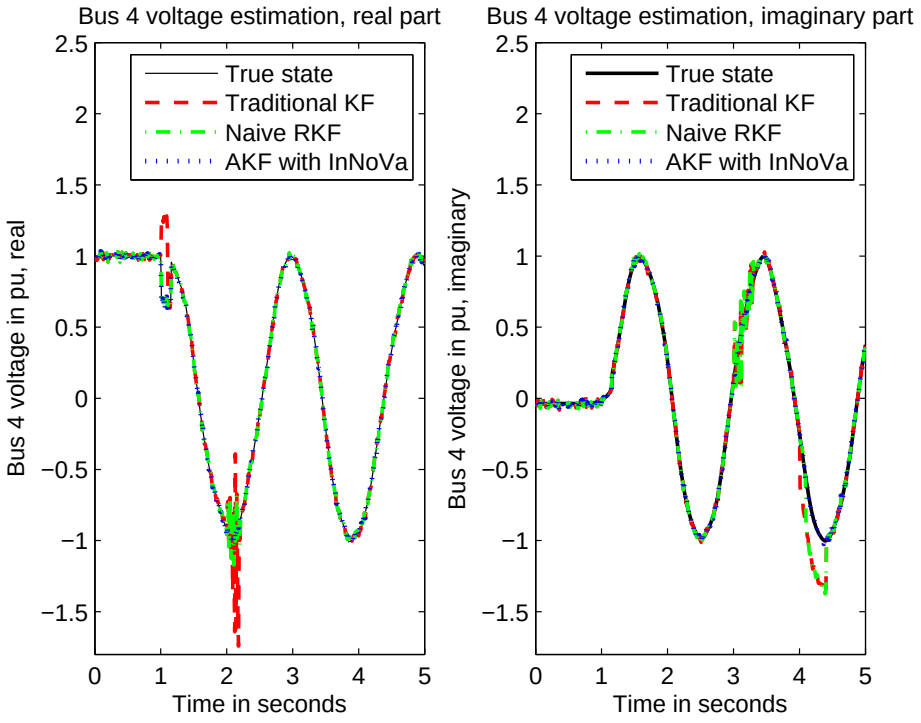
With the same large Q as in case 4, case 5 has more complicated noise interference: beyond PMU at bus 4, some adjacent PMUs also have bad measurements to affect the estimation. At $t = 1$, bus 1 voltage real-part measurements are corrupted by a constant bias value of 5 for 0.1 seconds. At $t = 2$, bus 4 voltage measurements are corrupted by the normally distributed noise $\mathcal{N}(0, 4^2)$ for 0.2 seconds. At $t = 3$, line 4-5 current measurements are corrupted by another normally distributed noise $\mathcal{N}(0, 10^2)$ for 0.3 seconds. At $t = 4$, bus 9 voltage imaginary-part measurements are corrupted by a constant bias value of -3 for 0.4 seconds. Without loss of generality, the tracking results for bus 4 can be seen in Fig. 4.11.

No doubt the traditional KF is the most vulnerable to this sort of interference. When the interference is large enough (such as at $t = 1$, which can be viewed more clearly in the left figure of Fig. 4.11(b).), the naive RKF is capable of detecting the bad measurements and calibrating the estimates; otherwise it does not help. Fortunately, the AKF with InNoVa remains robust to noise interference. Moreover, the inflated R provides useful information for people to inspect and repair devices, as the measurement noise variance of bus 1, 4, 9 voltages and line 4-5 current is abnormally larger than others.

Fig. 4.12(a) and Fig. 4.12(b) both verified that the AKF with InNoVa has the most consistent and highest accuracy level, which leads to its most outstanding performance.

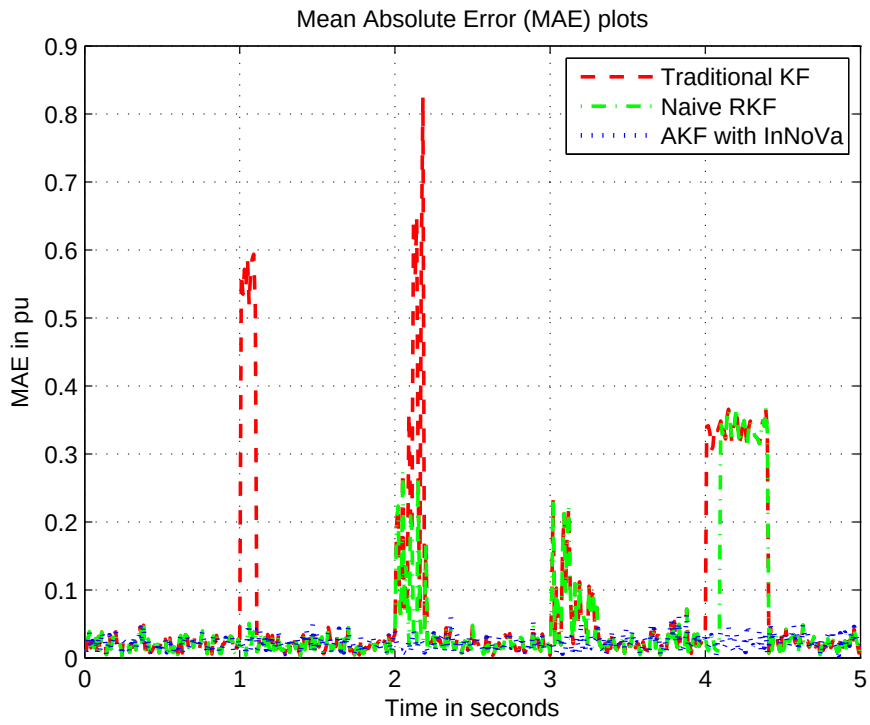


(a) Bus 4 voltage estimated by three filters in case 5

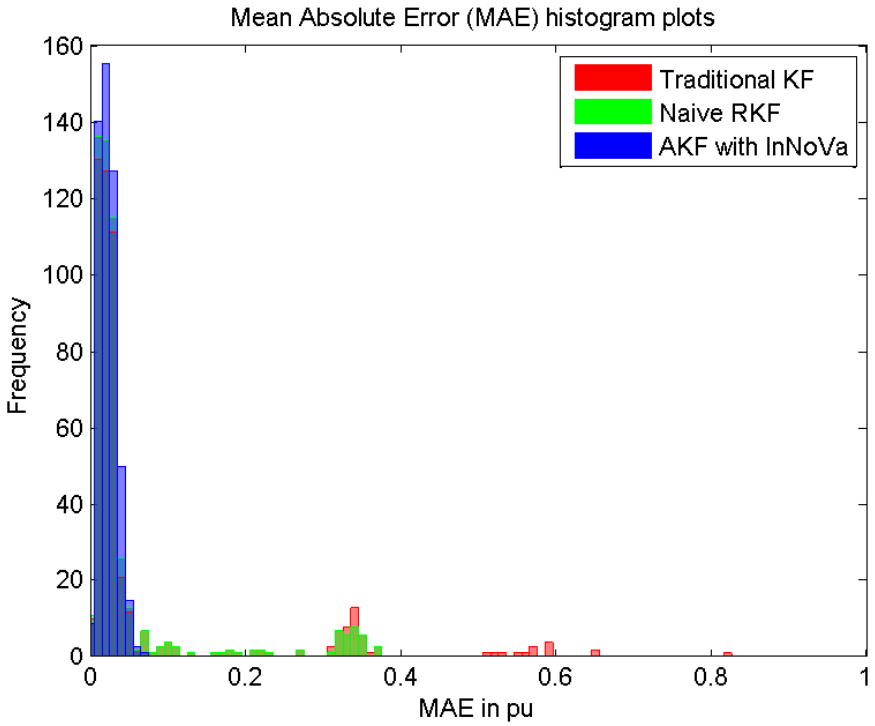


(b) Bus 4 voltage real and imaginary parts estimated by three filters in case 5

Figure 4.11: Voltage estimation results of three filters in case 5



(a) MAE comparison of three filters in case 5



(b) MAE histograms of three filters in case 5

Figure 4.12: Voltage estimation accuracy of three filters in case 5

Case 6: failure case

When introducing the typical steps in any state estimation procedure in Section 3.1, we have mentioned that in contrast to *critical measurements*, "*redundant measurements* can be removed without affecting the system observability". Now it is worth emphasizing that although the removal of redundant measurements will not affect observability, this does not mean these measurements are not useful or unnecessary.

In fact, redundant measurements are important in state estimation procedures. Without enough measurement redundancy, any algorithm will be vulnerable to measurement errors or telemetry failure. For this reason, in practice state estimation procedures make use of a set of redundant measurements to filter out such errors and find an optimal estimate.

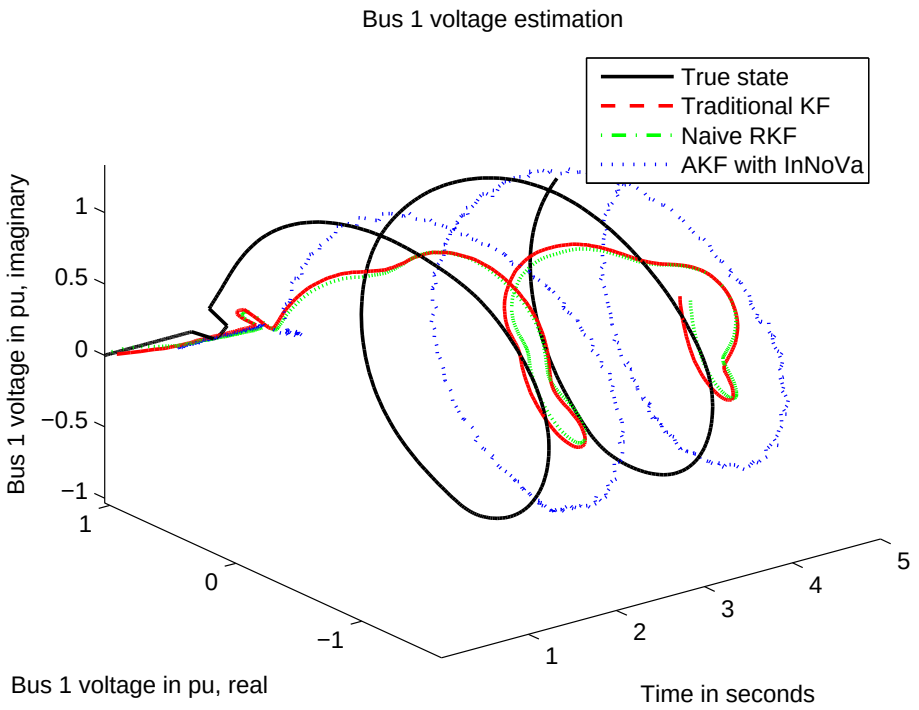
Although our AKF with InNoVa demonstrates its robustness in many test cases, we should still be aware that it is not a "silver bullet" if the redundancy of measurements is insufficient. When multiple interacting bad data are conforming with each other, *i.e.* multiple interacting measurements have errors that are coincidentally in agreement, then the AKF with InNoVa might fail to identify this bad data, and consequently produce incorrect estimation results.

For example in this failure case, we have line 1-4 current measurements (provided by the PMU at bus 1) which contain a constant error of $5 + 10i$, and accidentally, line 4-1 current measurements (provided by PMU at bus 4) contain a constant error of $-5 - 10i$. Furthermore, the process noise variance Q is under-valued like in case 1 and 2. Because of inadequate redundancy of measurements at bus 1, the good bus 1 voltage measurements will be falsely identified as the bad measurements by the AKF with InNoVa, hence the estimates of bus 1 voltages will largely deviate from the true state, as shown in Fig. 4.13.

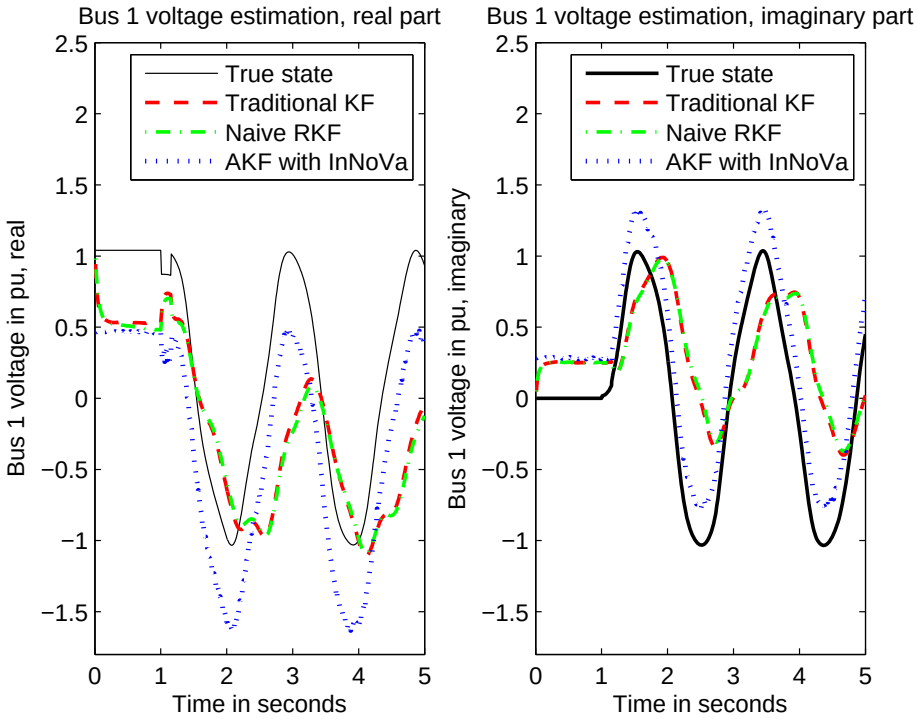
Note that since the interacting bad data problem is quite local, it does not affect the performance of the AKF with InNoVa at other buses significantly (see Fig. 4.14), however at bus

1 all three algorithms are strongly affected by this particular problem. An obvious solution would be to increase the redundancy, *e.g.* to place more PMUs, or to incorporate traditional RTU measurements in the estimation process, *etc*.

Similar to that in case 1, Naive RKF has again misjudged some measurements, due to the overly small Q initialization. As for the AKF with InNoVa, comparing to Fig. 4.4, 4.6, 4.8, 4.10 and 4.12 from previous cases, we can tell its *MAE* level has been definitely raised, which is mostly caused by the bad estimating at bus 1.

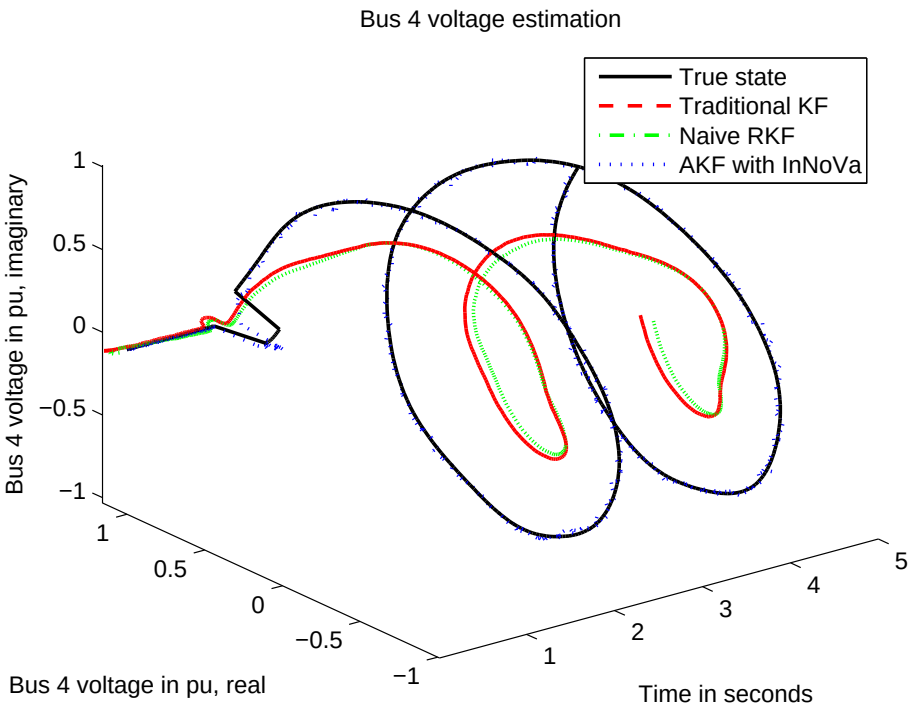


(a) Bus 1 voltage estimated by three filters in case 6

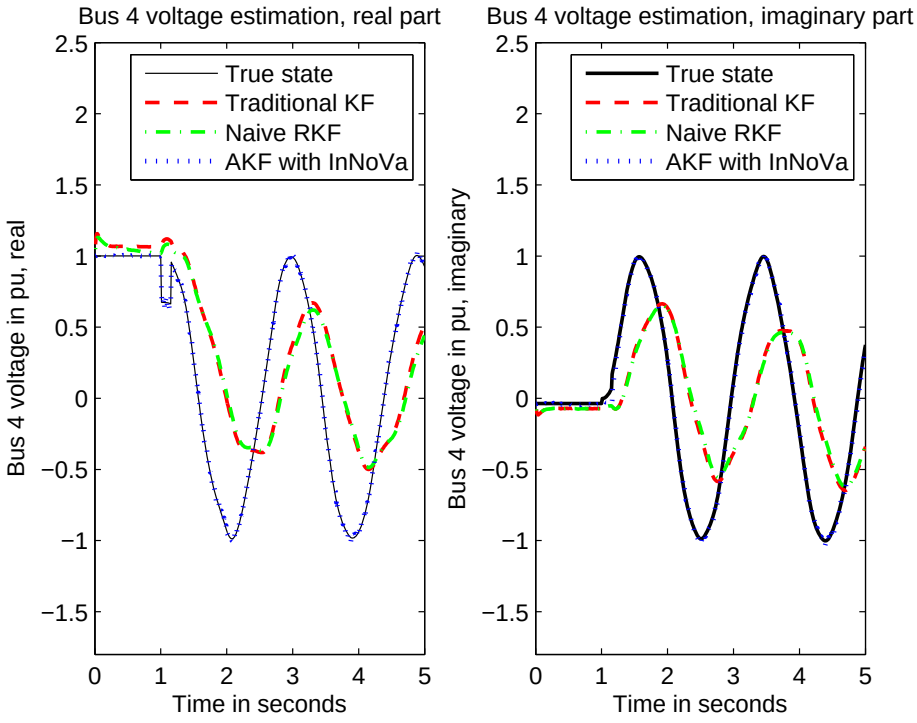


(b) Bus 1 voltage real and imaginary parts estimated by three filters in case 6

Figure 4.13: Bus 1 voltage estimation results of three filters in case 6

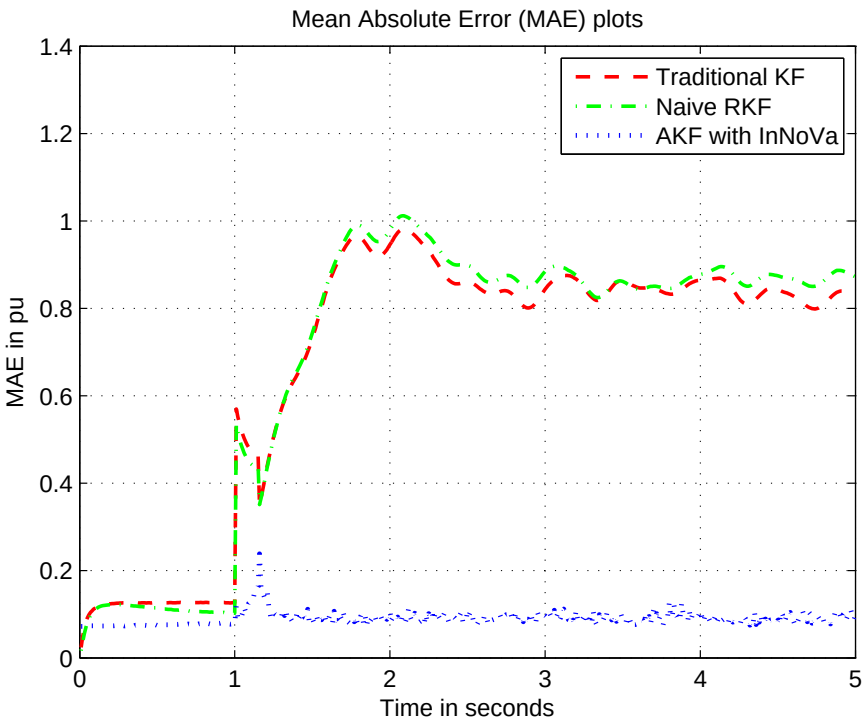


(a) Bus 4 voltage estimated by three filters in case 6

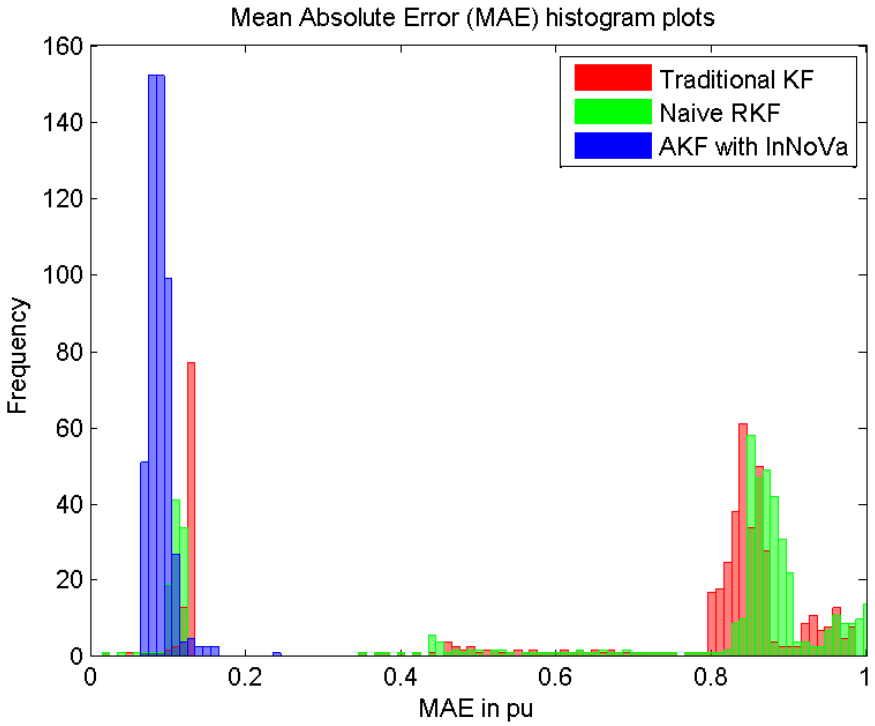


(b) Bus 4 voltage real and imaginary parts estimated by three filters in case 6

Figure 4.14: Bus 4 voltage estimation results of three filters in case 6



(a) MAE comparison of three filters in case 6



(b) MAE histograms of three filters in case 6

Figure 4.15: Voltage estimation accuracy of three filters in case 6

4.6 My Method: A Two-stage Kalman Filter Approach for Robust and Real-time Power Systems State Tracking

A quasi-steady-state assumption is typically applied to power system 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—electro-mechanical 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 and enables operation on standard computers to be feasible within the required operational time intervals. The problem with this assumption is that many studies cannot be performed in the operational environment. Development of the smart grid makes the grid 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. Encouraged by 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 continue at an accelerated speed. This will result in stochastic operating behaviors and dynamics the grid has never seen nor been designed for.

In this section we developed a two-stage Kalman filter approach to estimate the *traditional* states of bus voltage phasor, as well as the *dynamic* states of generator rotor angle and generator speed. Theoretically, if we have knowledge of the system noise characteristics, and they have desirable statistical properties including zero-mean, Gaussian, and spectrally white, then Kalman filters can achieve its optimal performance. In practice however, the system noise characteristics usually remain unknown to us, and may change from time to time. Thus an adaptive Kalman filter algorithm becomes favorable in such cases.

In the first stage, we estimate the *traditional* states from raw PMU measurements, using

the Adaptive Kalman Filter with Inflatable Noise Variances (AKF with InNoVa) we have proposed in Section 4.5. As mentioned previously, it is able to 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 estimation. Simulations demonstrate its robustness to sudden changes of system dynamics and erroneous PMU measurements.

4.6.1 Stage One

In stage one, the inputs are 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 hypothetical system models.

Ideally, the hypothetical models are the “true" models, with accurate noise statistical characteristics. However when the hypothetical models do not match the actual models, and/or PMU measurements contain significant errors, the estimated states could deviate from the true states rapidly. Thus we apply the AKF with InNoVa in the first stage to deal with unknown system dynamics and erroneous PMU measurements. The detailed implementation is described in Section 4.5.

4.6.2 Stage Two

In stage two, the estimation results from stage one are fed directly into an Extended Kalman Filter (EKF) as “measurements". The outputs 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 are modified as follows.

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

$$\begin{cases} \frac{d\delta_i}{dt} & = \omega_B(\omega_i - \omega_0) \\ \frac{d\omega_i}{dt} & = \frac{\omega_0}{2H_i}(P_{m_i} - \frac{|E_i|}{X'_{d_i}}\sin\delta_i|V_i|\cos\theta_i + \frac{|E_i|}{X'_{d_i}}\cos\delta_i|V_i|\sin\theta_i - D_i(\omega_i - \omega_0)) \\ & = \frac{\omega_0}{2H_i}(P_{m_i} - \frac{|E_i|}{X'_{d_i}}\sin\delta_i Re(V_i) + \frac{|E_i|}{X'_{d_i}}\cos\delta_i Im(V_i) - D_i(\omega_i - \omega_0)) \end{cases} \tag{4.24}$$

where state variables δ_i and ω_i are the generator rotor angle and speed respectively, ω_B and ω_0 are the speed base and the synchronous speed in per-unit quantities, P_{m_i} is the mechanical input, H_i is the machine inertia, 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 $\delta_1, \delta_2, ..., \delta_n$).

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

$$\frac{dx}{dt} = A_c x + w_c, \tag{4.25}$$

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 $i \in \{1, ..., n\}$ and $j \in \{1, ..., n\}$ $(i \neq j)$ are the corresponding partial derivatives

$$A_{c_{[2i-1,2i-1]}} = 0 \tag{4.26}$$

$$A_{c_{[2i-1,2i]}} = \omega_B \tag{4.27}$$

$$A_{c_{[2i,2i-1]}} = -\frac{\omega_0 |E_i|}{2H_i X'_{d_i}} \left[\cos\delta_i Re(V_i) + \sin\delta_i \frac{\partial Re(V_i)}{\partial \delta_i} + \sin\delta_i Im(V_i) - \cos\delta_i \frac{\partial Im(V_i)}{\partial \delta_i} \right] \tag{4.28}$$

$$A_{c_{[2i,2i]}} = -\frac{\omega_0}{2H_i} D_i \tag{4.29}$$

$$A_{c_{[2i-1,2j-1]}} = 0 \tag{4.30}$$

$$A_{c_{[2i-1,2j]}} = 0 \tag{4.31}$$

$$A_{c_{[2i,2j-1]}} = -\frac{\omega_0 |E_i|}{2H_i X'_{d_i}} \left[\sin\delta_i \frac{\partial Re(V_i)}{\partial \delta_j} - \cos\delta_i \frac{\partial Im(V_i)}{\partial \delta_j} \right] \tag{4.32}$$

$$A_{c_{[2i,2j]}} = 0. \tag{4.33}$$

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

$$A = I + A_c \cdot \Delta t. \tag{4.34}$$

The discrete-time process noise covariance Q can be formulated by integrating the continuous time process equation (4.25) over the time interval Δt as described before:

$$Q = \int_0^{\Delta t} e^{A_c t} Q_c e^{A_c^T t} dt \tag{4.35}$$

Hence the process model is written as

$$x_k = A x_{k-1} + w_{k-1} \tag{4.36}$$

$$= (I + A_c \cdot \Delta t) x_{k-1} + w_{k-1}, \tag{4.37}$$

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

The Measurement Model

The expanded system nodal equation can be expressed as:

$$Y_{exp}\begin{pmatrix} E \\ V \end{pmatrix} = \begin{pmatrix} Y_{GG} & Y_{GL} \\ Y_{LG} & Y_{LL} \end{pmatrix}\begin{pmatrix} E \\ V \end{pmatrix} = \begin{pmatrix} I_G \\ 0 \end{pmatrix}, \tag{4.38}$$

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:

$$V = -Y_{LL}^{-1}Y_{LG}E = R_V E \tag{4.39}$$

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

$$\begin{pmatrix} V_1 \\ V_2 \\ \vdots \\ V_m \end{pmatrix} = \begin{pmatrix} R_{V_{11}} & R_{V_{12}} & \cdots & R_{V_{1n}} \\ R_{V_{21}} & R_{V_{22}} & \cdots & R_{V_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ R_{V_{m1}} & R_{V_{m2}} & \cdots & R_{V_{mn}} \end{pmatrix}\begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_n \end{pmatrix}, \tag{4.40}$$

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

In rectangular form, we have

$$\begin{aligned} Re(V_i) = {} & |R_{V_{i1}}||E_1|\cos(\angle R_{V_{i1}} + \delta_1) + |R_{V_{i2}}||E_2|\cos(\angle R_{V_{i2}} + \delta_2) + \cdots \\ & + |R_{V_{in}}||E_n|\cos(\angle R_{V_{in}} + \delta_n) \end{aligned} \tag{4.41}$$

$$\begin{aligned} Im(V_i) = {} & |R_{V_{i1}}||E_1|\sin(\angle R_{V_{i1}} + \delta_1) + |R_{V_{i2}}||E_2|\sin(\angle R_{V_{i2}} + \delta_2) + \cdots \\ & + |R_{V_{in}}||E_n|\sin(\angle R_{V_{in}} + \delta_n) \end{aligned} \tag{4.42}$$

Those “measurements" obtained from stage-one AKF can now be written in this form:

$$z = h(x) + v \tag{4.43}$$

where $z = [Re(V_1), Im(V_1), Re(V_2), Im(V_2), ..., Re(V_m), Im(V_m)]^T$ is the $2m \times 1$ measurement vector that consists of real and imaginary parts of each estimated bus voltage, $x = [\delta_1, \omega_1, \delta_2, \omega_2, ..., \delta_n, \omega_n]^T$ is the state vector and v is the normally distributed measurement noise with covariance P estimated by stage-one AKF.

After linearization, the measurement model now has this form

$$z = Hx + v \tag{4.44}$$

where H is the corresponding Jacobian matrix:

$$H = \begin{pmatrix} \frac{\partial Re(V_1)}{\partial \delta_1} & \frac{\partial Re(V_1)}{\partial \omega_1} & \cdots & \frac{\partial Re(V_1)}{\partial \delta_n} & \frac{\partial Re(V_1)}{\partial \omega_n} \\ \frac{\partial Im(V_1)}{\partial \delta_1} & \frac{\partial Im(V_1)}{\partial \omega_1} & \cdots & \frac{\partial Im(V_1)}{\partial \delta_n} & \frac{\partial Im(V_1)}{\partial \omega_n} \\ \frac{\partial Re(V_2)}{\partial \delta_1} & \frac{\partial Re(V_2)}{\partial \omega_1} & \cdots & \frac{\partial Re(V_2)}{\partial \delta_n} & \frac{\partial Re(V_2)}{\partial \omega_n} \\ \frac{\partial Im(V_2)}{\partial \delta_1} & \frac{\partial Im(V_2)}{\partial \omega_1} & \cdots & \frac{\partial Im(V_2)}{\partial \delta_n} & \frac{\partial Im(V_2)}{\partial \omega_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial Re(V_m)}{\partial \delta_1} & \frac{\partial Re(V_m)}{\partial \omega_1} & \cdots & \frac{\partial Re(V_m)}{\partial \delta_n} & \frac{\partial Re(V_m)}{\partial \omega_n} \\ \frac{\partial Im(V_m)}{\partial \delta_1} & \frac{\partial Im(V_m)}{\partial \omega_1} & \cdots & \frac{\partial Im(V_m)}{\partial \delta_n} & \frac{\partial Im(V_m)}{\partial \omega_n} \end{pmatrix} \tag{4.45}$$

$$= \begin{pmatrix} -|R_{V_{11}}||E_1|\sin(\angle R_{V_{11}} + \delta_1) & 0 & \cdots & -|R_{V_{1n}}||E_n|\sin(\angle R_{V_{1n}} + \delta_n) & 0 \\ |R_{V_{11}}||E_1|\cos(\angle R_{V_{11}} + \delta_1) & 0 & \cdots & |R_{V_{1n}}||E_n|\cos(\angle R_{V_{1n}} + \delta_n) & 0 \\ -|R_{V_{21}}||E_1|\sin(\angle R_{V_{21}} + \delta_1) & 0 & \cdots & -|R_{V_{2n}}||E_n|\sin(\angle R_{V_{2n}} + \delta_n) & 0 \\ |R_{V_{21}}||E_1|\cos(\angle R_{V_{21}} + \delta_1) & 0 & \cdots & |R_{V_{2n}}||E_n|\cos(\angle R_{V_{2n}} + \delta_n) & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -|R_{V_{m1}}||E_1|\sin(\angle R_{V_{m1}} + \delta_1) & 0 & \cdots & -|R_{V_{mn}}||E_n|\sin(\angle R_{V_{mn}} + \delta_n) & 0 \\ |R_{V_{m1}}||E_1|\cos(\angle R_{V_{m1}} + \delta_1) & 0 & \cdots & |R_{V_{mn}}||E_n|\cos(\angle R_{V_{mn}} + \delta_n) & 0 \end{pmatrix}$$

4.6.3 Simulation Results

Our two-stage Kalman filter approach has been simulated on several multi-machine systems in various situations. For a more realistic and comprehensive study, this time we will verify the feasibility of our method on the 16-generator-68-bus New England Test System and New York Power System model (see Fig. 2.6) from Section 2.8.3, under different abnormal conditions.

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.

To make it closer to real-world applications, we assume that PMU data contain 1% random noise. The latest IEEE Standard for Synchrophasor Measurements for Power Systems, *IEEE C37.118.1-2011* [55], has established a criterion of 1% for the value of *TVE* during calibration. That means that if the PMU is to be deemed compliant, the observed samples should not lie outside the circle shown in Fig. 4.2, so that the values calculated by substitution into Equation (4.22) do not exceed 1%.

In the following five subsections, five abnormal cases are simulated to illustrate the robustness of our proposed algorithm. In each case we executed three two-stage Kalman filter approaches for a more informative comparison: Approach 1 uses the traditional KF at stage one, while Approach 2 uses the naive robust KF (RKF) (using largest normalized innovation test to identify and exclude bad measurements, without adjusting model parameters) and Approach 3 uses the AKF with InNoVa.

For each case at stage one, without loss of generality, we visualize the voltage tracking results for bus 22 in three-dimensional plots. Then the complex voltage *MAE* and their histogram plots are illustrated for a more comprehensive comparison. To maintain consistency of our experiments in this subsection, the data range of each histogram is set to be $[0, 0.15]$, and

the bin (*i.e.* interval) size is set to be 0.001.

At stage two, the outputs of these three filters are processed by separate EKFs to estimate the system's dynamic states. Again without loss of generality, for each case we chose generator 6 to demonstrate and compare the dynamic states tracking results. We have also plotted the *MAE* plots and their histograms for generator speed and generator angle estimates separately. For generator speed, the data range of each histogram is set to be $[0, 0.002]$, and the bin (*i.e.* interval) size is set to be $1e^{-5}$. For generator angles, the data range of each histograms is set to be $[0, 3.5]$, and the bin (*i.e.* interval) size is set to be 0.01.

These test cases are stated in the following subsections respectively.

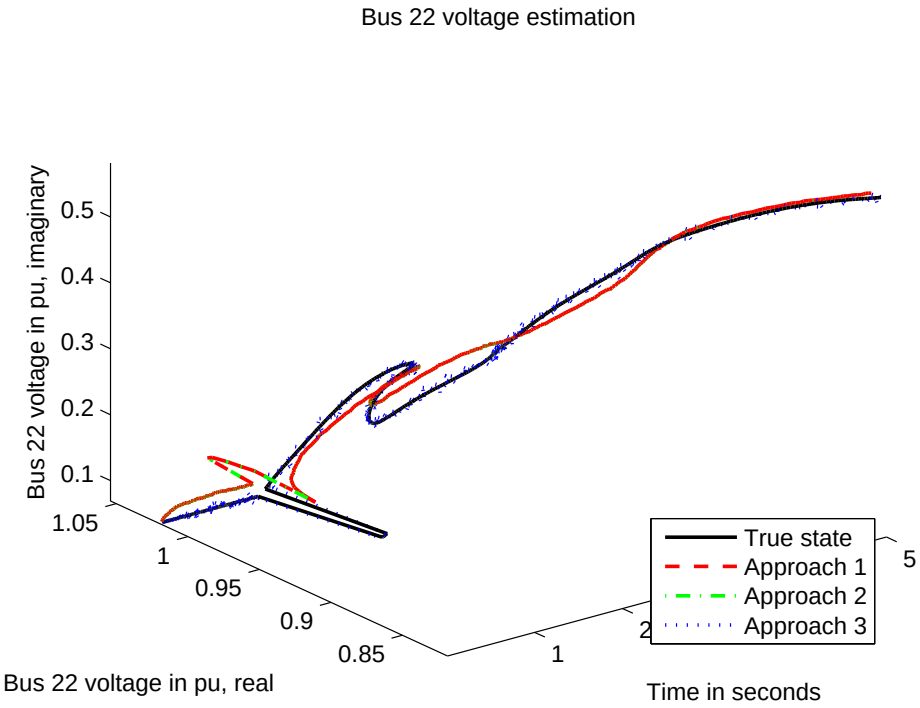
Case 1: under-valued initial Q and R

We first initialize each state variable with a small variance ($2.5e^{-5}$), meaning that the user assumes the system is stable. Then we initialize the PMU measurement noise to 1%, but simulate actual PMU measurement noise at 2%. In this and the following simulations, we always use a black solid line for the true state, a red dash-dash line for Approach 1, a green dash-dot line for Approach 2, and a blue dot-dot line for Approach 3—our proposed approach.

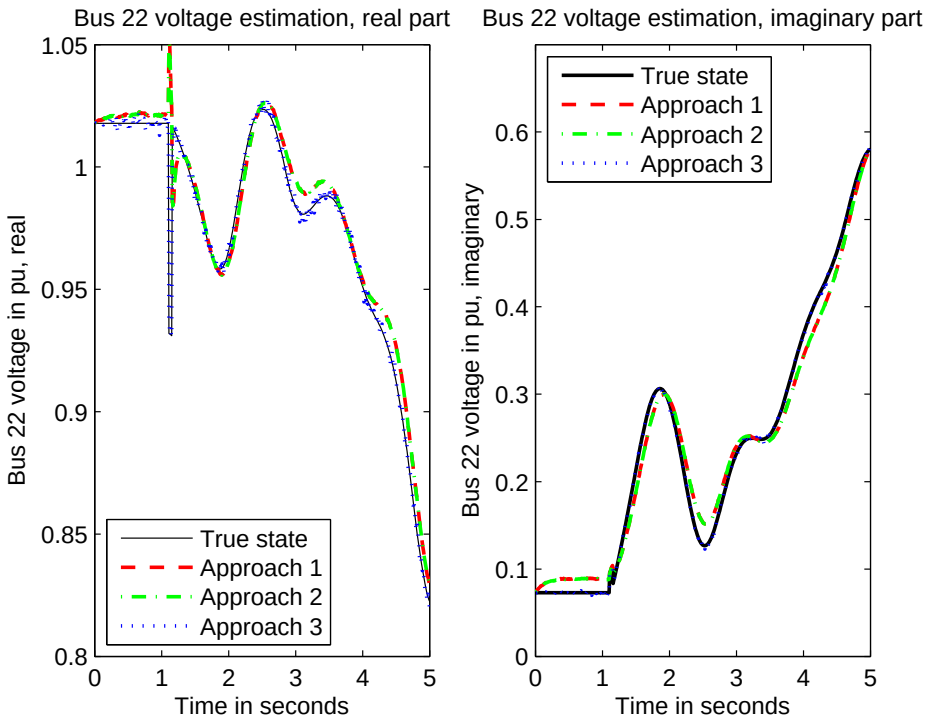
Fig. 4.16 shows the tracking results of the three approaches after stage one. Similar to those in Fig. 4.3, the small Q has dominated the behavior of the KF and the naive RKF, causing them to be overconfident with the quasi-static model, where the system states have slow and small oscillations. Yet by adjusting Q and R, 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 occurs at bus 29. The average measurement noise has grown to 1.4% in 5 seconds. The accuracy evaluation of stage one estimates can be found in Fig. 4.17.

Fig. 4.18 illustrates the corresponding dynamic state estimates of the three approaches,

which are obtained after stage two. Fig. 4.19 has displayed the *MAE* plots and their histogram plots respectively, for the generator speed estimates; while Fig. 4.20 has displayed those for the generator angle estimates. Naturally following the previous stage, Approach 3 still outperforms the others.

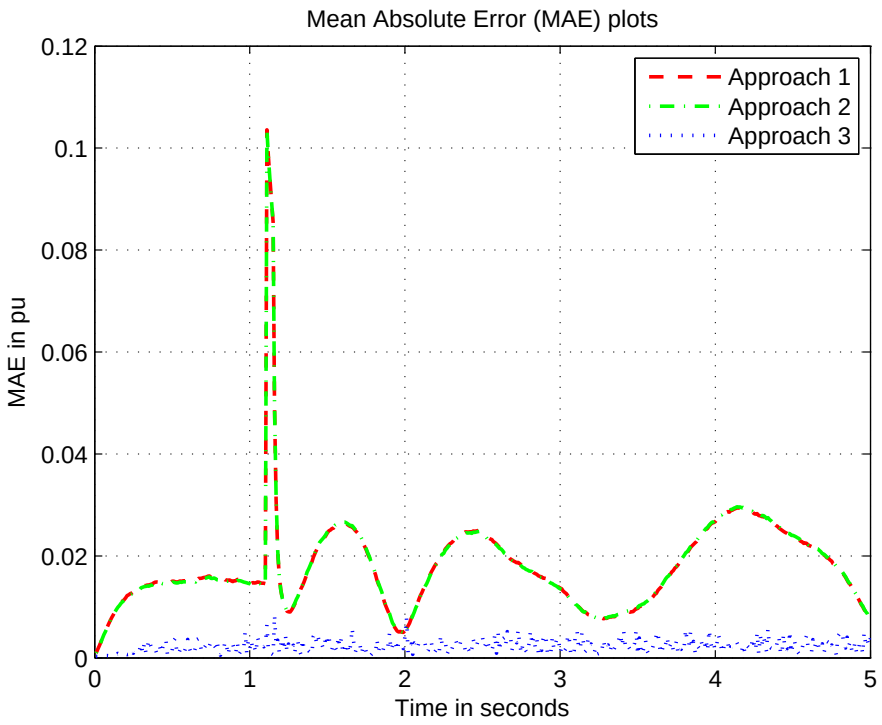


(a) Bus 22 voltage estimated by three approaches after stage one

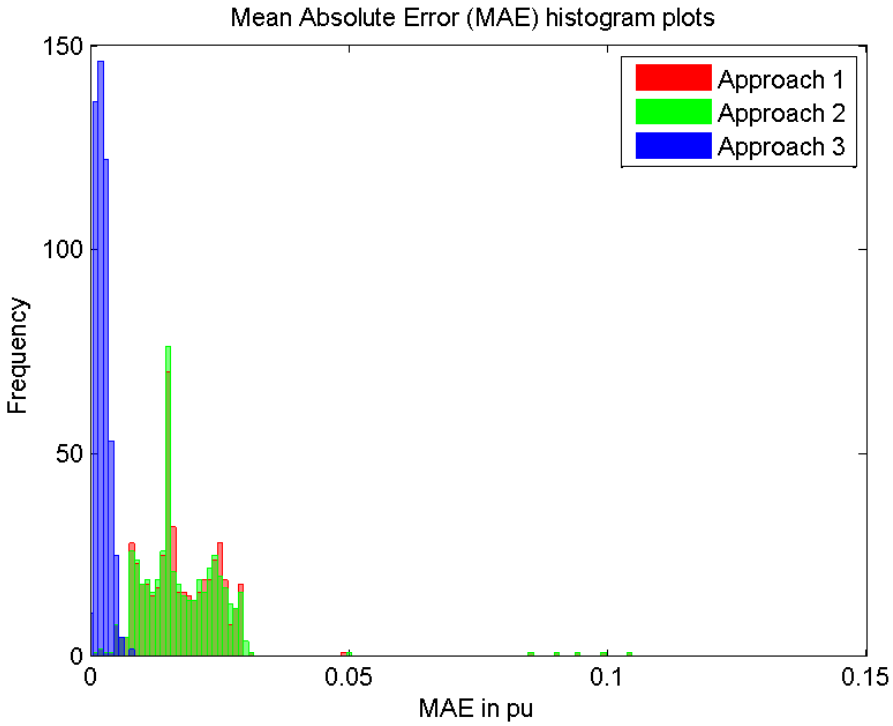


(b) Voltage real and imaginary parts estimated by three approaches after stage one

Figure 4.16: Voltage estimation results of three approaches in case 1

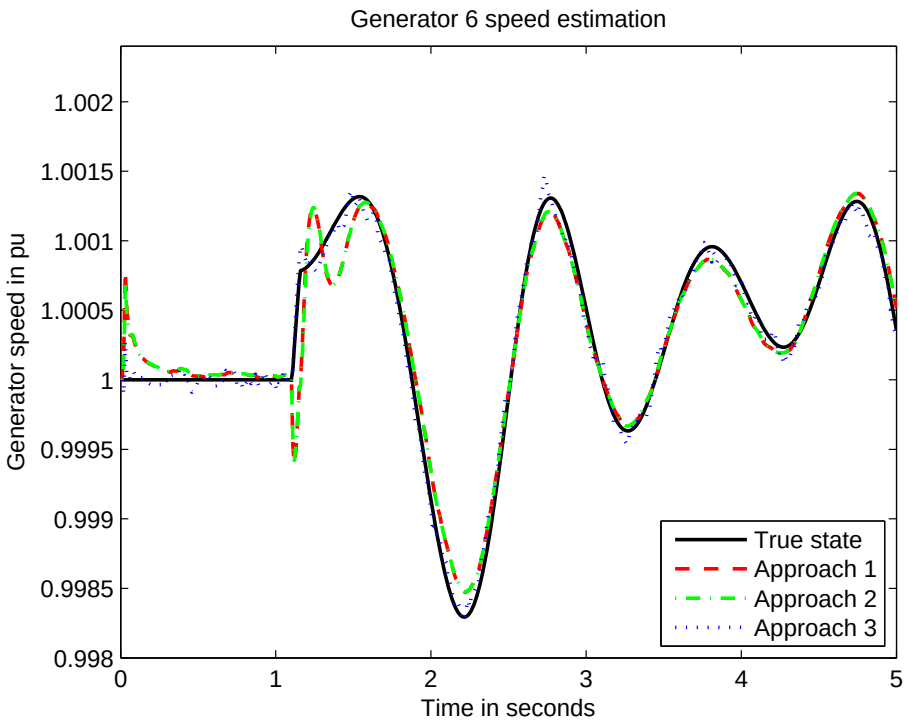


(a) MAE comparison of three approaches after stage one

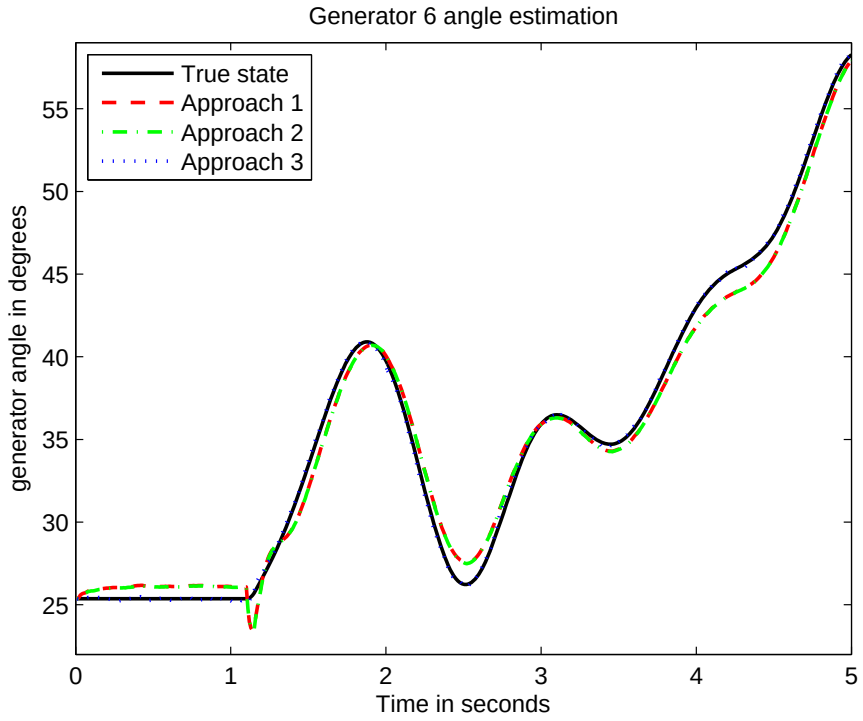


(b) MAE histograms of three approaches after stage one

Figure 4.17: Voltage estimation accuracy of three approaches in case 1

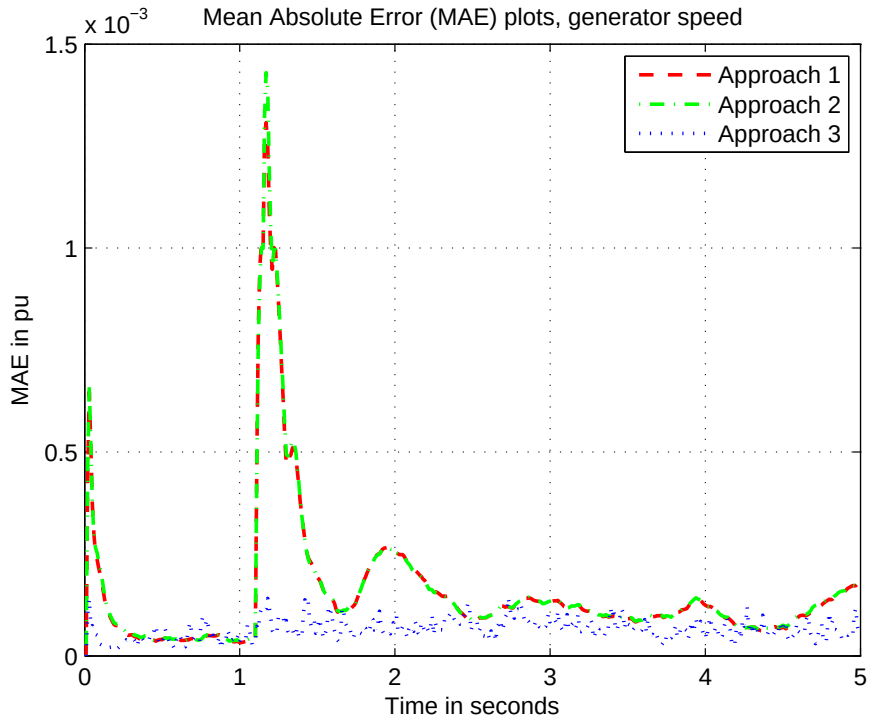


(a) Generator 6 rotor speed estimated by three approaches after stage two

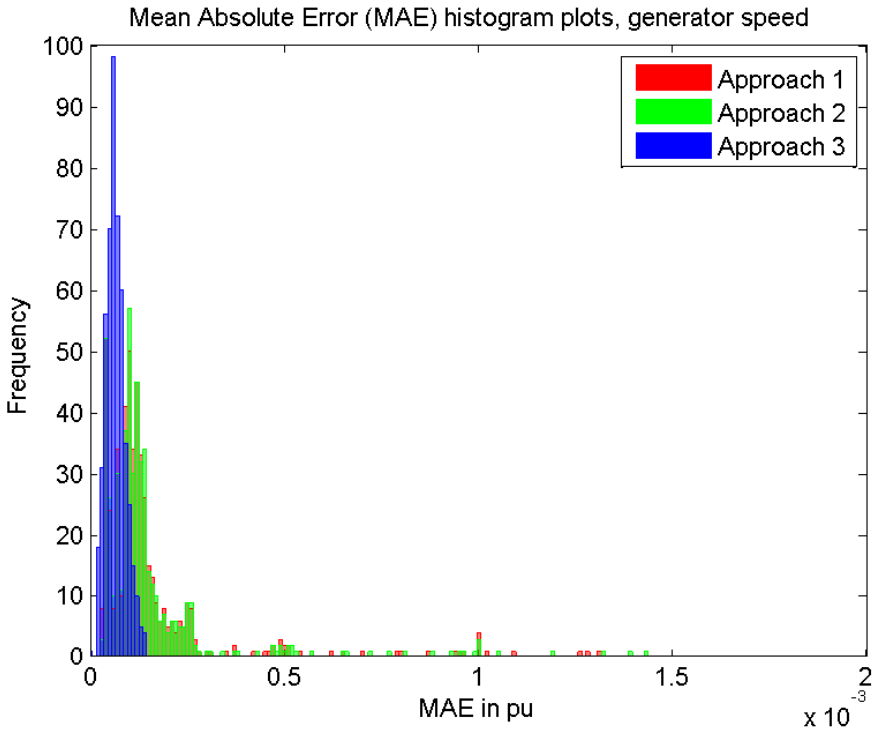


(b) Generator 6 rotor angle estimated by three approaches after stage two

Figure 4.18: Rotor speed and angle estimation results of three approaches in case 1

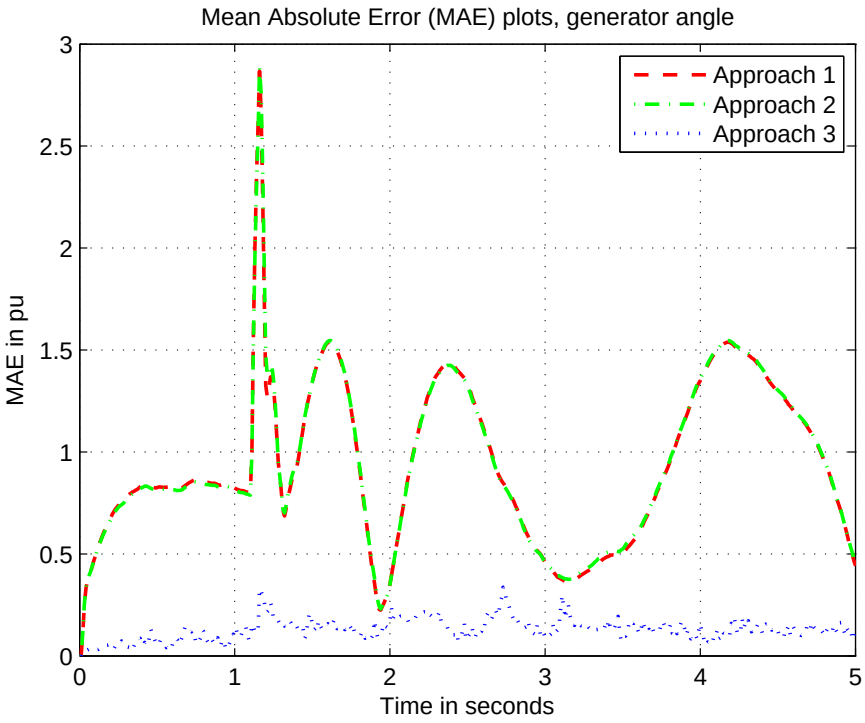


(a) Rotor speed MAE comparison of three approaches after stage two

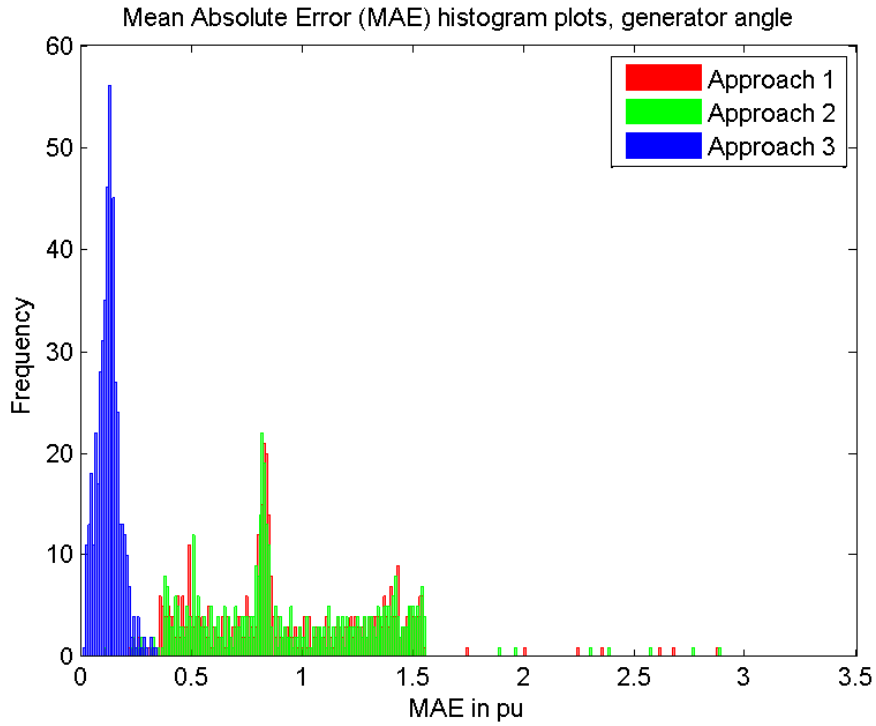


(b) Rotor speed MAE histograms of three approaches after stage two

Figure 4.19: Rotor speed estimation accuracy of three approaches in case 1



(a) Rotor angle MAE comparison of the three approaches after stage two



(b) Rotor angle MAE histograms of the three approaches after stage two

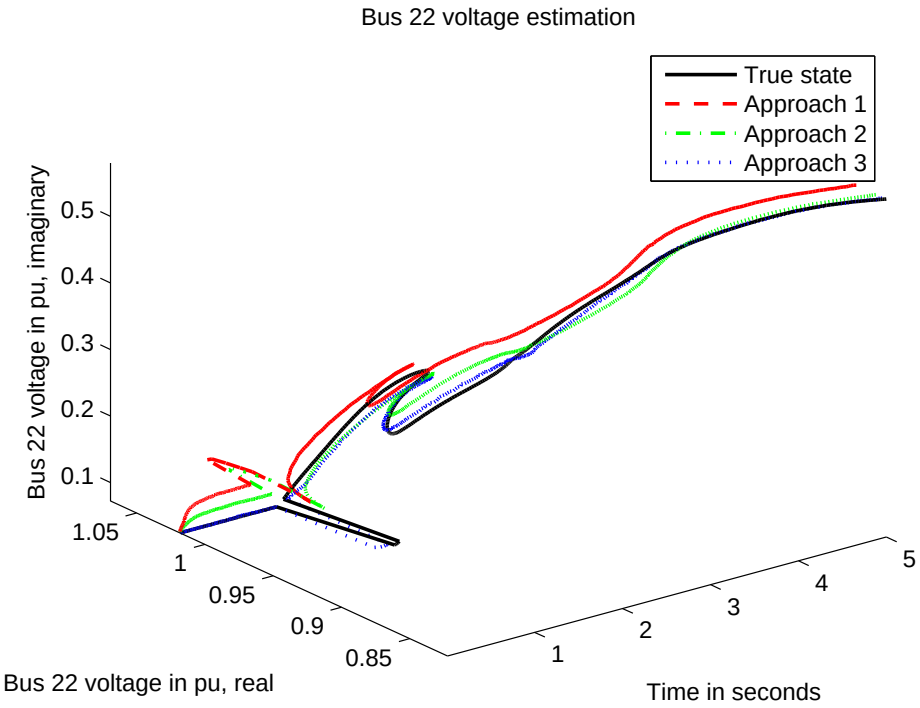
Figure 4.20: Rotor angle estimation results of three approaches in case 1

Case 2: under-valued initial Q and a malfunctioning device

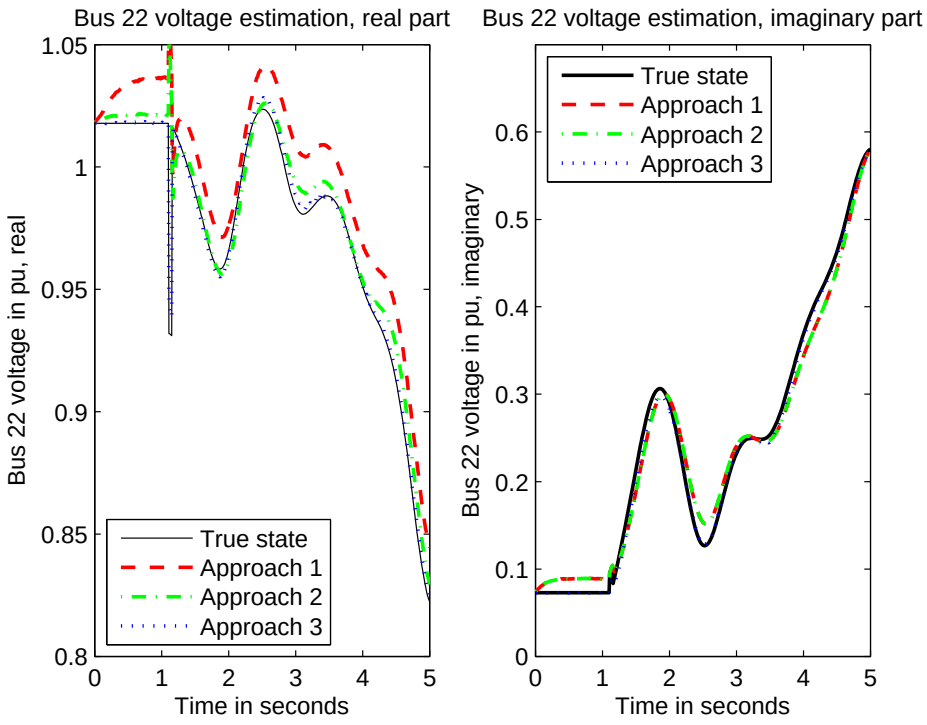
While Q is initialized with the same values as above, R is set correctly this time, i.e. PMU measurements are simulated with 1% noise. We introduce a new problem: a malfunctioning PMU at bus 22. The simulated voltage measurements provided by this PMU contain systematic errors, with a normal distribution $\mathcal{N}(1, 0.1^2)$. Fig. 4.21 illustrates the results after stage one.

Fig. 4.21 shows that at stage one the naive RKF has identified and dropped the bad measurements. However it is still affected by the under-valued initial Q. The AKF with InNoVa not only adjusted Q (bus 29 has process variance grown to 0.1165, which is still above the crowd), but also identified the bad measurement. Although our algorithm does not exclude the measurement—we seek to avoid unobservable conditions by maintaining sufficient redundancy—the noise of this erroneous measurement becomes 41.21%, which is significantly higher than the others (1%). Hence this measurement is not weighted as heavily, and has a negligible impact on the estimated states. Fig. 4.22 also confirms that at stage one, while Approach 2 does improve, Approach 3 stays the closest to the truth.

Consequently, the accuracy levels of these three filters have directly affected the estimation performances at the next stage. The tracking quality remains to be Approach $3 >$ Approach $2 >$ Approach 1, as seen in Fig. 4.23, Fig. 4.24 and Fig. 4.25, which show the dynamic states tracking results, *MAE* analysis for generator speed estimates and *MAE* analysis for generator angle estimates respectively.

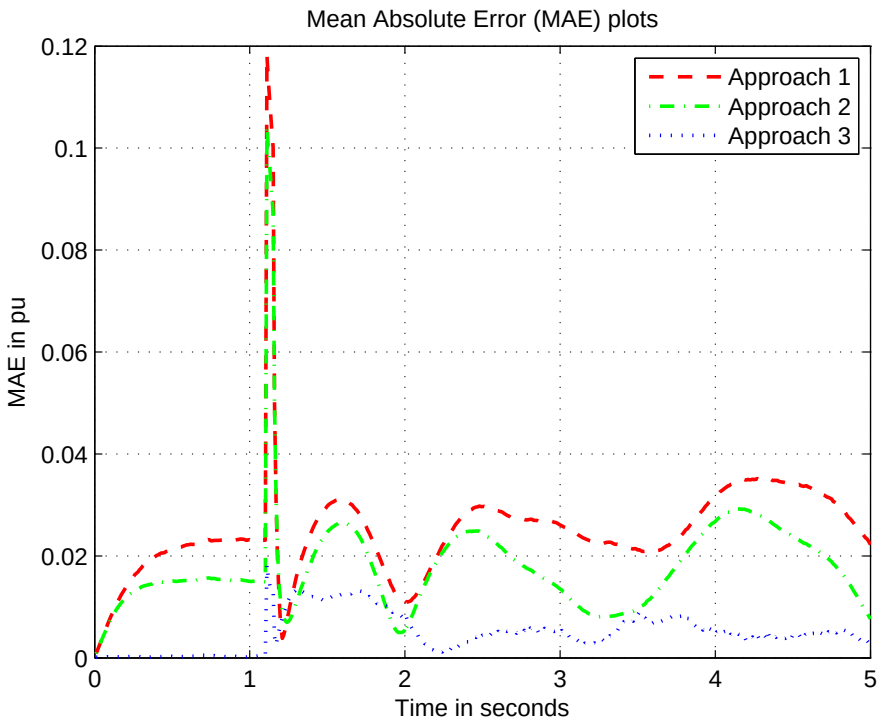


(a) Bus 22 voltage estimated by three approaches after stage one

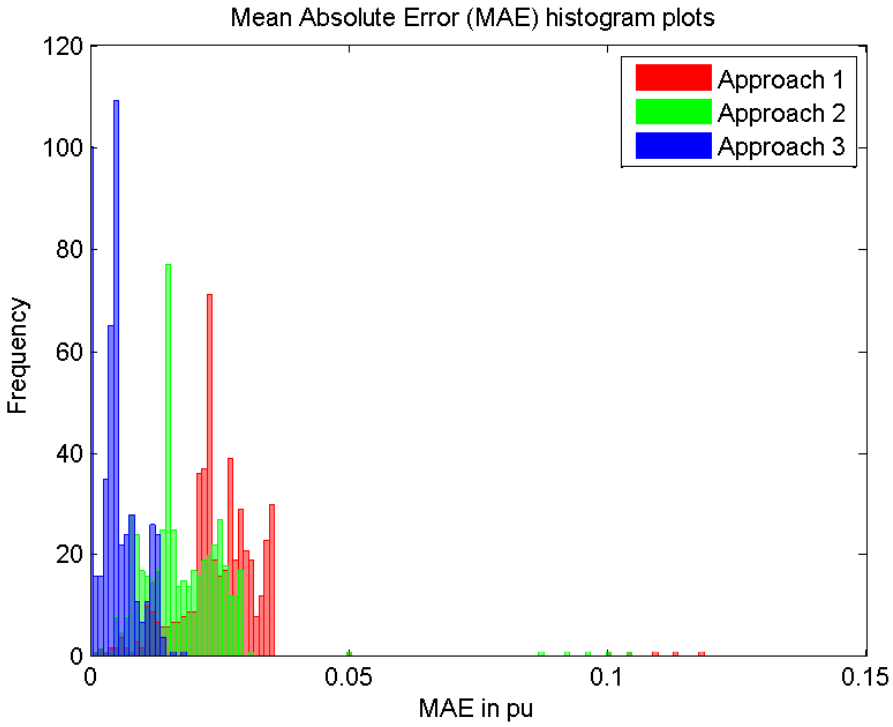


(b) Voltage real and imaginary parts estimated by three approaches after stage one

Figure 4.21: Voltage estimation results of three approaches in case 2

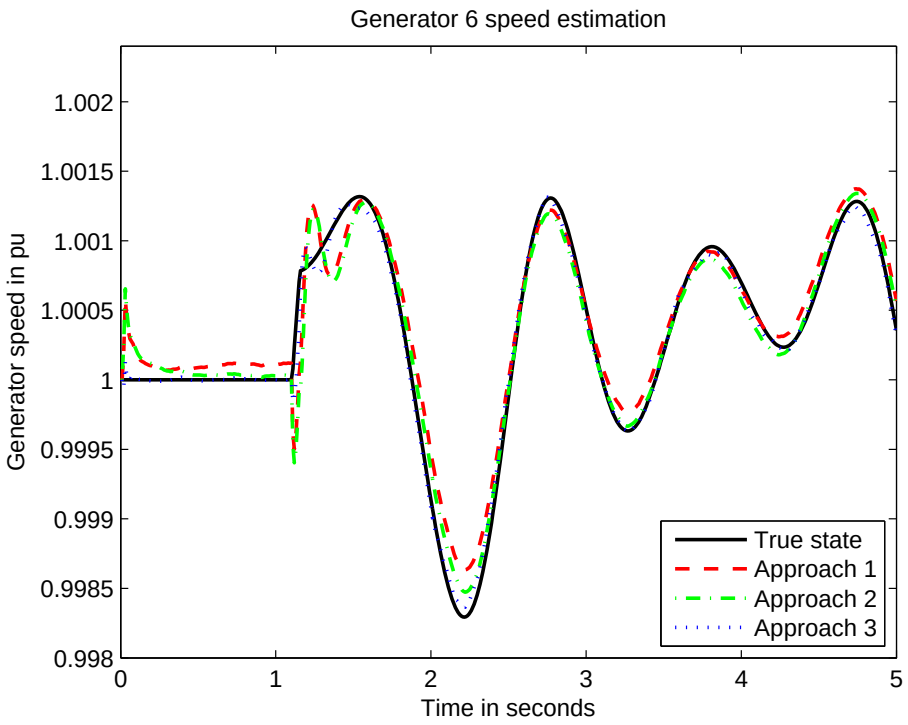


(a) MAE comparison of three approaches after stage one

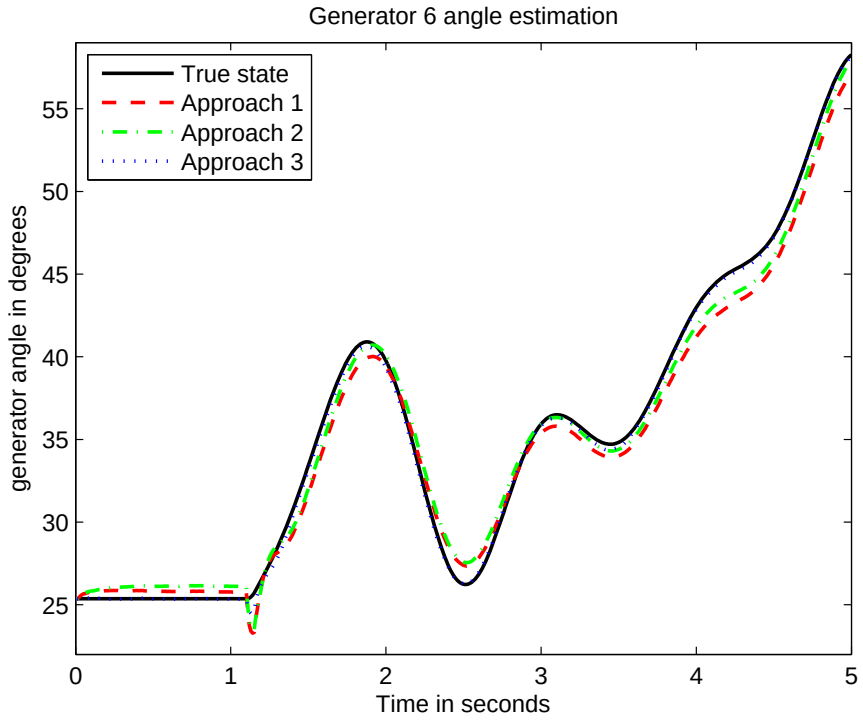


(b) MAE histograms of three approaches after stage one

Figure 4.22: Voltage estimation accuracy of three approaches in case 2

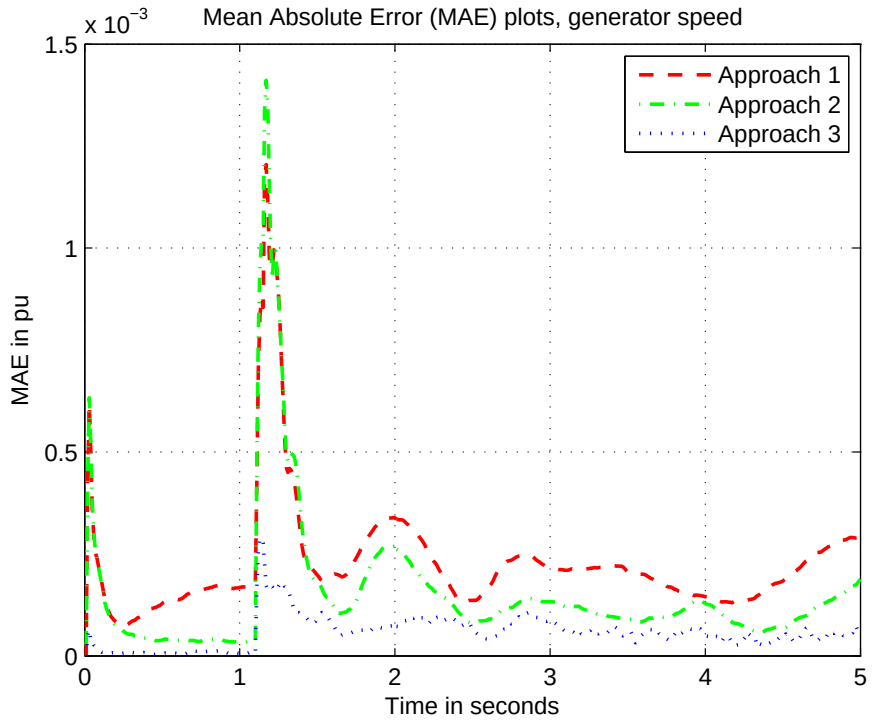


(a) Generator 6 rotor speed estimated by three approaches after stage two

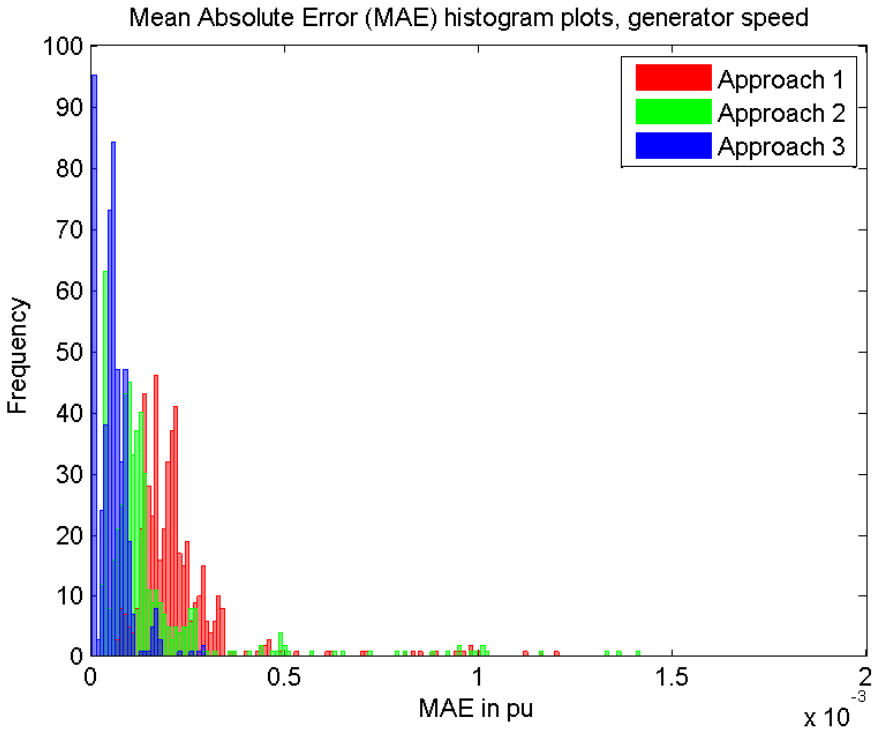


(b) Generator 6 rotor angle estimated by three approaches after stage two

Figure 4.23: Rotor speed and angle estimation results of three approaches in case 2

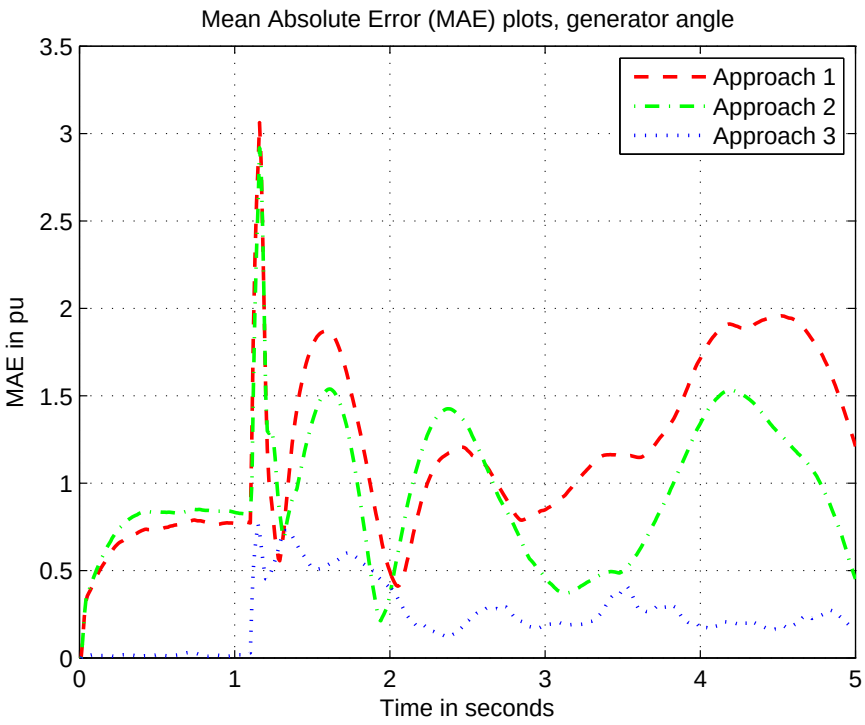


(a) Rotor speed MAE comparison of three approaches after stage two

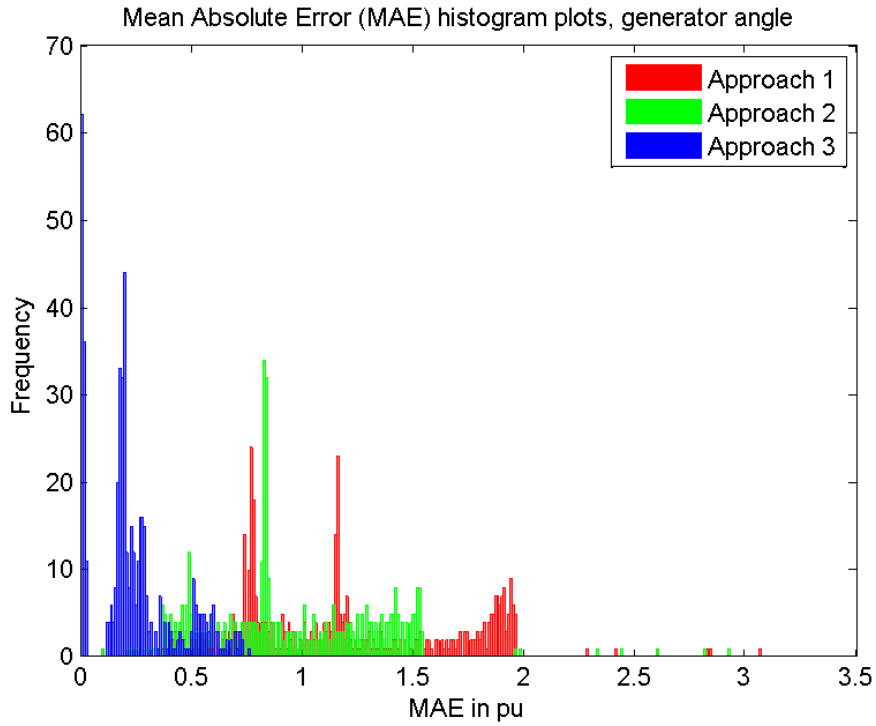


(b) Rotor speed MAE histograms of three approaches after stage two

Figure 4.24: Rotor speed estimation accuracy of three approaches in case 2



(a) Rotor angle MAE comparison of the three approaches after stage two



(b) Rotor angle MAE histograms of the three approaches after stage two

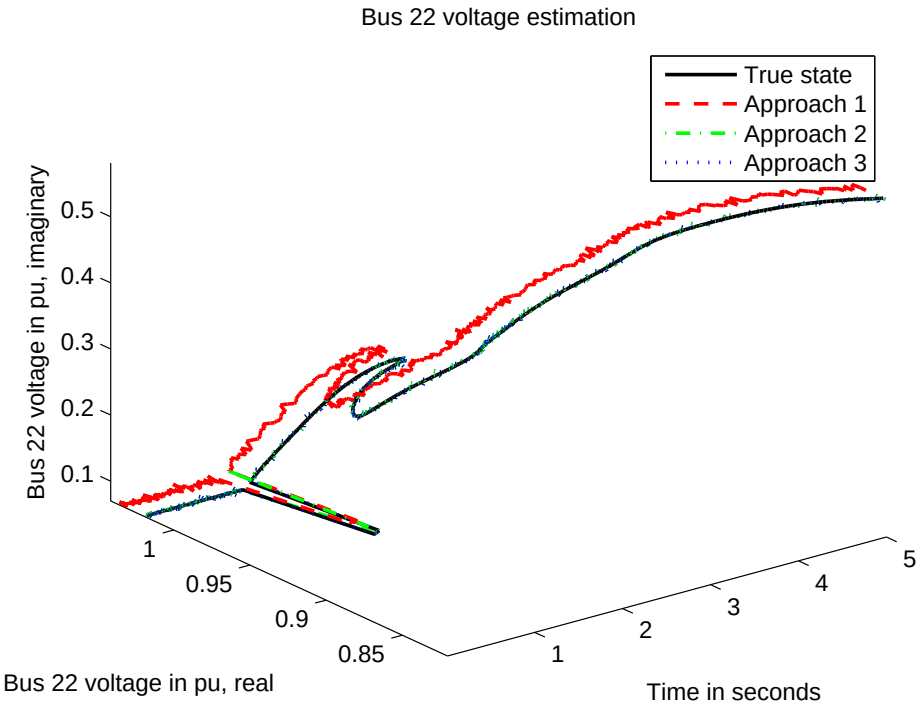
Figure 4.25: Rotor angle estimation results of three approaches in case 2

Case 3: fairly-valued initial Q and a malfunctioning device

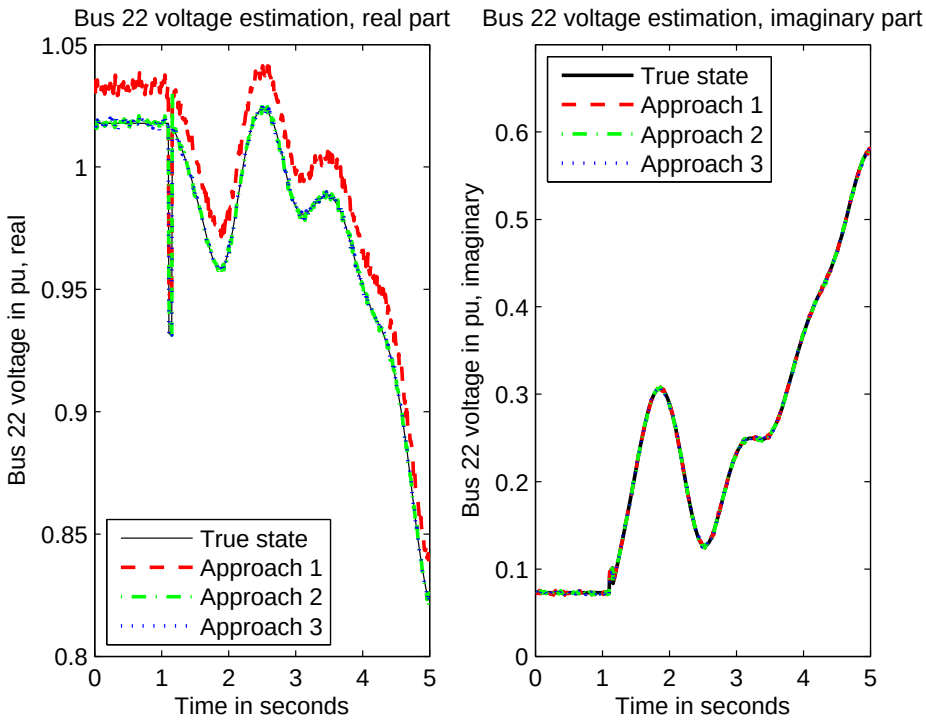
This time the PMU at bus 22 has the same malfunctioning behaviour as described in case 2, but the initial process noise variance of each state variable is 0.05, which is more appropriate for our system model.

At stage one, the AKF with InNoVa has its process noise variance at bus 29 grown to 0.1166 in 5 seconds; it also identifies the bad measurements by inflating the corresponding measurement noise level to 41.17%. As compared to others at 1%, it is quite obvious that we are having a malfunctioning device at bus 22. Fig. 4.26 tells us that with this fairly-valued initial Q, while the AKF with InNoVa still performs the best, the naive RKF is able to correctly detect, identify and eliminate the bad data most of the time. The *MAE* analysis in Fig.4.27 has shown that the naive RKF can behave quite close to the AKF with InNoVa in this case, while the traditional KF produces a shifted tracking result because of the mulfunctioned PMU data.

At stage two, Fig. 4.28 also proves that this Q initialization has greatly helped Approach 2 in its dynamic state estimation. Fig. 4.29 and Fig. 4.30 have provided detailed information regarding dynamic state estimation accuracy of these three approaches.

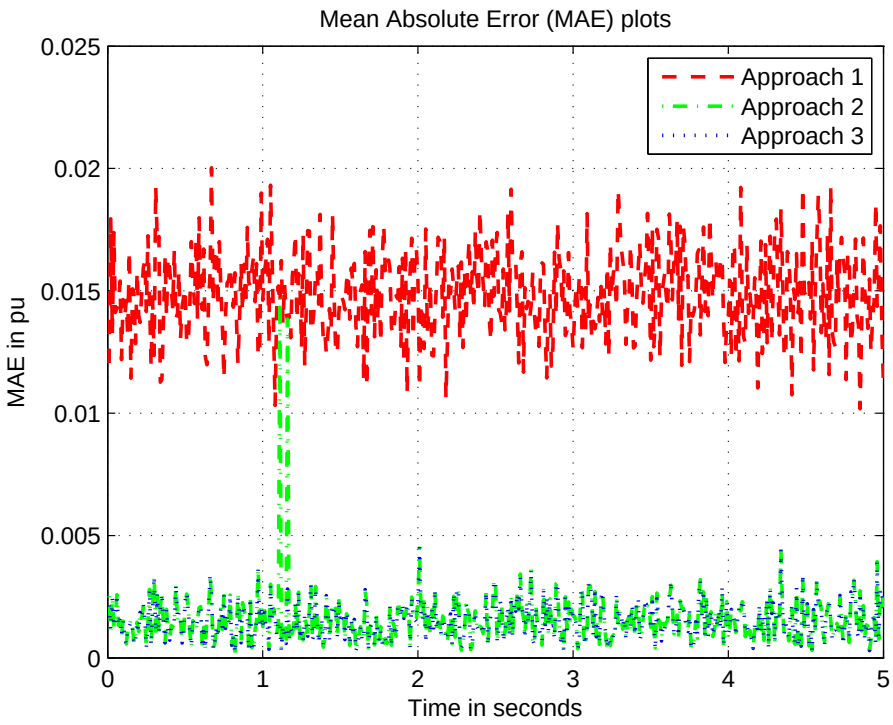


(a) Bus 22 voltage estimated by three approaches after stage one

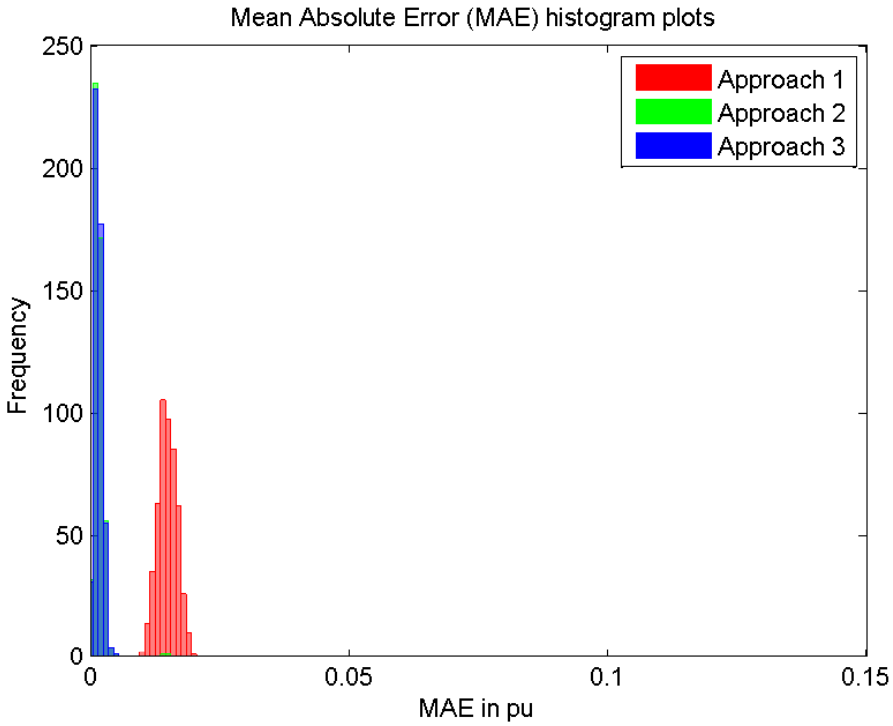


(b) Voltage real and imaginary parts estimated by three approaches after stage one

Figure 4.26: Voltage estimation results of three approaches in case 3

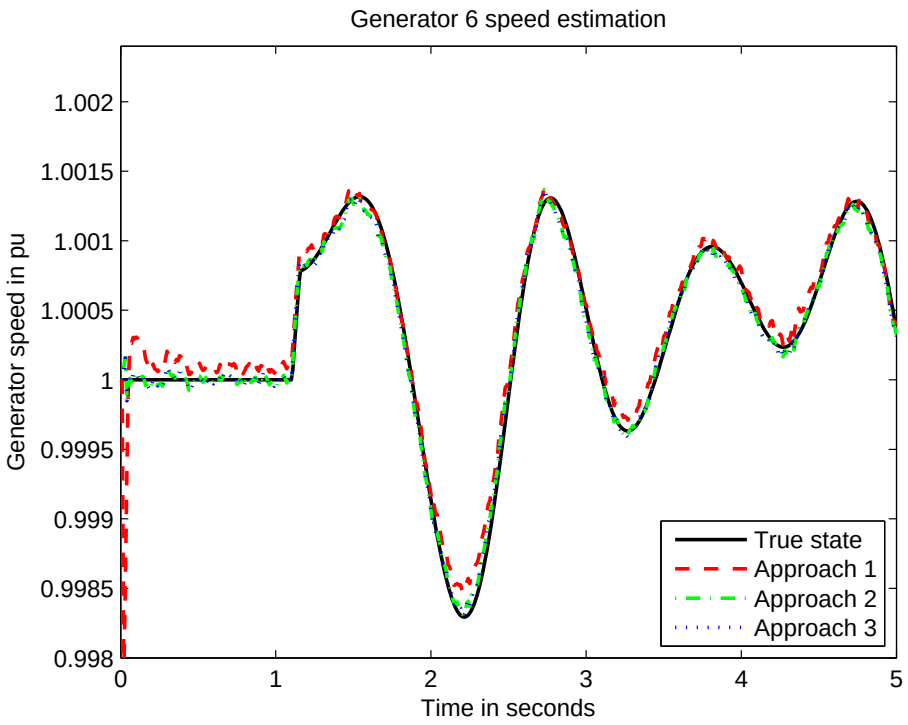


(a) MAE comparison of three approaches after stage one

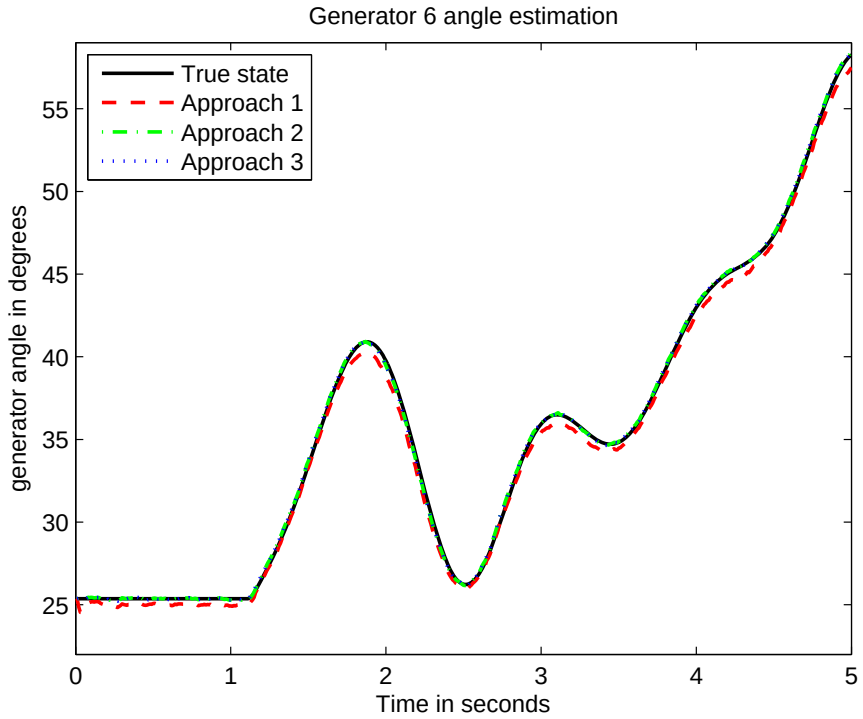


(b) MAE histograms of three approaches after stage one

Figure 4.27: Voltage estimation accuracy of three approaches in case 3

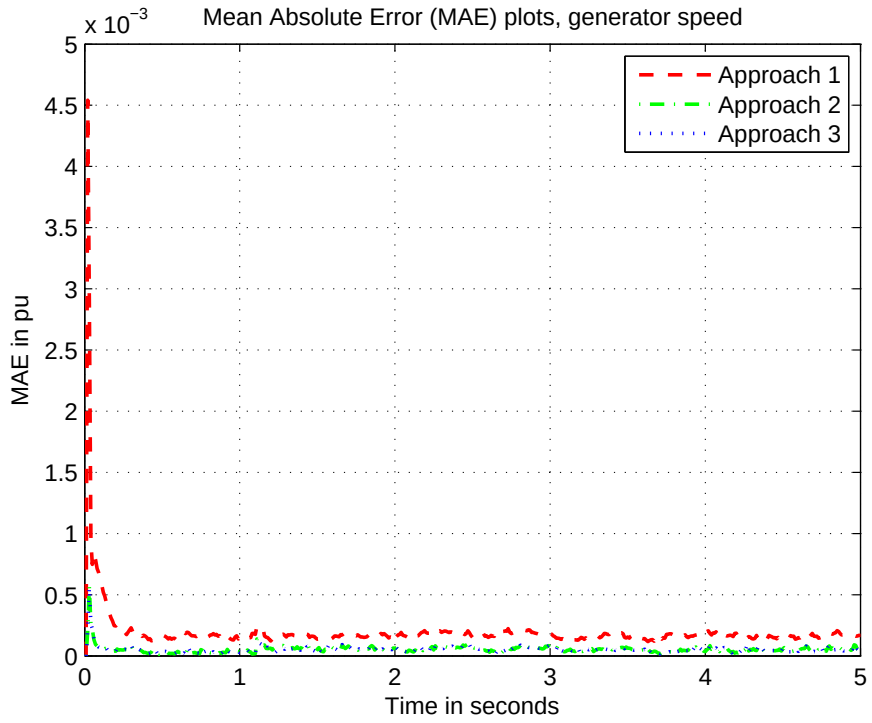


(a) Generator 6 rotor speed estimated by three approaches after stage two

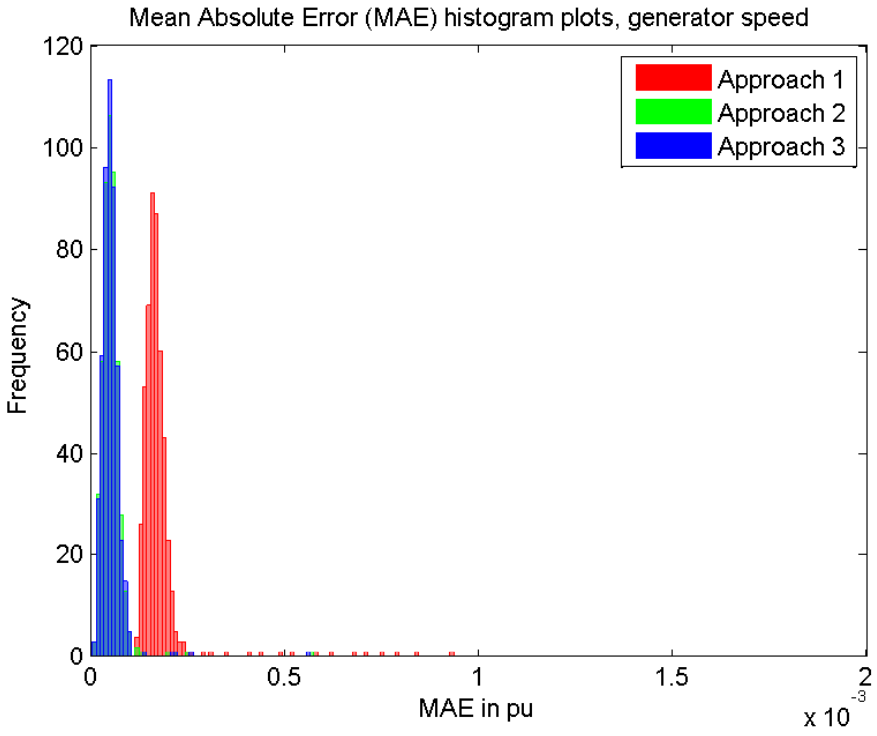


(b) Generator 6 rotor angle estimated by three approaches after stage two

Figure 4.28: Rotor speed and angle estimation results of three approaches in case 3

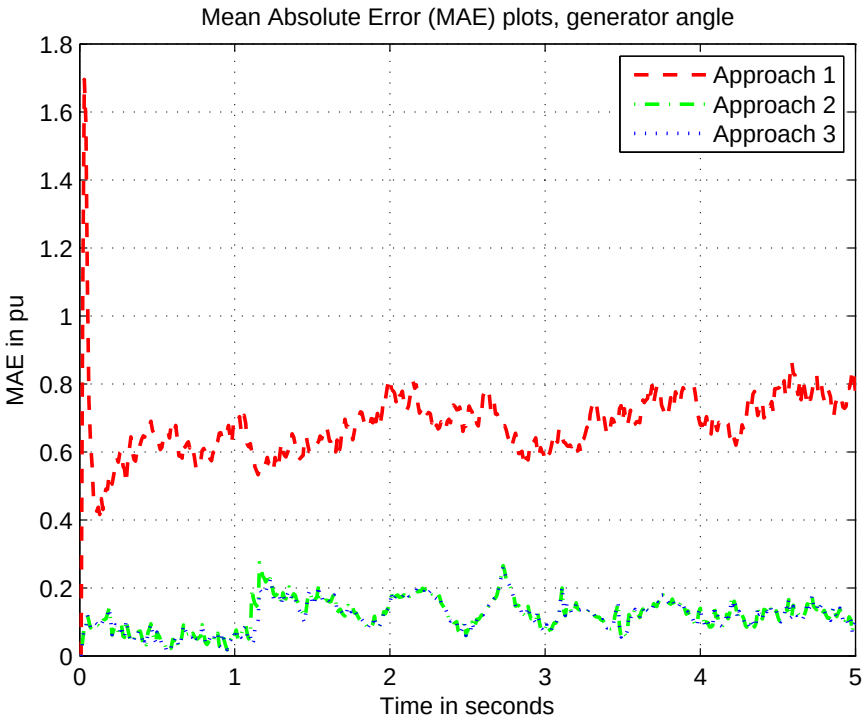


(a) Rotor speed MAE comparison of three approaches after stage two

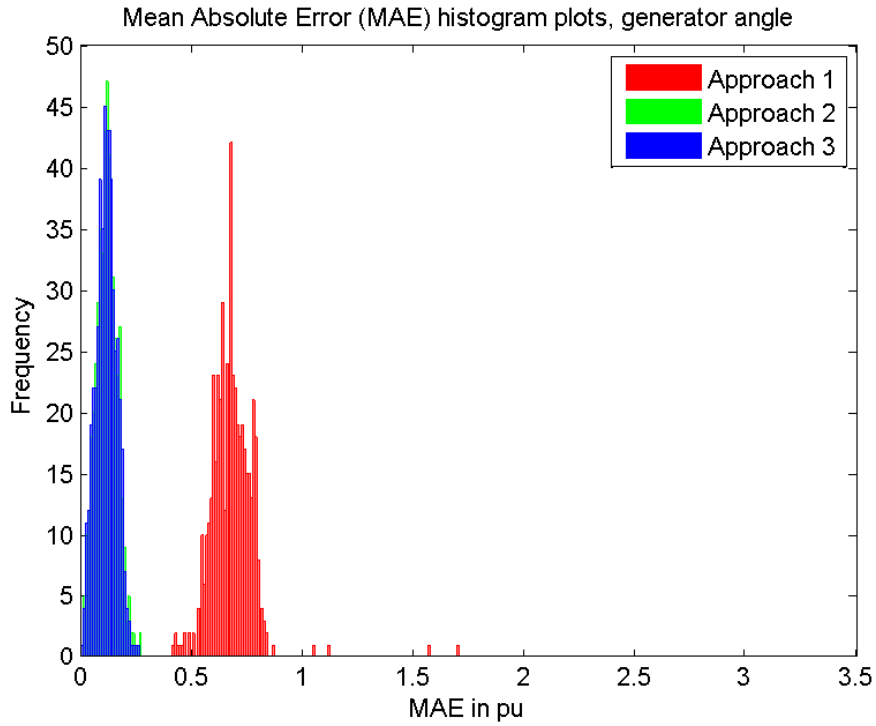


(b) Rotor speed MAE histograms of three approaches after stage two

Figure 4.29: Rotor speed estimation accuracy of three approaches in case 3



(a) Rotor angle MAE comparison of the three approaches after stage two



(b) Rotor angle MAE histograms of the three approaches after stage two

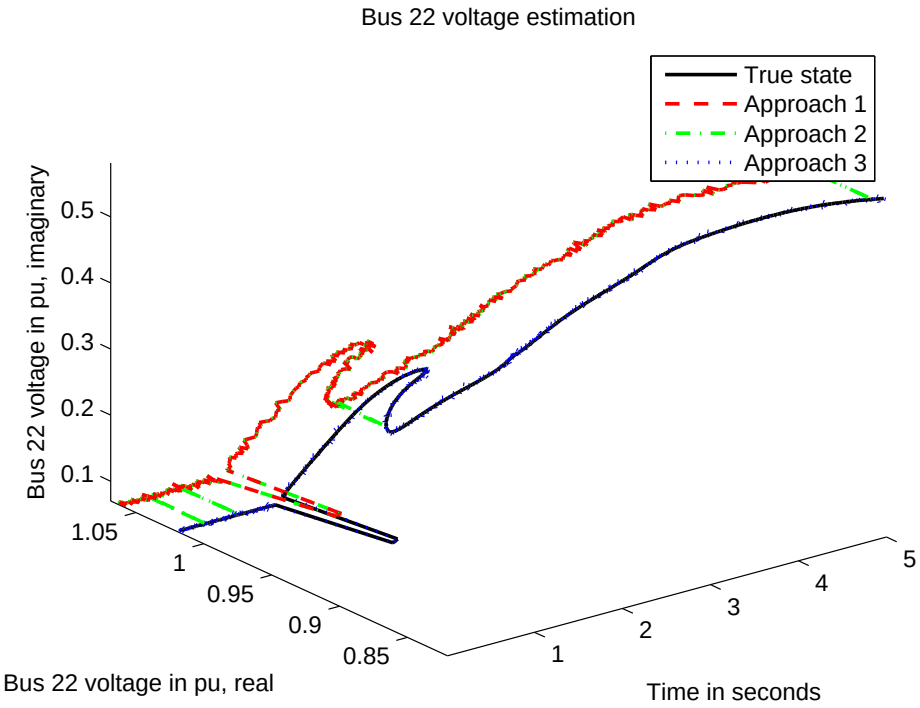
Figure 4.30: Rotor angle estimation results of three approaches in case 3

Case 4: over-valued initial Q and a malfunctioning device

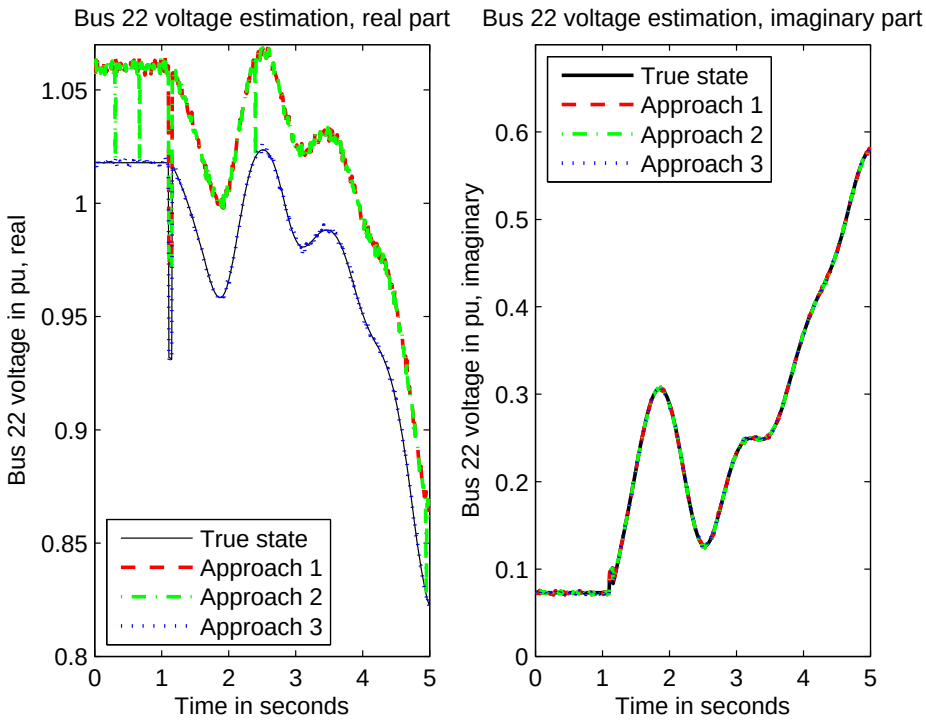
Now the initial process noise variances in Q are each assigned a much larger value 1, meaning that the user is not confident with the process model. This time the voltage measurement errors at bus 22 are even larger, with distribution $\mathcal{N}(2.8, 0.1^2)$.

The results obtained from stage one and stage two are depicted in Fig. 4.31 and Fig. 4.33 respectively. Interestingly in Fig. 4.31(a), the naive RKF rarely detects the systematic error. The reason is that Q is set so large (hence S is so large), the *normalized innovations* almost always fall under the threshold. However the error is always detected by the *normalized residual test* in the AKF with InNoVa. As a matter of fact, comparing to other measurement noise at 1%, the noise of this erroneous measurement has been inflated to 100.95%.

In *MAE* analyses plots Fig. 4.32 (for bus voltage estimates), Fig. 4.34 (for generator speed estimates) and Fig. 4.35 (for generator angle estimates), we can see that in either stage one or stage two, Approach 3 always performs the best, while Approach 2 performs as bad as Approach 1 almost all the time.

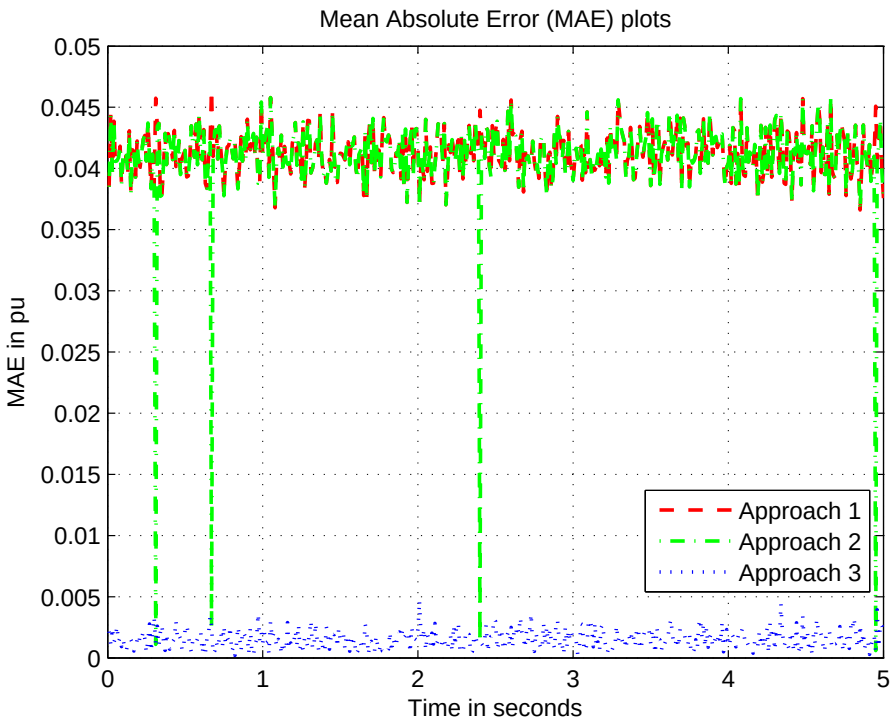


(a) Bus 22 voltage estimated by three approaches after stage one

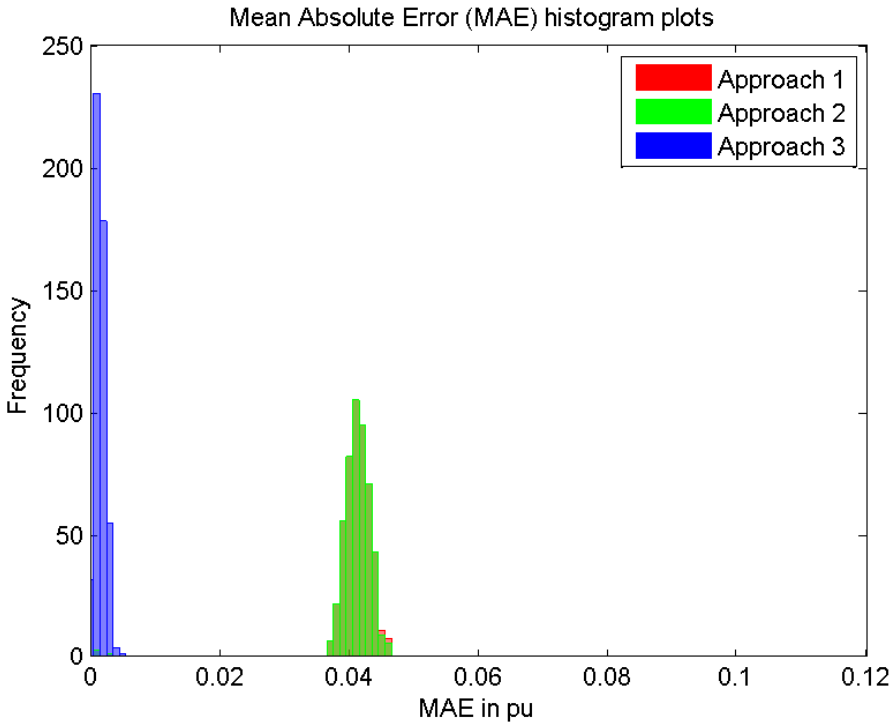


(b) Voltage real and imaginary parts estimated by three approaches after stage one

Figure 4.31: Voltage estimation results of three approaches in case 4

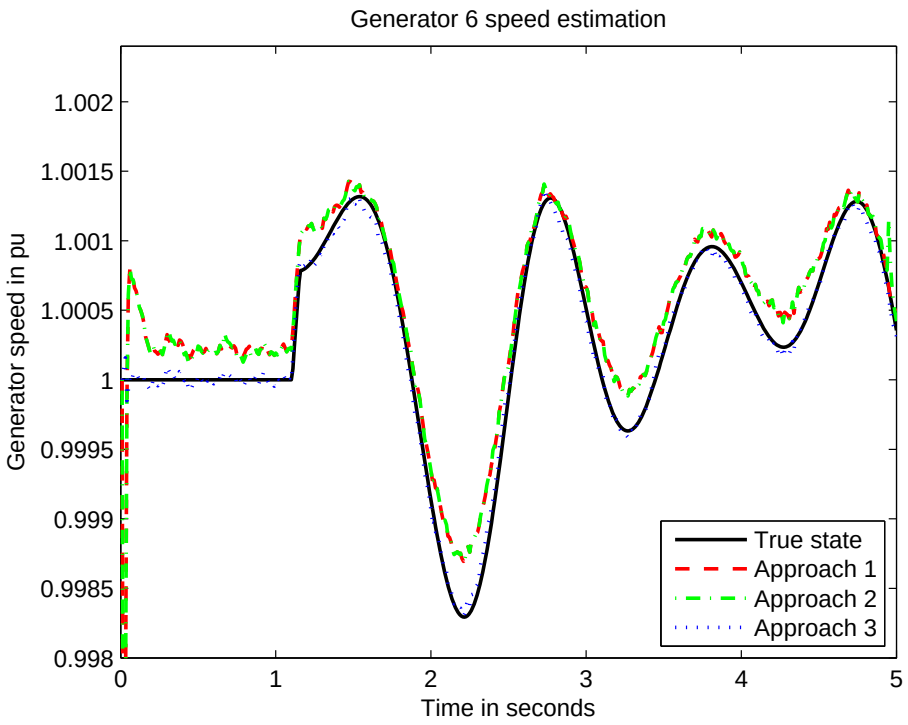


(a) MAE comparison of three approaches after stage one

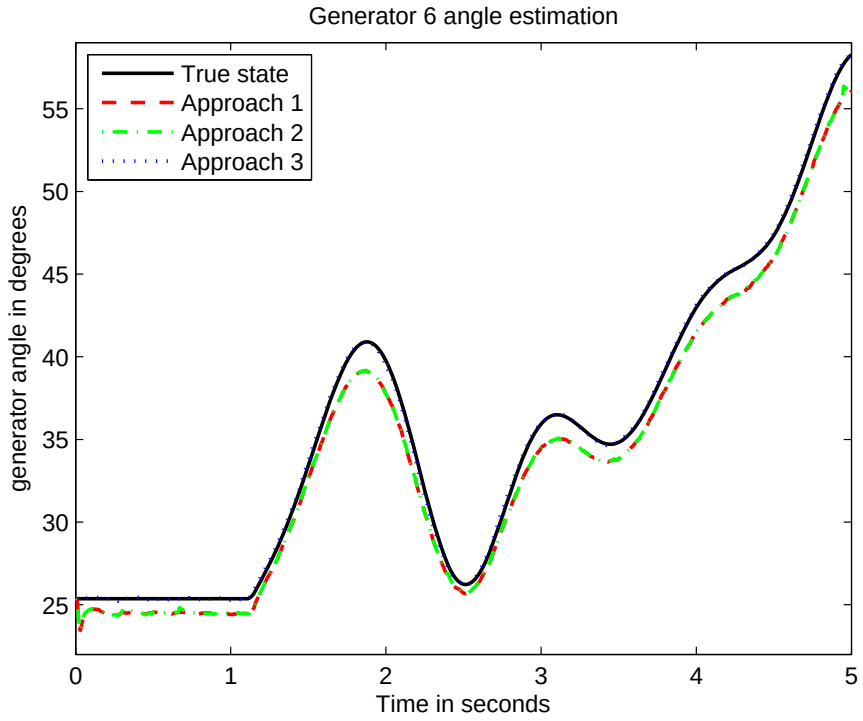


(b) MAE histograms of three approaches after stage one

Figure 4.32: Voltage estimation accuracy of three approaches in case 4

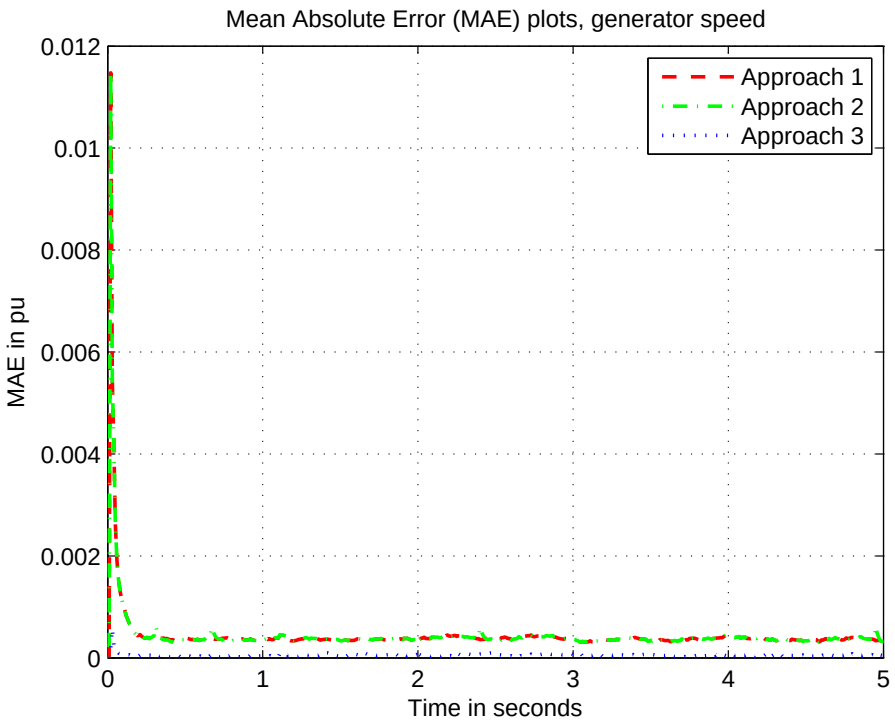


(a) Generator 6 rotor speed estimated by three approaches after stage two

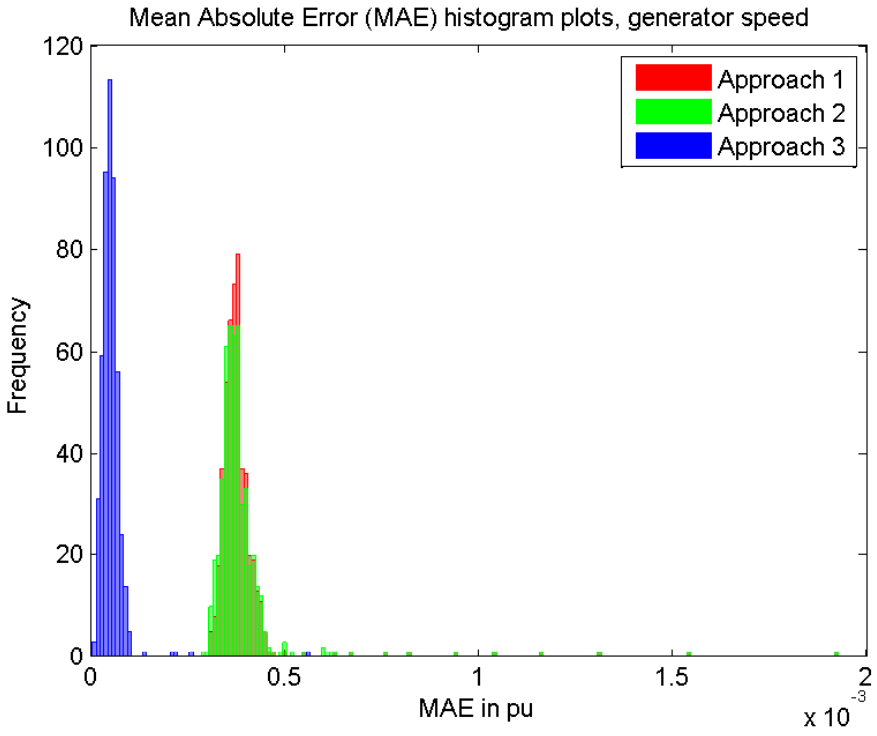


(b) Generator 6 rotor angle estimated by three approaches after stage two

Figure 4.33: Rotor speed and angle estimation results of three approaches in case 4

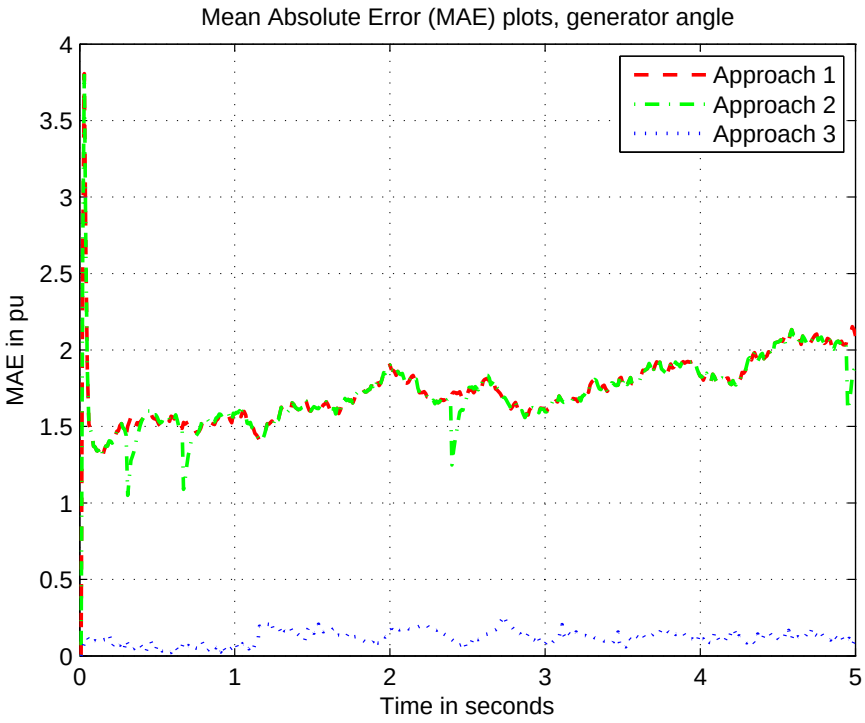


(a) Rotor speed MAE comparison of three approaches after stage two

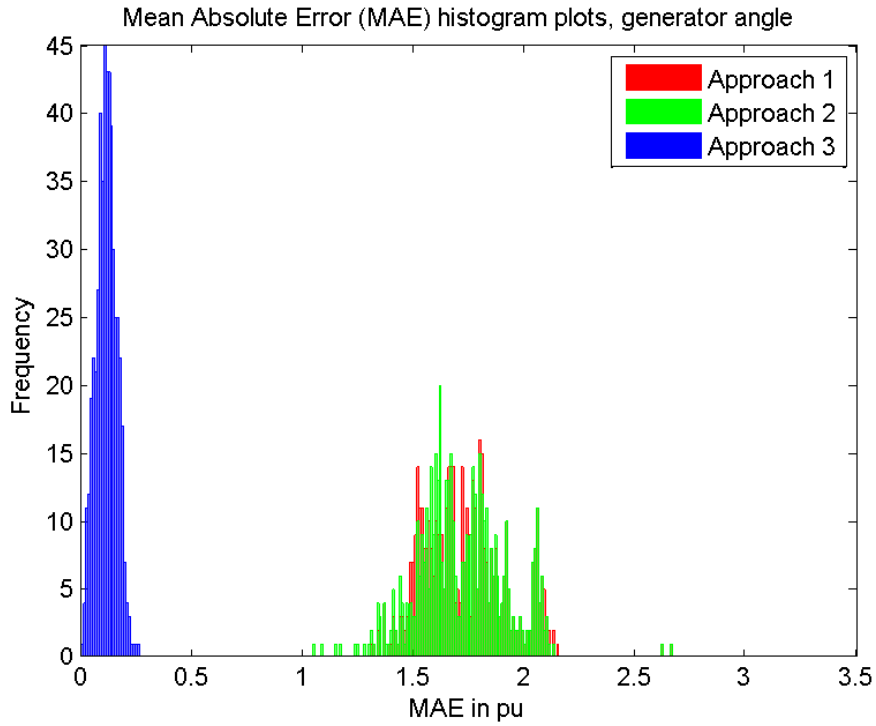


(b) Rotor speed MAE histograms of three approaches after stage two

Figure 4.34: Rotor speed estimation accuracy of three approaches in case 4



(a) Rotor angle MAE comparison of the three approaches after stage two



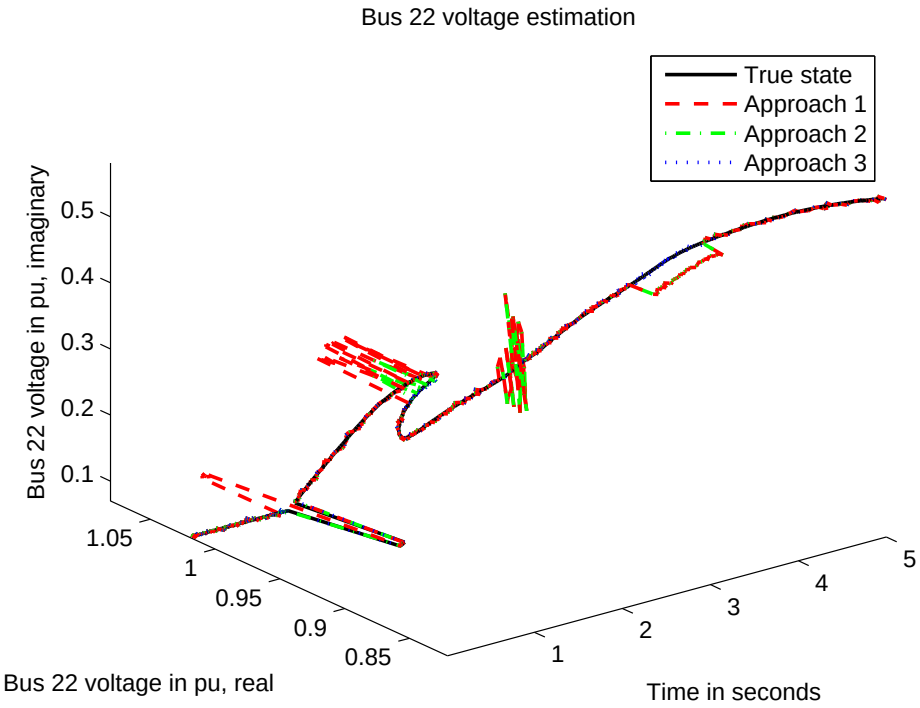
(b) Rotor angle MAE histograms of the three approaches after stage two

Figure 4.35: Rotor angle estimation results of three approaches in case 4

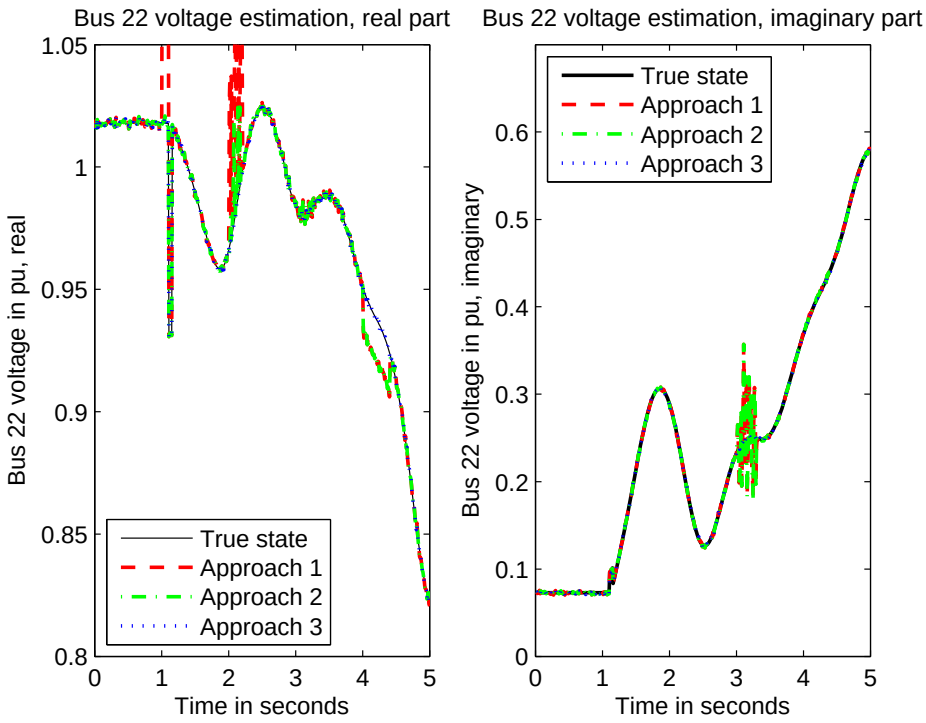
Case 5: over-valued initial Q and complex noise interference

Although the process noise covariance Q is initialized with the same large values as in case 4, case 5 has more complicated noise interference (note that although errors are treated as Gaussian white noise, they do NOT have to be Gaussian and white): at $t = 1$, bus 22 voltage real-part measurements are corrupted by a constant bias value of 4 for 0.1 second; at $t = 2$, bus 23 voltage measurements are corrupted by the normally distributed noise $\mathcal{N}(0, 5^2)$ for 0.2 second; at $t = 3$, line $22 - 21$ current measurements are corrupted by the uniformly distributed noise $\mathcal{U}(0, 10)$ for 0.3 second; at $t = 4$, line $23 - 22$ current imaginary-part measurements are corrupted by a constant bias value of -15 for 0.4 second. Fig. 4.36 and Fig. 4.38 plot the tracking results from stage one and stage two respectively.

It is clear that Approach 1 is the most vulnerable to these disturbances. When the disturbance is large enough (e.g., $t = 1$), the naive RKF in Approach 2 is capable of detecting the bad measurements and calibrating the estimates in stage one, but otherwise it underperforms. Fortunately, the AKF with InNoVa in Approach 3 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. 4.37 shows that various noise disturbances have negligible effect on the AKF with InNoVa. Therefore, Approach 3 still outputs the most impressive dynamic state estimating results with the highest accuracy level, as seen in Fig. 4.39 (for generator speed estimates) and Fig. 4.35 (for generator angle estimates).

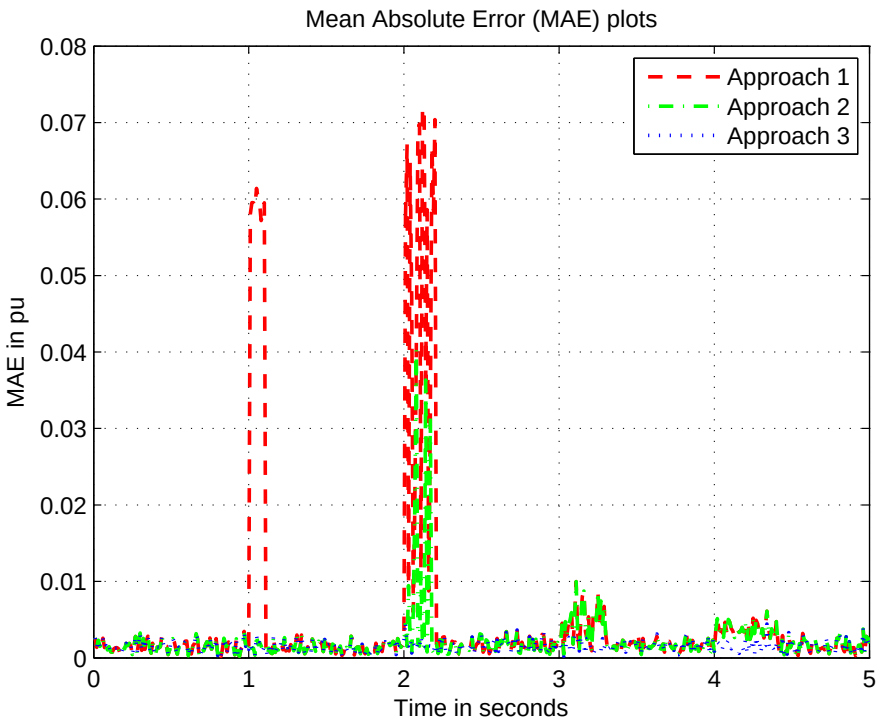


(a) Bus 22 voltage estimated by three approaches after stage one

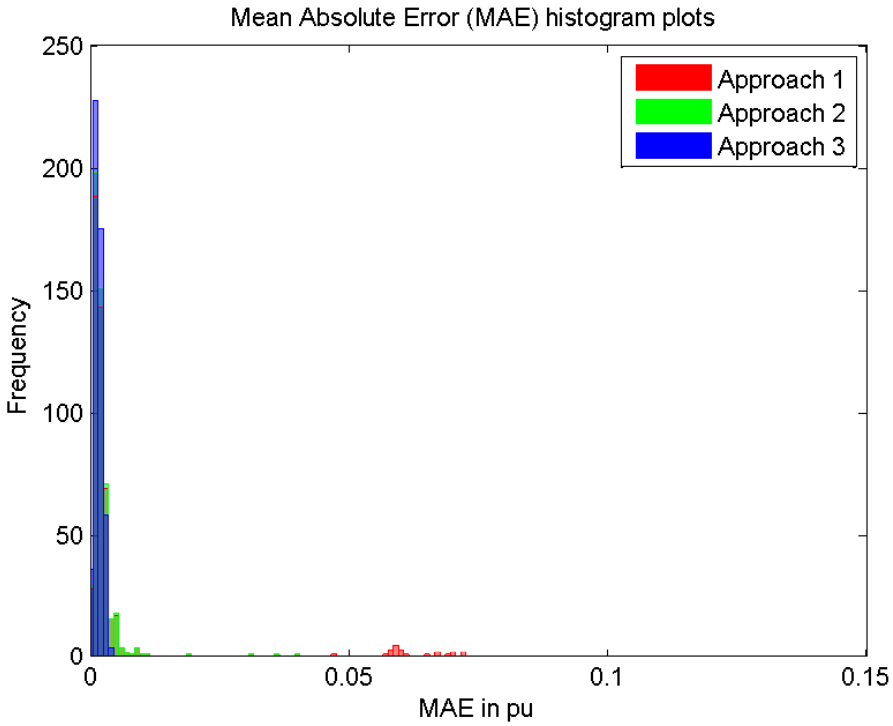


(b) Voltage real and imaginary parts estimated by three approaches after stage one

Figure 4.36: Voltage estimation results of three approaches in case 5

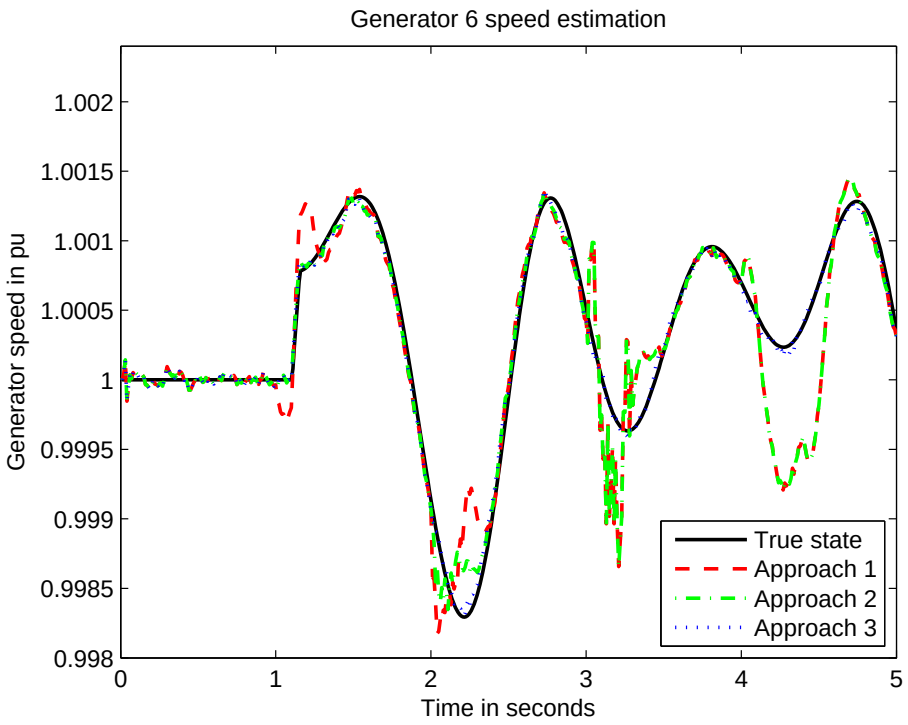


(a) MAE comparison of three approaches after stage one

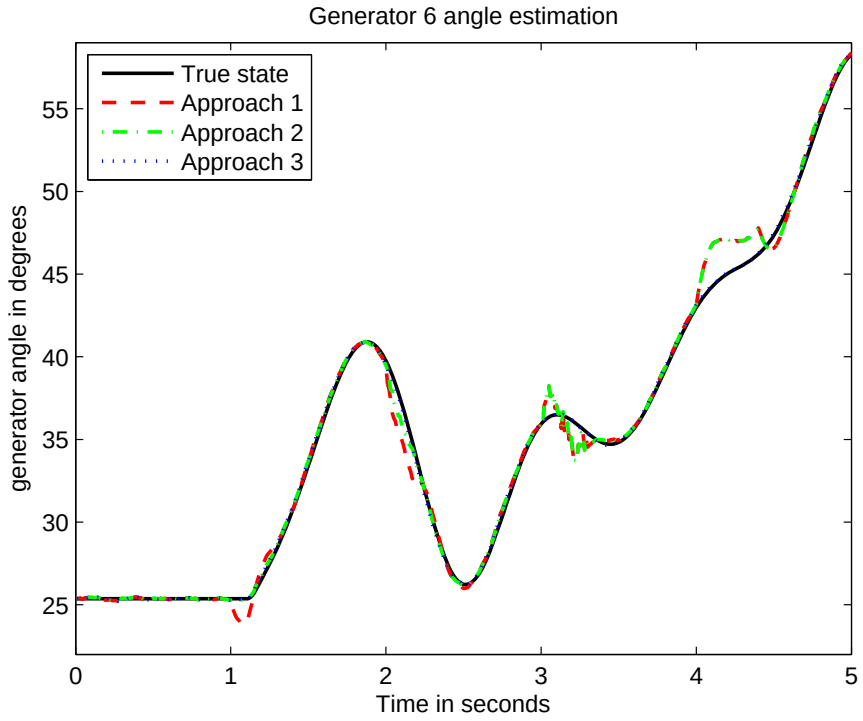


(b) MAE histograms of three approaches after stage one

Figure 4.37: Voltage estimation accuracy of three approaches in case 5

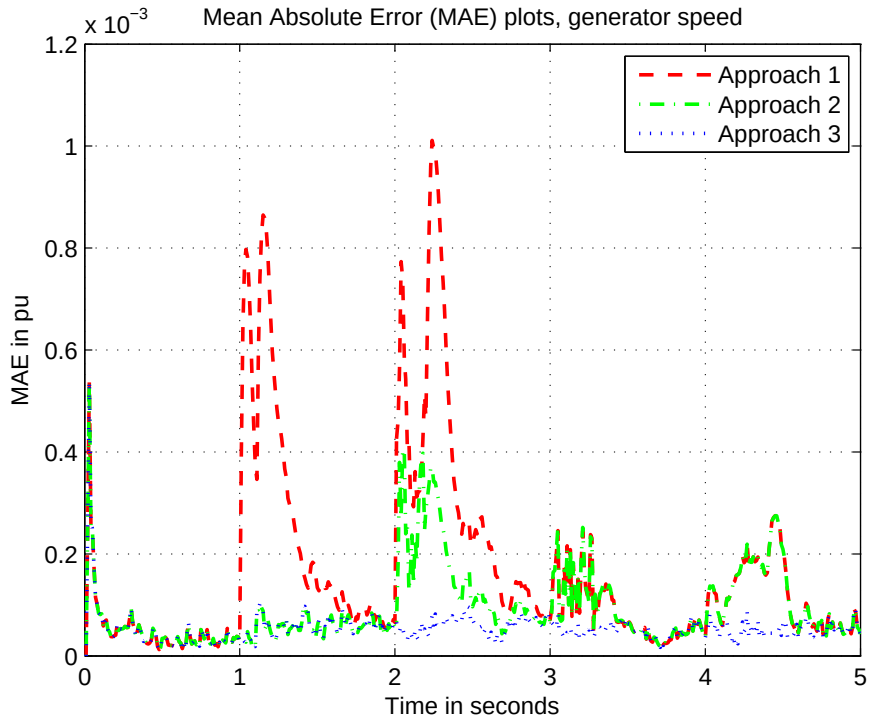


(a) Generator 6 rotor speed estimated by three approaches after stage two

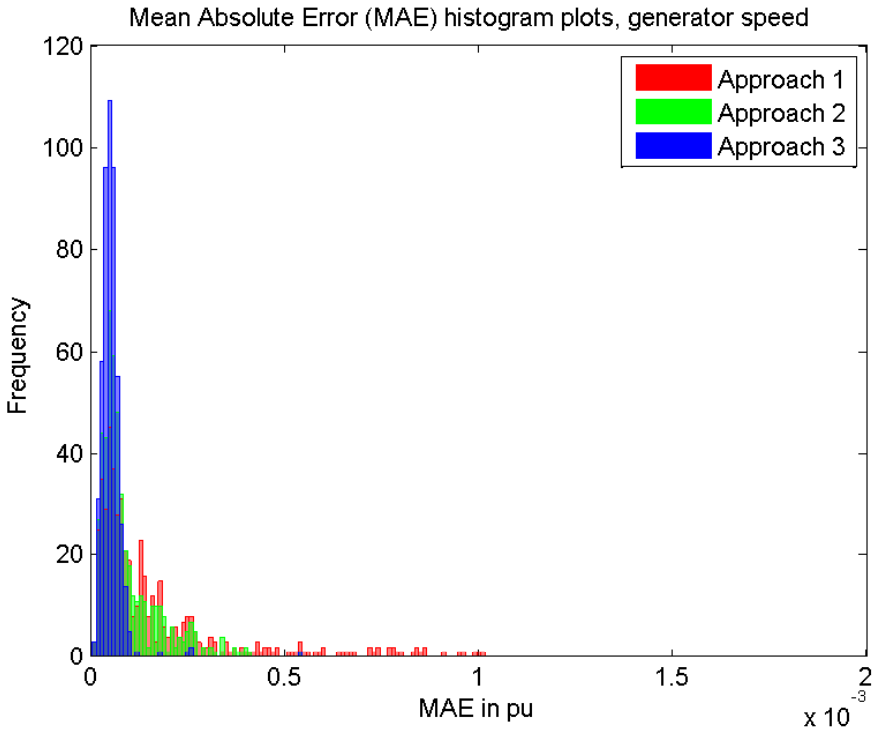


(b) Generator 6 rotor angle estimated by three approaches after stage two

Figure 4.38: Rotor speed and angle estimation results of three approaches in case 5

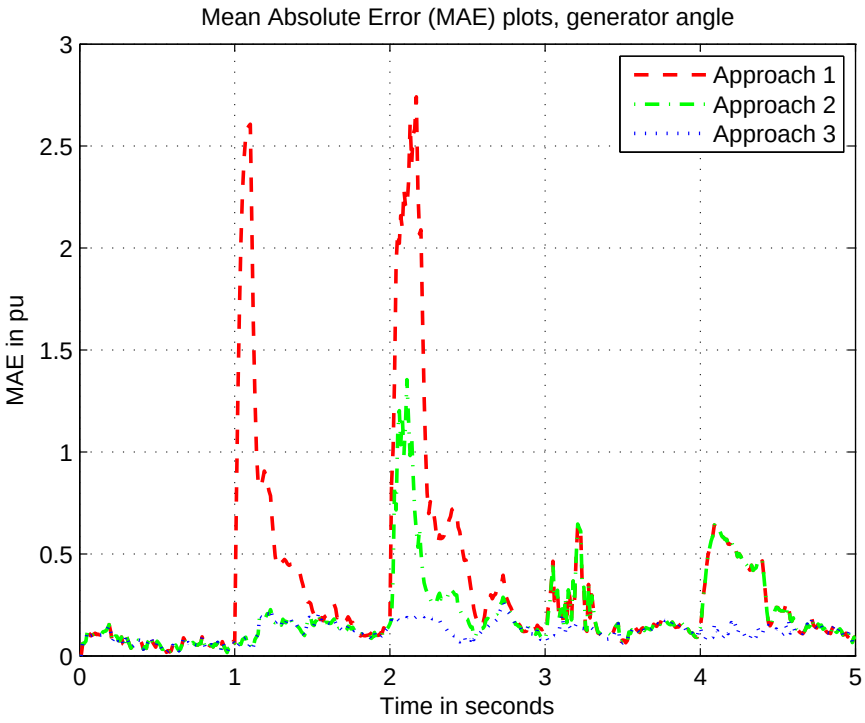


(a) Rotor speed MAE comparison of three approaches after stage two

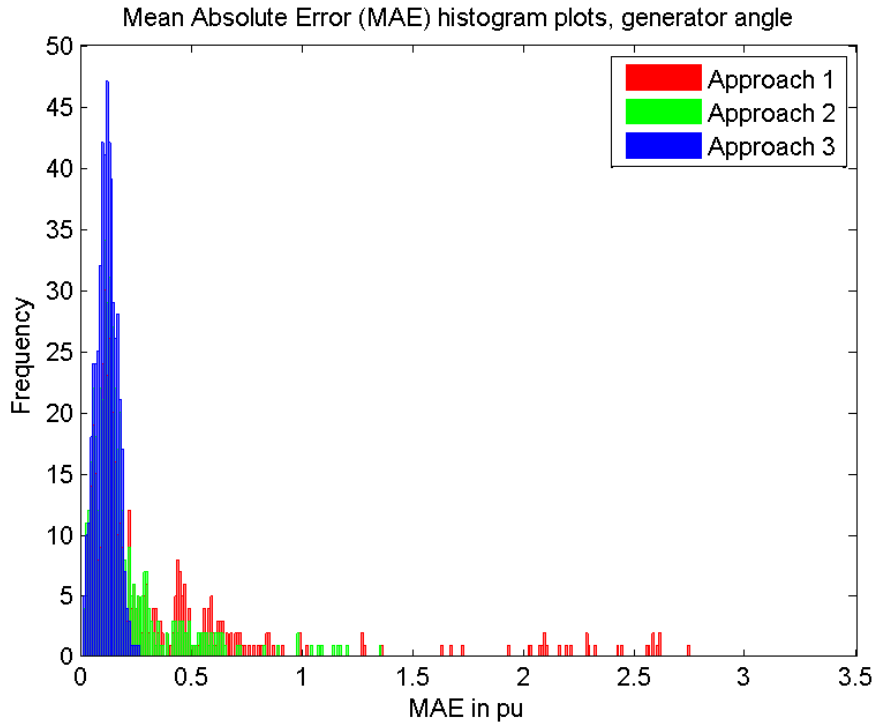


(b) Rotor speed MAE histograms of three approaches after stage two

Figure 4.39: Rotor speed estimation accuracy of three approaches in case 5



(a) Rotor angle MAE comparison of the three approaches after stage two



(b) Rotor angle MAE histograms of the three approaches after stage two

Figure 4.40: Rotor angle estimation results of three approaches in case 5

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

In this dissertation I have explored the benefits of on/off-line uncertainty analysis in three distinct but interrelated research topics, with applications in power systems:

1. *Grid Sensor Placement*

   Chapter 2 presented an approach that combines off-line uncertainty analysis and network topology analysis, to estimate the steady-state performance of candidate sensor placement design for discrete-space systems. Based on this framework, we restated the optimal PMU placement problem, applied our method on different system configurations, and visualized the results quantitatively. Taking one step further, we tied it in with dynamic state estimation to provide example solutions to the industry.

   The benefits of our approach include: 1) it can be performed efficiently without even running the actual estimation procedure; 2) it is generalized for any PMU placement design, regardless of achieving full observability.

   While we have initially chosen to work with PMU measurements only, the future work could extend the measurement set to include conventional RTU measurements. Moreover, this method is not restricted to PMU placement problems—it can be employed in different discrete-space system sensor placement optimization frameworks.

2. *Filter Computation Adaptation*

   Chapter 3 presented the *Lower Dimensional Measurement-space* (LoDiM) state estimation method, a Kalman-filter-based state estimation algorithm with a *measurement selection procedure* that can adapt to limited computational resources. Compared to the

traditional batch approach of KF/EKF methods, which handles the entire measurement-space, the LoDiM only needs to deal with a lower-dimensional measurement-space during each estimation cycle—the measurement selection procedure guides the LoDiM to dynamically choose a “most critical" (most effectively reduce the largest estimation uncertainty component) measurement subset. A smaller measurement-space incorporates less information each cycle, and hence leads to reduced computation time and increased reporting rate, which is favorable in various state tracking applications.

Specifically for power system dynamic state estimation problems, where the measurement-spaces are exceptionally large, we have proposed the *Reduced Measurement-space Dynamic State Estimation* (ReMeDySE) method derived from the LoDiM. Together with high frequency and accuracy PMU data, the ReMeDySE can facilitate dynamic state estimation in large-scale power systems.

When some emergency tasks are competing for computational resources with the estimation process, our method allows the estimation to continue with a reduced computational load, and release some resources for the tasks, without sacrificing the reporting rate. Then when the tasks are completed, the estimation process can reclaim the resources, so that the filter computation can go back up as desired. We can optimally adapt our algorithm to the available computation budget by analyzing the trade-offs.

Although the LoDiM approach is presented in the context of the Kalman filter, the associated measurement selection procedure is not filter-specific, *i.e.* it can be used with other state estimation methods such as particle and unscented filters. Selecting certain subspaces inherently delays the selection of the entire (remaining) measurement-space, which could potentially delay the observance of an important state change. One possible solution is to investigate the addition of measurement prioritization based on factors such as unexpected (deviating from prediction) state changes throughout the entire state space. For example, one can adjust the uncertainty parameters as discussed in our next topic, to affect priorities for parts of the state space, causing measurement selection to

prioritize certain areas.

With all the matrix arithmetic operations, it is possible to further parallelize and optimize the algorithm by taking advantage of modern computational hardware such as the GPUs. Moreover, we can combine our approach with hierarchical/distributed estimation methods, making it particularly attractive in large-scale real-time state estimation, for wide-area power systems and beyond.

3. *Adaptive and Robust Estimation*

   Chapter 4 presented a novel *Adaptive Kalman Filter with Inflatable Noise Variances* (AKF with InNoVa). With real-time phasor measurements, this algorithm enables convenient and on-line robust power system state estimation under various adverse conditions, given sufficient redundancy among measurements.

   The AKF with InNoVa is designed to identify and reduce the impact of incorrect system modeling, as well as erroneous measurements. It is capable of doing so because besides the regular *normalized innovation test*, we also adopt a *normalized residual test* to help separating the process and measurement factors. With these tests, our algorithm efficiently adjusts the process and measurement noise parameters on-the-fly separately for more robust estimates. Specifically, the inflation of process noise covariance Q indicates fast changing state or even wrong model, making it practical for people to localize system anomalies, *e.g.* faults, load changes, electricity theft, *etc*; while the inflation of measurement noise covariance R implies potentially bad measurements. At the same time, an exponential decay process is employed to enable automatic deflation of the parameters if the modeling/devices problems are resolved.

   To produce real-time and robust estimation of more *dynamic* states besides traditional static states, we have combined the *AKF with InNoVa* with *power system dynamic state estimation* from our previous work. The result is a more comprehensive power system state estimation technique with a two-stage Kalman filter approach. The AKF with In-

NoVa in stage one deals with system modeling and measurement errors. It gives more accurate estimates on bus voltage phasors, which are also served as the input to the EKF in stage two, to generate more accurate estimates on generator rotor speed and angles. The output of both stages are essential for myriad time-critical applications. For example in contingency analysis, signatures in different types of contingency may be detected from the change in state estimates. In fact, nowadays in the age of social media, besides the “official" measurements from the industrial control systems, public information available on the internet is a new source of measurements that can help to confirm the contingency and hence alert the operators.

The AKF with InNoVa should not be limited to power system state estimation. Furthermore, with the advance of computational hardware technologies and optimization algorithms, our approach is subject to further modifications.

REFERENCES

[1] Greg Welch, Gary Bishop, Leandra Vicci, Stephen Brumback, Kurtis Keller, and D'nardo Colucci. High-performance wide-area optical tracking: The hiball tracking system. *Teleoperators and Virtual Environments*, 10(1), February 2001.

[2] A.G. Phadke, J.S. Thorp, and M.G. Adamiak. A new measurement technique for tracking voltage phasors, local system frequency, and rate of change of frequency. *IEEE Transactions on Power Apparatus and Systems*, PAS-102(5), May 1983.

[3] A. G. Phadke, J. S. Thorp, and K. J. Karimi. State estimlatjon with phasor measurements. *IEEE Transactions on Power Systems*, 1(1), February 1986.

[4] T.L. Baldwin, L. Mili, M.B. Boisen Jr., and R. Adapa. Power system observability with minimal phasor measurement placement. *IEEE Transactions on Power Systems*, 8(2), May 1993.

[5] Jian Chen and Ali Abur. Placement of pmus to enable bad data detection in state estimation. *IEEE Transactions on Power Systems*, 21(4), November 2006.

[6] Bei Gou. Generalized integer linear programming formulation for optimal pmu placement. *IEEE Transactions on Power Systems*, 23(3), August 2008.

[7] Bei Xu and Ali Abur. Observability analysis and measurement placement for systems with pmus. In *Power Systems Conference and Exposition, 2004. IEEE PES*, October 2004.

[8] 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, USA*, September 2010.

[9] Bonnie Danette Allen and Greg Welch. A general method for comparing the expected performance of tracking and motion capture systems. *VRST '05: the ACM symposium on Virtual reality software and technology, Monterey, CA, USA*, 2005.

[10] Bonnie Danette Allen. *Hardware Design Optimization for Human Motion Tracking Systems*. Ph.d. dissertation, The University of North Carolina at Chapel Hill, Department of Computer Science, Chapel Hill, NC, USA, November 2007.

[11] Zhenyu Huang, Kevin Schneider, and Jarek Nieplocha. Feasibility studies of applying kalman filter techniques to power system dynamic state estimation. *International Power Engineering Conference 2007 (IPEC 2007)*, December 2007.

[12] Jinghe Zhang, Greg Welch, Gary Bishop, and Zhenyu Huang. Optimal pmu placement evaluation for power system dynamic state estimation. *IEEE PES Conference on Innovative Smart Grid Technologies Europe (ISGT 2010), Goteborg, Sweden*, October 2010.

[13] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Transaction of the ASME—Journal of Basic Engineering, Series D*, 82, 1960.

[14] Gregory Francis Welch. *SCAAT: Incremental Tracking with Incomplete Information*. Ph.d. dissertation, The University of North Carolina at Chapel Hill, Department of Computer Science, Chapel Hill, NC, USA, October 1996.

[15] Jinghe Zhang, Greg Welch, and Gary Bishop. Power system state estimation with dynamic optimal measurement selection. *2011 IEEE Symposium on Computational Intelligence Applications in Smart Grid, Paris, France,* April 2011.

[16] Jinghe Zhang, Greg Welch, and Gary Bishop. Lodim: A novel power system state estimation method with dynamic measurement selection. *2011 IEEE Power & Energy Society General Meeting, Detroit, MI, USA,* July 2011.

[17] Jinghe Zhang, Greg Welch, Gary Bishop, and Zhenyu Huang. Reduced measurement-space dynamic state estimation (remedyse) for power systems. *IEEE PowerTech 2011, Trondheim, Norway,* June 2011.

[18] Chakphed Madtharad, Suttichai Premrudeepreechacharn, and Neville R. Watson. Power system state estimation using singular value decomposition. *Electric Power Systems Research*, 67(2), 2003.

[19] Greg Welch and Gary Bishop. An introduction to the kalman filter. Technical report, University of North Carolina at Chapel Hill, Department of Computer Science, 1995.

[20] Mohinder S. Grewal and Angus P. Andrews. *Kalman Filtering Theory and Practice*. Information and System Sciences Series. Prentice Hall, Upper Saddle River, NJ USA, 1993.

[21] Ali Abur and Antonio Gomez Exposito. *Power System State Estimation: Theory and Implementation*. Marcel Dekker, 2004.

[22] Jin Xu, Miles H. F. Wen, Victor O. K. Li, and Ka-Cheong Leung. Optimal pmu placement for wide-area monitoring using chemical reaction optimization. *Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES*, February 2013.

[23] Ashwani Kumara, Biswarup Dasb, and Jaydev Sharmab. Dynamic state estimation of power system harmonics with bad data. *Electric Power Components and Systems*, 33(12), December 2005.

[24] Graham Rogers. *Power System Oscillation*. Academic Publishers, Boston, 2000.

[25] PNNL. Kalman equations. *Personal Communications with the Electricity Infrastructure group, Pacific Northwest National Laboratory*, 2008.

[26] Barbara Chapman, Gabriele Jost, and Ruud van der Pas. *Using OpenMP: Portable Shared Memory Parallel Programming*. the MIT Press, 2007.

[27] Jarek Nieplocha, Bruce Palmer, Manojkumar Krishnan, and P. Saddayappan. Overview of the global arrays parallel aoftware development toolkit. *2006 ACM/IEEE conference on Supercomputing SC '06, Tampa, Florida*, 2006.

[28] Anant Oonsivilai and Kenedy A. Greyson. Optimization of distributed processors for power system: Kalman filters using petri net. *World Academy of Science, Engineering and Technology 53*, 2009.

[29] M.J. Goris, D.A. Gray, and I.M.Y. Mareels. Reducing the computational load of a kalman filter. *Electronics Letters*, 33(18), August 1997.

[30] Shoujun Bian and Michael A. Henson. Measurement selection for on-line estimation of nonlinear wave models for high purity distillation columns. *Chemical Engineering Science*, 61(10), 2006.

[31] Ning Zhou, Zhenyu Huang, Greg Welch, and Jinghe Zhang. Identifying optimal measurement subspace for ensemble kalman filter. *Electronics Letters*, 48(11), May 2012.

[32] Atif S. Debs and Robert E. Larson. A dynamic estimation for tracking the state of a power system. *IEEE Transactions on Power Apparatus and Systems, vol. PAS-89, no. 7*, September/October 1970.

[33] Amit Jain and Shivakumar N. R. Power system tracking and dynamic state estimation. In *Power Systems Conference and Exposition (PSCE '09). IEEE/PES*, March 2009.

[34] 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, Alberta, Canada*, July 2009.

[35] Greg Welch and Gary Bishop. *An Introduction to the Kalman Filter*. SIGGRAPH 2001 course 8. In Computer Graphics, Annual Conference on Computer Graphics & Interactive Techniques. ACM Press, Addison-Wesley, Los Angeles, CA, USA, 2001.

[36] Frank L. Lewis. *Optimal Estimation: With an Introduction to Stochastic Control Theory*. John Wiley and Sons, Inc., 1986.

[37] Gary Bishop. *The Self-Tracker: A Smart Optical Sensor on Silicon*. Ph.d. dissertation, The University of North Carolina at Chapel Hill, Department of Computer Science, Chapel Hill, NC, USA, 1984.

[38] Ricardo Todling. Estimation theory and foundations of atmospheric data assimilation. Technical report, Data Assimilation Office, Goddard Laboratory for Atmospheres, 1999.

[39] Mohinder S. Grewal and Augus P. Andrews. *Kalman Filtering: Theory and Practice Using MATLAB*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2008.

[40] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins Univ. Press, Baltimore, MD, 1989.

[41] Saman A. Zonouz and William H. Sanders. A kalman-based coordination for hierarchical state estimation: Algorithm and analysis. *The 41st Hawaii International Conference on System Sciences*, 2008.

[42] S. Lakshminarasimhan and A. A. Girgis. Hierarchical state estimation applied to wide-area power systems. *IEEE Power Engineering Society General Meeting*, June 2007.

[43] M. S. Kurzyn. Real-time state estimation for large power systems using multicomputer systems. *Computers & Electrical Engineering*, 8(4), 1981.

[44] M. S. Kurzyn. Real-time state estimation for large-scale power systems. *IEEE Transactions on Power Apparatus and Systems*, July 1983.

[45] Jinghe Zhang, Greg Welch, Gary Bishop, and Zhenyu Huang. A two-stage kalman filter approach for robust and real-time power system state estimation. *IEEE Transactions on Sustainable Energy*, 2013.

[46] G. E. P. Box and N. R. Draper. *Empirical Model Building and Response Surfaces*. John Wiley & Sons, New York, NY, 1987.

[47] Zhongzhi Li and Xuegang Wang. Reverse prediction adaptive kalman filtering algorithm for maneuvering target tracking. *Journal of Computational Information Systems 6:10 (2010) 3257-3265.*, 2010.

[48] Kuang-Rong Shih and Shyh-Jier Huang. Application of a robust algorithm for dynamic state estimation of a power system. *IEEE Power Engineering Review , vol. 22, no. 1*, 2002.

[49] M. Oussalah and J. De Schutter. Adaptive kalman filter for noise identification. *International Conference on Noise and Vibration Engineering*, September 2000.

[50] Ashwani Kumar, Biswarup Das, and Jaydev Sharma. Robust dynamic state estimation of power system harmonics. *International Journal of Electrical Power and Energy Systems Volume 28, Issue 1*, January 2006.

[51] Jun Zhu and Ali Abur. Bad data identification when using phasor measurements. *Power Tech, 2007 IEEE Lausanne*, July 2007.

[52] Hui Xue, Qing quan Jia, Ning Wang, Zhi qian Bo, Hai tang Wang, and Hong xia Ma. A dynamic state estimation method with pmu and scada measurement for power systems. *the 8th International Power Engineering Conference (IPEC 2007), Singapore*, 2007.

[53] Amit Jain and Shivakumar N. R. Impact of pmus in dynamic state estimation of power systems. *North American Power Symposium (NAPS), Calgary, Canada*, September 2008.

[54] Antonio Gomez-Exposito, Ali Abur, Patricia Rousseaux, Antonio de la Villa Jaen, and Catalina Gomez-Quiles. On the use of pmus in power system state estimation. *17th Power Systems Computation Conference, Stockholm, Sweden*, August 2011.

[55] Power System Relaying Committee. C37.118.1-2011: Ieee standard for synchrophasor measurements for power systems. *IEEE Standard Association*, December 2011.

[56] J. Peppanen, T. Alquthami, D. Molina, and R. Harley. Optimal pmu placement with binary pso. *IEEE Energy Conversion Congress and Exposition (ECCE)*, 2012.

[57] F. Aminifar, M. Fotuhi-Firuzabad, and A. Safdarian. Optimal pmu placement based on probabilistic cost/benefit analysis. *IEEE Transactions on Power Systems*, February 2013.

[58] Paul D. Groves. *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems*. Artech House, 2013.