A Robust Dynamic State and Parameter Estimation Framework for Smart Grid Monitoring and Control

Junbo Zhao

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

Doctor of Philosophy
in
Electrical Engineering

Lamine Mili, Chair
Saifur Rahman
Adrian Sandu
Yue Wang
Vasileious Kekatos

May 2, 2018
Arlington, Virginia

Keywords: Kalman filter, Robust statistics, Dynamic state estimation, Unscented transformation, Robust control theory, Estimation theory, Power system dynamics and control, Outliers, Cyber attacks, Phasor measurement units

A Robust Dynamic State and Parameter Estimation Framework for Smart Grid Monitoring and Control

Junbo Zhao

(ABSTRACT)

The enhancement of the reliability, security, and resiliency of electric power systems depends on the availability of fast, accurate, and robust dynamic state estimators. These estimators should be robust to gross errors on the measurements and the model parameter values while providing good state estimates even in the presence of large dynamical system model uncertainties and non-Gaussian thick-tailed process and observation noises. It turns out that the current Kalman filter-based dynamic state estimators given in the literature suffer from several important shortcomings, precluding them from being adopted by power utilities for practical applications. To be specific, they cannot handle (i) dynamic model uncertainty and parameter errors; (ii) non-Gaussian process and observation noise of the system nonlinear dynamic models; (iii) three types of outliers; and (iv) all types of cyber attacks. The three types of outliers, including observation, innovation, and structural outliers are caused by either an unreliable dynamical model or real-time synchrophasor measurements with data quality issues, which are commonly seen in the power system.

To address these challenges, we have pioneered a general theoretical framework that advances both robust statistics and robust control theory for robust dynamic state and parameter estimation of a cyber-physical system. Specifically, the generalized maximum-likelihood-type (GM)-estimator, the unscented Kalman filter (UKF), and the H-infinity filter are integrated into a unified framework to yield various centralized and decentralized robust dynamic state estimators. These new estimators include the GM-iterated extended Kalman filter (GM-IEKF), the GM-UKF, the H-infinity UKF and the robust H-infinity UKF. The GM-IEKF is able to handle observation and innovation outliers but its statistical efficiency is low in the presence of non-Gaussian system process and measurement noise. The GM-UKF addresses this issue and achieves a high statistical efficiency under a broad range of non-Gaussian process and observation noise while maintaining the robustness to observation and innovation outliers. A reformulation of the GM-UKF with multiple hypothesis testing further enables it to handle structural outliers. However, the GM-UKF may yield biased state estimates in presence of large system uncertainties. To this end, the H-infinity UKF that relies on robust control theory is proposed. It is shown that H-infinity is able to bound the system uncertainties but lacks of robustness to outliers and non-Gaussian noise. Finally, the robust H-infinity filter framework is proposed that leverages the H-infinity criterion to bound system uncertainties while relying on the robustness of GM-estimator to filter out non-Gaussian noise and suppress outliers. Furthermore, these new robust estimators are applied for system bus frequency monitoring and control and synchronous generator model parameter calibration. Case studies of several different IEEE standard systems show the efficiency and robustness of the proposed estimators.

A Robust Dynamic State and Parameter Estimation Framework for Smart Grid Monitoring and Control

Junbo Zhao

(GENERAL AUDIENCE ABSTRACT)

The enhancement of the reliability, security, and resiliency of electric power systems depends on the availability of fast, accurate, and robust dynamic state estimators. These estimators should be robust to gross errors on the measurements and the model parameter values while providing good state estimates even in the presence of large dynamical system model uncertainties and non-Gaussian thick-tailed process and observation noises. There are three types of gross errors or outliers, namely, observation, innovation, and structural outliers. They can be caused by either an unreliable dynamical model or real-time synchrophasor measurements with data quality issues, which are commonly seen in the power system. The system uncertainties can be induced in several ways, including i) unknowable system inputs, such as noise, parameter variations and actuator failures, to name a few; ii) unavailable inputs, such as unmeasured mechanical power, field voltage of the exciter, unknown fault location; and iii) inaccuracies of the model parameter values of the synchronous generators, the loads, the lines, and the transformers, to name a few. It turns out that the current Kalman filter-based dynamic state estimators suffer from several important shortcomings, precluding them from being adopted by power utilities for practical applications.

To address these challenges, this dissertation has proposed a general theoretical framework that advances both robust statistics and robust control theory for robust dynamic state and parameter estimation. Specifically, the robust generalized maximum-likelihood-type (GM)-estimator, the nonlinear filter, i.e., unscented Kalman filter (UKF), and the H-infinity filter are integrated into a unified framework to produce various robust dynamic state estimators. These new estimators include the robust GM-IEKF, the robust GM-UKF, the H-infinity UKF and the robust H-infinity UKF. Specifically, the GM-IEKF deals with the observation and innovation outliers but achieving relatively low statistical efficiency in the presence of non-Gaussian system process and measurement noise. To address that, the robust GM-UKF is proposed that is able to achieve a high statistical efficiency under a broad range of non-Gaussian noise while maintaining the robustness to observation and innovation outliers. A reformulation of the GM-UKF with multiple hypothesis testing further enables it to handle three types of outliers. However, the GM-UKF may yield biased state estimates in presence of large system uncertainties. To this end, the H-infinity UKF that depends on robust control theory is proposed. It is able to bound the system uncertainties but lacks of robustness to outliers and non-Gaussian noise. Finally, the robust H-infinity filter framework is proposed that relies on the H-infinity criterion to bound system uncertainties while leveraging the robustness of GM-UKF to filter out non-Gaussian noise and suppress outliers. These new robust estimators are applied for system bus frequency monitoring and control and synchronous generator model parameter calibration. Case studies of several different IEEE standard systems show the efficiency and robustness of the proposed estimators.

Acknowledgments

I would like to express my sincere gratitude to my advisor and mentor Dr. Lamine Mili for offering me the opportunity to work on this topic at Virginia Tech, and the continuous support of my Ph.D. studies and research. I am very grateful for his patience, motivation, enthusiasm, and immense knowledge. Without his guidance and help, I may not complete such a successful PhD dissertation. I could not imagine having a better advisor for my Ph.D. studies. Additionally, I would like to thank for the rest of my committee members, Dr. Saifur Rahman, Dr. Adrian Sandu, Dr. Yue Wang and Dr. Vassilis Kekatos for their encouragement, insightful comments, and useful feedbacks.

My sincere thanks also goes to Dr. Zhenyu Huang, Dr. Ruisheng Diao and Dr. Shaobu Wang for offering me the summer internship opportunities at PNNL and giving me a lot help on this research topic. In addition, special thanks to my colleagues, Dr. Rabih Jabr, Dr. Federico Milano, Dr. Massimo La Scala, Dr. Robson Pires, Marcos Netto, Yijun Xu, Srivats Shukla, etc., who have supported me over years.

Finally, I would like to express my heartfelt thanks to my parents, my wife, Dr. Fangni Lei, for their love, support and encouragement.

This work is supported in part by the U.S. National Science Foundation under Grant ECCS-1711191; the Advanced Scientific Computing Research (ASCR) program of the U.S. Department of Energy (DOE) Office of Science; the Transmission Reliability Program of the U.S. Department of Energy (DOE) Office of Electricity (GMLC0070) and the Pacific Northwest National Laboratory.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The widespread deployment of synchrophasor measurement units (PMUs) on power transmission grids has made possible the real-time monitoring and control of power system dynamics. However, these functions may not be reliably achieved without the development of fast and robust dynamic state estimator (DSE). Indeed, the benefits of using a DSE include:

- Improved system oscillation monitoring [1–6]: a reliable estimation of power system state variables will allow us to track the fast dynamics taking place in a power grid and to carry out, for example, Koopman mode analysis [1–3] or Prony analysis [4–6]. The identified modes can then be utilized as control inputs to tune power system stabilizers (PSS) to achieve better damping of inter-area modes of oscillation, resulting in improved system stability.
- Enhanced local and global system control [7–11]: for instance, the estimated rotor speeds and other estimated state variables can be used as input control signals of FACTS devices to ensure global asymptotic stability of a power system by satisfying the Lyapunov sufficient conditions for stability.
- Improved reliability of protection systems [12–16]: for instance, by testing in real-time the consistency between the PMU measurements and the dynamical model of the protection zone identified by the DSE, both internal and external faults can be effectively detected without any a priori protection settings, yielding more reliable protection systems compared with the traditional coordinated settings-based schemes [12, 13]; the estimated real-time dynamic state variables can be used to initiate effective generator out-of-step protections based on the equal-area criterion or the energy function approach [15, 16].
- Enhanced reliability of dynamic security assessment (DSA) [18, 19, 135]: as highlighted by a PSERC team [19], DSA requires the availability of accurate models for generators and associated controls, loads, special protection schemes, etc. By utilizing a DSE, the dynamic system model and its associated parameters can be validated on-line [20], yielding more reliable DSA.

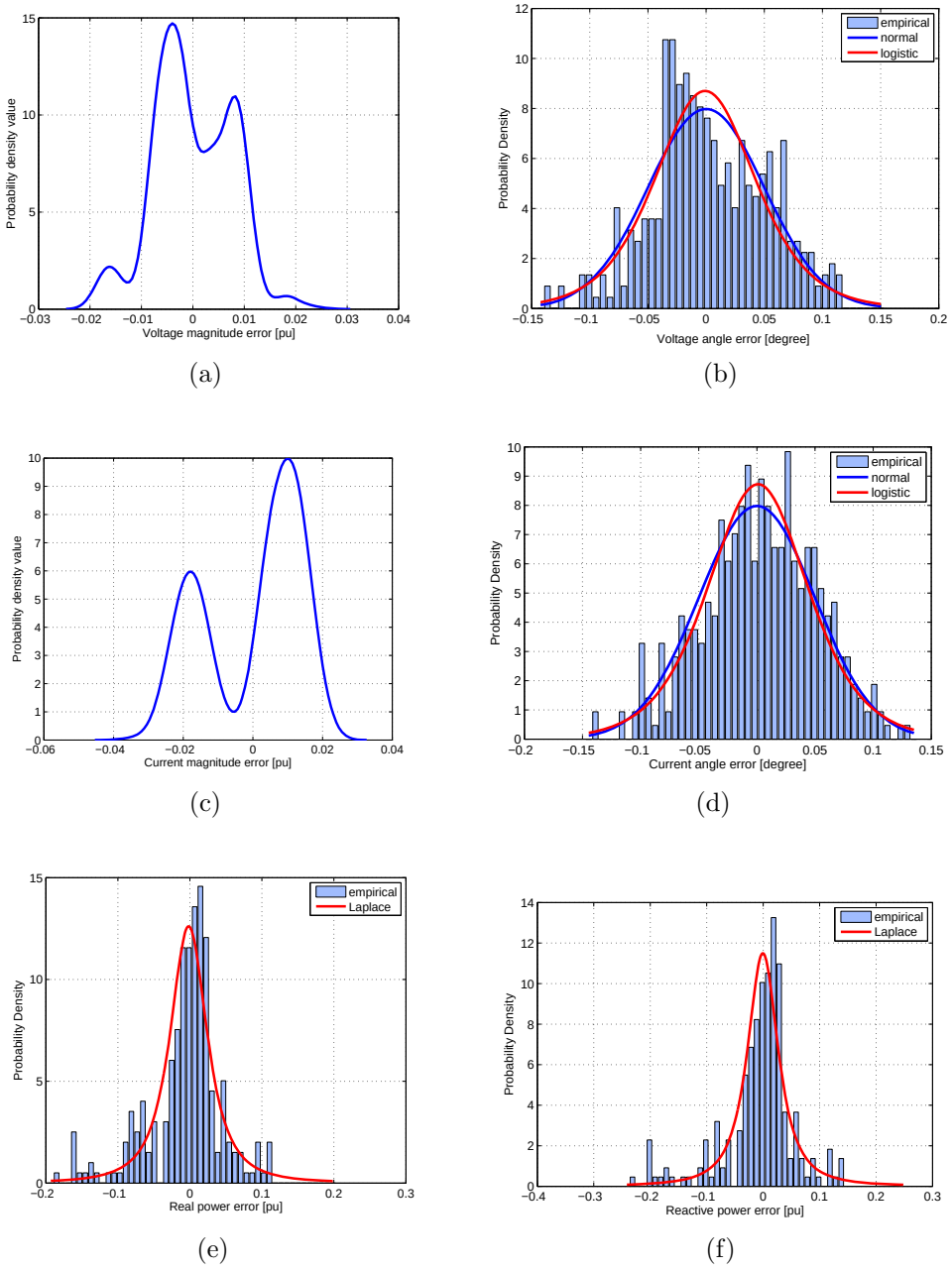


Figure 1.1: Error distributions of the phasor voltage angle and magnitude, the phasor current angle and magnitude, the real and reactive power using field PMU data provided by Pacific Northwest National Laboratory.

To date, a variety of dynamic state estimators have been proposed in the literature; they are based on the extended Kalman filter (EKF) [21–23], the iterated EKF (IEKF) [24], the unscented Kalman filter (UKF) [25–27], and the particle filter (PF) [28–30], to cite a

few. However, all these methods suffer from several important shortcomings, precluding them from being adopted by power utilities for power system real-time applications. To be specific, they cannot handle **i) dynamical model uncertainty and parameter errors; ii) non-Gaussian process and observation noise of the system nonlinear dynamic models; iii) any type of outliers, including innovation, observation and structural outliers; iv) all types of cyber attacks**.

There are several reasons for these shortcomings. Firstly, besides the uncertainties induced by the inherent approximations of the mathematical models on which these methods rely, there are several other uncertainties due to the following: i) unknowable system inputs, including noise, parameter variations and actuator failures [136], to name a few; ii) unavailable inputs, such as unmeasured mechanical power, field voltage of the exciter, unknown fault location; and iii) inaccuracies of the model parameter values of the synchronous generators, the loads, the lines, and the transformers, to name a few. To cope with these uncertainties, we will develop new DSEs based on robust generalized maximum-likelihood (GM) UKF combined with H-infinity filtering techniques. The GM-UKF is able to filter out non-Gaussian noise and suppress gross errors in the measurements, in the predictions and in the model parameters and system topologies, which will be called observation, innovation and structural outliers. But this filter produces strongly biased estimation results in presence of a large number of inaccurate model parameters. On the other hand, H-infinity filter cannot handle any type of outliers but it provides good results in presence of system uncertainties [32–34].

Secondly, the current DSE methods assume that both the process and the observation noise of the system nonlinear dynamic models are Gaussian. However, two recent investigations conducted by PNNL [35, 36] revealed that the PMU measurement errors of the voltage and current magnitudes follow non-Gaussian probability distributions. This is demonstrated in Fig.1.1 using real PMU data provided to us by PNNL. This figure displays histograms and parametric probability density estimates of PMU errors on nodal voltage magnitudes and angles, line current magnitudes and angles, and line real and reactive power. As observed in Fig.1.1, except for the measurement errors on nodal voltage and line current angles, which roughly follow a Gaussian distribution, the measurement errors on both nodal voltage and line current magnitudes follow a bimodal Gaussian mixture distribution. As for the measurement errors of real and reactive power calculated by voltage and current phasors, they follow a thick tailed distribution that may be approximated by either the Laplace or the Cauchy distribution. Recall that in contrast to the Gaussian distribution, which is a short-tailed distribution, a thick-tailed distribution is the one that allows the associated random variable to take large values with a non-negligible probability. Evidently, the presence of non-Gaussian noise calls for new research and development in robust power system dynamic state estimation based on robust statistics.

Thirdly, three types of outliers associated with a given dynamical system model have been defined by Gandhi and Mili [37], namely observation outliers, which affect the metered values, innovation outliers, which corrupt the predicted state estimates, and structural outliers, which affect the system dynamic states and the observation functions. Observation outliers

may result from large biases in PMU measurements due to infrequent calibration, or instrument failures, or impulsive communication noise [39–41]. As for innovation outliers, they may occur in several different ways. For example, some of the generator models may not be well calibrated, resulting in highly inaccurate model outputs that are inconsistent with the measurements. This was precisely the case in the 1996 blackout, where the model being used predicted system stability while in reality the system was undergoing numerous cascading failures, which resulted in a rapid system collapse that occurred within minutes [42, 43]. Innovation outliers may also be induced by the approximations in the state prediction model or by a system process impulsive noise. By contrast, structural outliers are gross errors in circuit breaker statuses or in the parameters of the transmission lines, or the synchronous machines, or the automatic voltage regulators. In [44], it is reported that wrong estimates of the parameters of the synchronous machine models may result from the use of erroneous metered values. It turns out that the conventional filters, namely the EKF, the UKF and the PF are not robust to any type of outliers. For instance, it is demonstrated in [45] that their performances are significantly degraded in presence of observation outliers. In this research project, we will develop new robust DSE methods based on the generalized maximum likelihood-type estimator that are able to handle the three types of outliers listed above.

Last but not least, with the strong reliance of smart grid functions on communications networks, cyber attacks have become a major concern. Typically, they are classified as bias injection attack, denial of service attack, and replay attack [46]. Bias injection attack occurs when an adversary attempts to corrupt the content of either the measurement or the control signals; for example, the man-in-the-middle intercepts the PMU measurement signals and corrupts them with large biases. Denial of service attack occurs when the actuator and sensor data are prevented from reaching their respective destinations, resulting in the absence of data for the DSE; for instance, this would be the case if the PMU metered values do not reach the phasor data concentrator (PDC). Replay attack occurs when a hacker first performs a disclosure attack from a certain time period, gathering sequences of data, and then begins replaying the data during a certain period; for instance, the current PMU measurements processed by a dynamic state estimator are replaced by past values. Those attacks will induce outliers, or loss of measurements, or uncertainties in the model. Obviously, mitigating the impacts of cyber attacks is an important task that calls for the development of a robust DSE.

1.1 Research Objectives and Achievements

The enhancement of the reliability, security, and resiliency of electric power systems depends on the availability of fast, accurate, and robust dynamic state estimators. These estimators should be robust to gross errors on the measurements and the model parameter values while providing good state estimates even in the presence of large dynamical system model uncertainties and non-Gaussian thick-tailed process and observation noises. The research

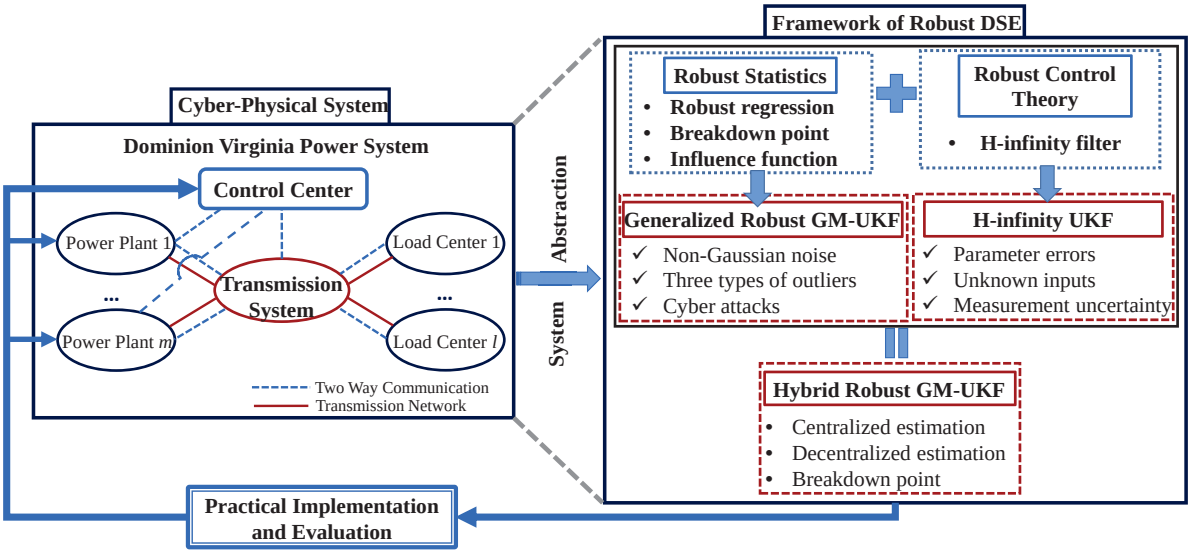


Figure 1.2: The research scheme of this dissertation for power system monitoring and control.

scheme is summarized in Fig. 1.2. To this end, we have pioneered a general theoretical framework that advances both robust statistical theory and robust control theory for robust dynamic state and parameter estimation of a cyber-physical system. Specifically, the main contribution of this dissertation is outlined in below:

1. The robust statistics and the robust control theory are integrated into a unified framework to yield various centralized and decentralized robust dynamic state estimators. These new estimators are able to handle large system uncertainties as well as suppress three types of outliers while achieving good statistical efficiency under a broad range of non-Gaussian process and observation noise.

2. A centralized robust iterated extended Kalman filter (EKF) based on the generalized maximum likelihood approach (termed GM-IEKF) is proposed for estimating power system state dynamics when subjected to disturbances. The proposed GM-IEKF dynamic state estimator is able to track system transients in a faster and more reliable way than the conventional EKF and the unscented Kalman filter (UKF) thanks to its batch-mode regression form and its robustness to innovation and observation outliers, even in position of leverage. Good robustness and high statistical efficiency under Gaussian noise are achieved.

3. The theory of a new robust Generalized Maximum-likelihood-type Unscented Kalman Filter (GM-UKF) is developed that is able to suppress observation and innovation outliers while filtering out non-Gaussian process and measurement noise. Because the errors of the real and reactive power measurements calculated using Phasor Measurement Units (PMUs) follow long-tailed probability distributions, the conventional UKF

provides strongly biased state estimates since it relies on the weighted least squares estimator. By contrast, the state estimates and residuals of our GM-UKF are proved to be roughly Gaussian, allowing the sigma points to reliably approximate the mean and the covariance matrices of the predicted and corrected state vectors. The asymptotic error covariance matrix of the GM-UKF state estimates is derived from the total influence function. The robustness our GM-UKF to outliers and cyber attacks is investigated; it is shown that it can handle *at least* 25% of incorrect PMU measurements and predicted state variables.

4. A fast and robust unscented Kalman filter-based decentralized dynamic state estimator (DSE) using multiple hypothesis testing is proposed to detect, identify and suppress three types of outliers, namely the observation, innovation and structural outliers. Specifically, three hypotheses corresponding to the occurrence of three types of outliers are assumed by constructing three innovation matrices; these matrices are made up by time-correlated innovation vectors and/or predicted states and/or measurements; then projection statistics are applied to each of the innovation matrix and its calculated projection values are checked by a statistical test to validate the assumed hypothesis. The identified outliers are further suppressed by a generalized maximum-likelihood type (GM)-estimator.

5. Accurate local bus frequency is essential for power system frequency regulation provided by distributed energy sources, flexible loads and among others. This dissertation proposes a robust frequency divider (RFD) for online bus frequency estimation and control. Our RFD is independent of the load models, and the knowledge of swing equation parameters, transmission line parameters and local PMU measurements at each generator terminal bus is sufficient. In addition, it is able to handle several types of data quality issues, such as measurement noise, gross measurement errors, cyber attacks and measurement losses.

6. The existing Kalman filter-type power system dynamic state estimator (DSE) achieves a good performance if the dynamical model is accurate and the system process and measurement noises are known. However, these assumptions may not be satisfied in practice as the dynamical model and noises are usually subject to uncertainties. To address these problems, this dissertation proposes a decentralized H-infinity unscented Kalman filter that leverages the strength of the H-infinity criteria developed in robust control for handling system uncertainties with the advantage of the UKF for addressing strong model nonlinearities. The statistical linerization approach is used to derive a linear batch-mode regression model similar to the linear Kalman filter, which allows us to derive the H-infinity UKF.

7. A novel theoretical framework for robust dynamic state estimation is proposed that integrates both robust statistics and robust control theory. The GM-estimator, the unscented Kalman filter (UKF), and the H-infinity filter are integrated into a unified framework to yield the general robust H-infinity UKF. The latter is able to handle

large system uncertainties as well as suppress outliers while achieving good statistical efficiency under a broad range of non-Gaussian process and observation noise. Specifically, it leverages the H-infinity criterion to bound system uncertainties while relying on the robustness of GM-estimator to filter out non-Gaussian noise and suppress outliers. We also show that the H-infinity UKF is based on the Krein space least squares estimator and thus lacks robustness to outliers and non-Gaussian noise.

1.2 Dissertation Organization

The remainder of this dissertation is organized as follows. A centralized robust GM-IEKF is proposed in Chapter 2 to demonstrate the value of robust statistics in handling innovation and structural outliers. The GM-IEKF is able to track the power system dynamics reliably if the system nonlinearity is not very strong. In addition, its statistical efficiency for non-Gaussian system process and observation noise is not very high. To address this issue, we develop the GM-UKF in Chapter 3. Its theory is first built by resorting to the robust statistics and the statistical linearization strategy. Then, its break down point and value of power system DSE application is investigated. Although GM-UKF is able to address innovation and structural outliers while achieving a high statistical efficiency under a broad range of non-Gaussian process and observation noise, it fails to suppress the structural outliers that are typically caused by model parameters of generator, excitor, governor or power system stabilizer, etc. To this end, a decentralized robust DSE using multiple hypothesis testing is proposed in Chapter 4. Its extension to power system bus frequency monitoring is shown in Chapter 5. It is worth pointing out that thanks to its statistical robustness, the GM-UKF is able to handle outliers and non-Gaussian noise. However, it may produce large estimation biases in presence of model uncertainties or inaccurate model parameters. To address these problems, Chapter 6 presents a decentralized H-infinity unscented Kalman filter that leverages the strength of the H-infinity criteria developed in robust control for handling system uncertainties with the advantage of the UKF for addressing strong model nonlinearities. Although H-infinity unscented Kalman filter can handle model and observation uncertainties, it lacks robustness to any type of outliers. To this end, Chapter 7 proposes a novel theoretical framework that integrates both robust statistics and robust control theory. The GM-estimator, the unscented Kalman filter (UKF), and the H-infinity filter are integrated into a unified framework to yield the robust H-infinity UKF. It leverages the H-infinity criterion to bound system uncertainties while relying on the robustness of GM-estimator to filter out non-Gaussian noise and suppress outliers. Finally, several interesting future research topics unveiled during this research efforts are discussed in Chapter 8.

Chapter 2

A Centralized Robust Iterated Extended Kalman Filter for Power System Dynamic State Estimation

2.1 Introduction

With the widespread deployment of wide-area synchrophasor measurements, new techniques have been proposed to effectively track power system dynamics [21, 55, 56] via a dynamic state estimator (DSE). By using the estimated system dynamic states, improved real-time control schemes can be implemented, for example through actions on FACTS devices and wide-area power system stabilizers, to cite a few, thus enhancing power system stability.

For tracking power system dynamics subject to large disturbances, the extended Kalman filter (EKF) and the unscented Kalman filter (UKF) have been recently proposed. For instance, Huang *et al.* [21] and Fang and Wehbe [24] investigated the benefits of adopting PMU data for real-time state and parameter estimation using Kalman filter techniques. Following their work, Ghahremani and Kamwa [22] proposed a modified EKF-based DSE to cope with cases where the field voltage is not accessible to metering due to brushless excitation systems. This work was later extended with the development of a decentralized DSE while relaxing the assumption of a known mechanical torque [23].

To circumvent first-order approximation errors of the EKF, which may be large under strong nonlinearities of the model, the iterated EKF (IEKF) [24] and the UKF have been proposed as alternative methods. Specifically, the IEKF linearizes the system nonlinear equations iteratively to compensate for the higher-order terms, whereas the UKF leverages the unscented transformation by deterministically providing sigma points to approximate the mean and covariance matrix of a random state vector, thus achieving better accuracy than the EKF [57, 58]. An UKF-based DSE using a fourth-order generator model is proposed in [59]

to estimate the states of a single-machine infinite-bus power system. Along the same lines, a centralized UKF is developed in [25] for a multi-machine system, while a decentralized strategy that does not require transmission of local signals is advocated in [26], significantly increasing the computational efficiency. However, it has been demonstrated in [45] that the performance of either EKF or UKF is greatly degraded in the presence of observation outliers due to their lack of robustness. To mitigate this issue, a normalized innovation vector-based test is advocated in [26] to detect observation outliers despite the vulnerability of this test to innovation outliers. In [27], Rouhani and Abur developed a distributed two-stage robust UKF-based DSE using the least-absolute-value (LAV) estimator that can handle observation outliers in PMU measurements. However, the authors do not address the vulnerability of the dynamic state estimator to innovation outliers that are induced by approximations in the state prediction model or by impulsive system process noise.

Gandhi and Mili [37] proposed a robust Kalman filter for linear dynamical model. In this chapter [38], we have extended that work to general nonlinear dynamic state estimation problems with several new features. The latter includes the derivations of a new batch-mode regression form to enhance data redundancy, a new outlier detection method based on projection statistics and applied to two time-sequence of the prediction and the innovation vectors, and the robust state covariance matrix of the proposed robust generalized maximum likelihood iterated EKF (GM-IEKF) method. These features allow our GM-IEKF to track power system dynamics more reliably and more rapidly than the conventional EKF, even in presence of observation and innovation outliers or non-Gaussian PMU noise.

The rest of the paper is organized as follows. Section 2.2 deals with the problem formulation. Section 2.3 presents the proposed GM-IEKF, while Section 2.4 shows numerical results on the IEEE 39-bus test system. Section 2.5 concludes the paper.

2.2 Problem formulation

2.2.1 Problem statement

Our objective is to estimate the system state variables using the discrete-time state space model with additive white noise that is expressed as

$$\boldsymbol{x}_k = \boldsymbol{f}\left(\boldsymbol{x}_{k-1}\right) + \boldsymbol{w}_k, \tag{2.1}$$

$$\boldsymbol{z}_k = \boldsymbol{h}\left(\boldsymbol{x}_k\right) + \boldsymbol{v}_k, \tag{2.2}$$

where $\boldsymbol{x}_k \in \mathbb{R}^n$ is the state vector; $\boldsymbol{z}_k \in \mathbb{R}^m$ is the measurement vector composed by terminal active and reactive power, and voltage magnitude and angle obtained or calculated from PMUs; $\boldsymbol{h}(\cdot)$ is the vector-valued measurement function; $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are process and measurement noise, respectively, and are assumed to be white and independent of each other, i.e., $\mathbb{E}\left[\boldsymbol{w}_k\right] = \boldsymbol{0}$, $\mathbb{E}\left[\boldsymbol{w}_k \boldsymbol{w}_k^T\right] = \boldsymbol{W}_k$, $\mathbb{E}\left[\boldsymbol{v}_k\right] = \boldsymbol{0}$, $\mathbb{E}\left[\boldsymbol{v}_k \boldsymbol{v}_k^T\right] = \boldsymbol{R}_k$, $\mathbb{E}\left[\boldsymbol{v}_j \boldsymbol{v}_k^T\right] = \mathbb{E}\left[\boldsymbol{w}_k \boldsymbol{w}_j^T\right] =$

$\mathbf{0}$ $(j \neq k)$ and $\mathbb{E}\left[\boldsymbol{w}_k \boldsymbol{v}_j^T\right] = \mathbf{0}$; $\boldsymbol{f}(\cdot)$ is the vector-valued function that relates $\boldsymbol{x}_k$ to $\boldsymbol{x}_{k-1}$. In this chapter, the classical generator model and the more detailed two-axis model with IEEE-DC1A exciter and TGOV1 turbine-governor are implemented and tested. When the classical generator model is considered, $\boldsymbol{f}(\cdot)$ is expressed as follows:

Differential equations:

$$\frac{d\delta}{dt} = \omega - \omega_s, \tag{2.3}$$

$$\frac{2H}{\omega_s}\frac{d\omega}{dt} = T_M - P_e - D\left(\omega - \omega_s\right), \tag{2.4}$$

Algebraic equation:

$$P_e = \frac{|E|}{X'_d}(|V_j|\sin\delta\cos(\theta_j) - |V_j|\cos\delta\sin(\theta_j)), \tag{2.5}$$

where the subscript j denotes the j-th generator bus; δ denotes the rotor angle in radians; ω and ω_s are respectively actual and synchronous rotor speed in radians/s; T_M, H and D denote the mechanical power input, the inertia constant and the damping factor, respectively; P_e is the real power output; $|E|\angle\delta$ represents the voltage behind the transient reactance X'_d, and $|V_j|\angle\theta_j$ is the bus voltage.

When the detailed model is considered, $\boldsymbol{f}(\cdot)$ is expressed as follows:

Differential equations:

$$T'_{do}\frac{dE'_q}{dt} = -E'_q - \left(X_d - X'_d\right)I_d + E_{fd}, \tag{2.6}$$

$$T'_{qo}\frac{dE'_d}{dt} = -E'_d - \left(X_q - X'_q\right)I_q, \tag{2.7}$$

$$\frac{d\delta}{dt} = \omega - \omega_s, \tag{2.8}$$

$$\frac{2H}{\omega_s}\frac{d\omega}{dt} = T_M - P_e - D\left(\omega - \omega_s\right), \tag{2.9}$$

$$T_E\frac{dE_{fd}}{dt} = -\left(K_E + S_E\left(E_{fd}\right)\right)E_{fd} + V_R, \tag{2.10}$$

$$T_{\mathrm{F}}\frac{dV_F}{dt} = -V_F + \frac{K_F}{T_E}V_R - \frac{K_F}{T_E}\left(K_E + S_E\left(E_{fd}\right)\right)E_{fd}, \tag{2.11}$$

$$T_A\frac{dV_R}{dt} = -V_R + K_A\left(V_{ref} - V_F - V\right), \tag{2.12}$$

$$T_{CH}\frac{dT_M}{dt} = -T_M + P_{SV}, \tag{2.13}$$

$$T_{SV}\frac{dP_{SV}}{dt} = -P_{SV} + P_C - \frac{1}{R_D}\left(\frac{\omega}{\omega_s} - 1\right), \tag{2.14}$$

Algebraic equations:

$$V_d = V\sin(\delta-\theta), V_q = V\cos(\delta-\theta), \tag{2.15}$$

$$I_d = \frac{E'_q - V_q}{X'_d}, I_q = \frac{V_d - E'_d}{X'_q}, \tag{2.16}$$

$$P_e = V_d I_d + V_q I_q, Q_e = -V_d I_q + V_q I_d, \tag{2.17}$$

where T'_{do}, T'_{qo}, T_E, T_F, T_A, T_{CH} and T_{SV} are time constants, in seconds; K_E, K_F and K_A are controller gains; V_{ref} and P_C are known control inputs [60]; E'_q, E'_d, E_{fd}, V_F, V_R, T_M and P_{SV} are the q-axis and d-axis transient voltages, field voltage, scaled output of the stabilizing transformer and scaled output of the amplifier, synchronous machine mechanical torque and steam valve position, respectively; X_d, X'_d, X_q and X'_q are generator parameters; V and θ are the terminal bus voltage magnitude and phase angle, respectively; P_e and Q_e are the active and reactive electrical power outputs; I_d and I_q are the d and q axis currents, respectively.

The above estimation problem can be solved using the well-known EKF, which is a two-step prediction-correction process; it can be summarized as follows:

State Prediction

Calculate the one-step prediction of the system state along with the associated covariance matrix of the prediction state estimation error. Formally, we have

$$\hat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{f}(\hat{\boldsymbol{x}}_{k-1|k-1}), \tag{2.18}$$

$$\boldsymbol{\Sigma}_{k|k-1} = \boldsymbol{F}_{k-1}\boldsymbol{\Sigma}_{k-1|k-1}\boldsymbol{F}_{k-1}^T + \boldsymbol{W}_k, \tag{2.19}$$

where $\boldsymbol{F}_{k-1} = {}^{\partial \boldsymbol{f}}/_{\partial \boldsymbol{x}}\big|_{\boldsymbol{x}=\widehat{\boldsymbol{x}}_{k-1|k-1}}$ is the Jacobian matrix; $\boldsymbol{\Sigma}_{k|k-1} = \mathbb{E}\left[\left(\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k-1}\right)\left(\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k-1}\right)^T\right]$ is the state prediction error covariance matrix and can be derived using a first-order Taylor series expansion of $\boldsymbol{f}\left(\boldsymbol{x}_{k-1}\right)$ about $\widehat{\boldsymbol{x}}_{k|k-1}$.

State Correction

Calculate the Kalman filter gain and update the state estimate and the estimation error covariance matrix using

$$\boldsymbol{K}_k = \boldsymbol{\Sigma}_{k|k-1}\boldsymbol{H}_k^T(\boldsymbol{H}_k\boldsymbol{\Sigma}_{k|k-1}\boldsymbol{H}_k^T + \boldsymbol{R}_k)^{-1}, \tag{2.20}$$

$$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k[\boldsymbol{z}_k - \boldsymbol{h}(\hat{\boldsymbol{x}}_{k|k-1})], \tag{2.21}$$

$$\boldsymbol{\Sigma}_{k|k} = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{\Sigma}_{k|k-1}, \tag{2.22}$$

where $\boldsymbol{H}_k = {}^{\partial \boldsymbol{h}}/_{\partial \boldsymbol{x}}\big|_{\boldsymbol{x}=\widehat{\boldsymbol{x}}_{k|k-1}}$.

2.2.2 Motivation for the Use of a Robust Dynamic State Estimator

As pointed out by Gandhi and Mili [37], outliers can be classified into three types, namely observation, innovation and structural outliers as shown in Table I. In power system dynamic state estimation, observation outliers may originate from either communication impulsive noise, or the loss of the PMU's communications links, or an imperfect phasor synchronization in PMU measurements [61, 62], or large biases caused by the saturation of metering current transformers or by metering coupling capacitor voltage transformers (CCVTs), to name a few. As for the innovation outliers, they may be induced in several different ways. For instance, some of the generators are not well calibrated and their simulated outputs are not consistent with the measurements, like the 1996 blackout, where the model predicted a stable system state, while in reality it was highly unstable, resulting in the collapse of the system within minutes [42, 43]. Innovation outliers may also be induced by the approximations in the state prediction model or impulsive system process noise [63,64]. By contrast, structural outliers may be caused by system parameter or topological errors. In this chapter, we propose a robust GM-IEKF that can handle observation and innovation outliers but not structural outliers, which require a different formulation that is not addressed here.

Table 2.1: Definition of the three types of outliers

Outlier type	Types of noise	Affected components
Observation outlier	Observation noise, $\boldsymbol{v}_k$	$\boldsymbol{z}_k$
Innovation outlier	System process noise, $\boldsymbol{w}_k$	$\hat{\boldsymbol{x}}_{k\|k-1}$
Structural outlier	Structural errors in $\boldsymbol{f}(\boldsymbol{x}_{k-1})$, $\boldsymbol{h}(\boldsymbol{x}_k)$	$\boldsymbol{z}_k$, $\hat{\boldsymbol{x}}_{k\|k-1}$, $\boldsymbol{\Sigma}_{k\|k-1}$, $\boldsymbol{\Sigma}_{k\|k}$

2.3 Proposed GM-IEKF

The proposed GM-IEKF method is developed in four main steps as follows. Firstly, the classical recursive approach is converted into a batch-mode regression form such that the observations and the predictions are processed simultaneously, resulting in an enhanced data redundancy. This redundancy allows our estimator to suppress both innovation and observation outliers while exhibiting good tracking capabilities of the state dynamics. Secondly, a robust prewhitening of the data is achieved by means of Projection Statistics (PS) [37,65,66] to uncorrelate the prediction and the observation errors even in presence of outliers. Thirdly, a generalized maximum likelihood criterion based on the Huber convex ρ-function is minimized via the iteratively reweighted least squares (IRLS) algorithm. Fourthly, the error covariance matrices of the predictions and the state estimates are updated using expressions derived from the total influence function of the GM-estimator.

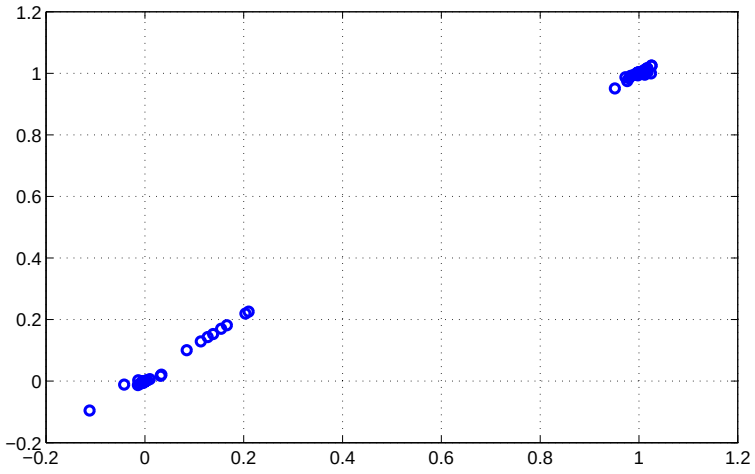


Figure 2.1: Scatter plot of the matrix $\boldsymbol{Z}$ without outliers

2.3.1 Derivation of the Batch-Mode Regression Form

Given the filtered state vector $\widehat{\boldsymbol{x}}_{k-1|k-1}$ at time step $k-1$, with covariance matrix $\boldsymbol{\Sigma}_{k-1|k-1}$, the state at the next time step k is predicted through (2.18) and (2.19). Define $\widehat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{x}_k - \boldsymbol{\eta}_k$, where $\boldsymbol{x}_k$ is the true state vector and $\boldsymbol{\eta}_k$ is the state estimate error with zero mean and covariance matrix $\mathbb{E}\left[\boldsymbol{\eta}_k \boldsymbol{\eta}_k^T\right] = \boldsymbol{\Sigma}_{k|k-1}$. By putting it with (2.2) in matrix form, we obtain

$$\begin{bmatrix} \widehat{\boldsymbol{x}}_{k|k-1} \\ \boldsymbol{z}_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_k \\ \boldsymbol{h}\left(\boldsymbol{x}_k\right) \end{bmatrix} + \begin{bmatrix} -\boldsymbol{\eta}_k \\ \boldsymbol{v}_k \end{bmatrix}, \tag{2.23}$$

which can be rewritten in the following compact form:

$$\widetilde{\boldsymbol{z}}_k = \widetilde{\boldsymbol{h}}(\boldsymbol{x}_k) + \widetilde{\boldsymbol{e}}_k, \tag{2.24}$$

with the error covariance matrix given by

$$\mathbb{E}\left[\widetilde{\boldsymbol{e}}_k \widetilde{\boldsymbol{e}}_k^T\right] = \begin{bmatrix} \boldsymbol{\Sigma}_{k|k-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R}_k \end{bmatrix} = \boldsymbol{S}_k \boldsymbol{S}_k^T. \tag{2.25}$$

Here $\boldsymbol{S}_k$ is calculated by the Cholesky decomposition technique.

2.3.2 Robust Prewhitening

If we directly apply the matrix $\boldsymbol{S}_k$ in the prewhitening step, the outliers will corrupt the results [37]. Instead, we first detect and downweight the outliers in the derived batch-mode regression form by means of a statistical test applied to the PS. The PS values are some

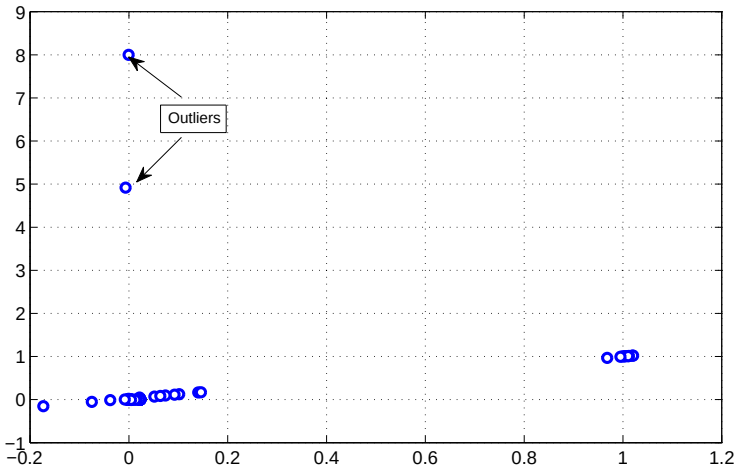


Figure 2.2: Scatter plot of the matrix $\boldsymbol{Z}$ with two outliers

kinds of robust distances of a collection of data points, $\boldsymbol{\ell}_i$, defined as realizations of a random vector. Outliers are data points that are distant from the bulk of the point cloud. To detect outliers, we propose to apply the PS to a 2-dimensional matrix $\boldsymbol{Z}$ that contains serially correlated samples of the innovations and of the predicted state vector. Formally, we have

$$\boldsymbol{Z} = \begin{bmatrix} \widehat{\boldsymbol{x}}_{k-1|k-2} & \widehat{\boldsymbol{x}}_{k|k-1} \\ \boldsymbol{z}_{k-1} - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k-1|k-2}) & \boldsymbol{z}_k - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k|k-1}) \end{bmatrix}, \tag{2.26}$$

where $\boldsymbol{z}_{k-1} - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k-1|k-2})$ and $\boldsymbol{z}_k - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k|k-1})$ are the innovation vectors while $\widehat{x}_{k-1|k-2}$ and $\widehat{x}_{k|k-1}$ are the predicted state vectors at time instants k-1 and k, respectively. We may also apply the PS to higher dimensional samples, but we found that 2 dimensions are enough to identify outliers. The PS values of the predictions and of the innovations are separately calculated because, as shown in Fig. 2.1, the values taken by the former are centered around one while those taken by the latter are centered around zero. Before calculating the PS, the real powers produced by the generators are normalized using their own rated MVA values.

The PS of the ith row vector, $\boldsymbol{\ell}_i$, of the predictions (respectively the innovations) in $\boldsymbol{Z}$ is defined as the maximum of the standardized projections of all the $\boldsymbol{\ell}_i$'s on every direction $\boldsymbol{l}$ that originates from the coordinatewise medians of the predictions (respectively the innovations) and that passes through every data point, and where the standardized projections are based on the sample median and the median-absolute-deviation [65]. Formally we have

$$PS_i = \max_{\|\boldsymbol{l}\|=1} \frac{\left|\boldsymbol{\ell}_i^T l - med_j\left(\boldsymbol{\ell}_j^T l\right)\right|}{1.4826\ med_k\left|\boldsymbol{\ell}_k^T l - med_j\left(\boldsymbol{\ell}_j^T l\right)\right|}. \tag{2.27}$$

Once the PS values are calculated, they are compared to a threshold to identify the outliers. Fig. 2.2 provides a typical example where the outliers stand far away from the bulk of

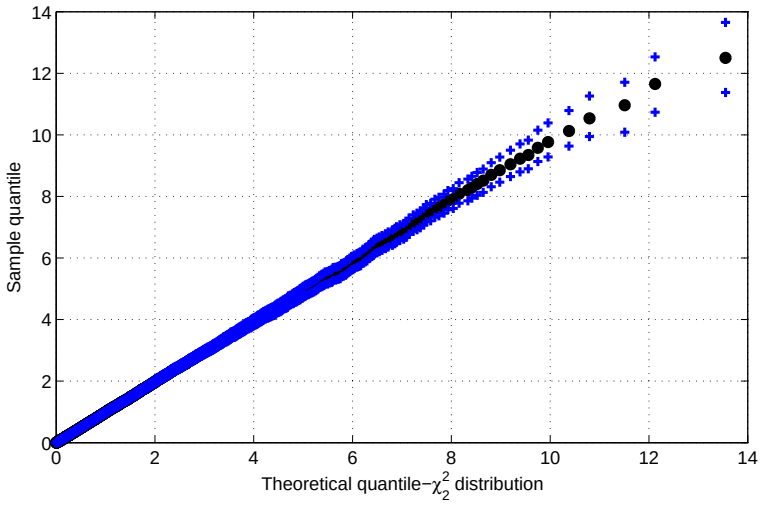


Figure 2.3: QQ plots of the sample quantiles of the PS vs. the corresponding quantiles of the χ^2_2 distribution

innovations. The threshold is determined from the probability distribution of the PS under the assumption that the data points follow a bivariate Gaussian probability distribution.

To this end, we conduct extensive Monte Carlo simulations and QQ-plots. Consider two random variables $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ that are independent and identically distributed according to $\mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$. Then generate 1000 realizations of these two random variables, apply the PS to $[\boldsymbol{v}_1 \;\; \boldsymbol{v}_2]$ and repeat the procedure 100 times. The sample medians and the interquantile ranges of the empirical PS quantiles are finally plotted versus the corresponding quantiles of the chi-square distribution with 2 degrees of freedom. The QQ plots displayed in Fig. 2.3 provide evidence that PS follows that distribution. We pick the threshold ζ of the statistical test equal to $\chi^2_{\nu,\beta} = \chi^2_{2,0.975}$ at a significance level of 97.5%. The detected outliers, whose values satisfy $PS_i > \zeta$, are downweighted via

$$\varpi_i = \min\left(1, d^2/PS_i^2\right). \tag{2.28}$$

Choosing d=1.5 yields good statistical efficiency at the Gaussian distribution without increasing too much the bias induced by outliers [37]. Finally, the prewhitening of $\widetilde{\boldsymbol{e}}_k$ in (7.32) is performed by pre-multiplying that nonlinear regression model by $\boldsymbol{S}_k^{-1}$, resulting in

$$\boldsymbol{S}_k^{-1}\widetilde{\boldsymbol{z}}_k = \boldsymbol{S}_k^{-1}\widetilde{\boldsymbol{h}}(\boldsymbol{x}_k) + \boldsymbol{S}_k^{-1}\widetilde{\boldsymbol{e}}_k, \tag{2.29}$$

which can be written in a compact form as

$$\boldsymbol{y}_k = \boldsymbol{\varphi}(\boldsymbol{x}_k) + \boldsymbol{\xi}_k. \tag{2.30}$$

2.3.3 Robust Filtering and Solution

To bound the influence of outliers on the state estimates, a GM-estimator is used. It minimizes an objective function given by

$$J(\boldsymbol{x}) = \sum_{i=1}^{m+n} \varpi_i^2 \rho(r_{S_i}), \tag{2.31}$$

where ϖ_i is calculated as in (5.20); $r_{S_i} = r_i / s\varpi_i$ is the standardized residual; $r_i = y_i - \varphi_i(\widehat{\boldsymbol{x}})$ is the residual; $s = 1.4826 \cdot b_m \cdot \text{median}_i |r_i|$ is the robust scale estimate; b_m is a correction factor for unbiasedness at the Gaussian distribution [65]; $\rho(\cdot)$ is the convex Huber ρ-function made up of a quadratic and a tangent function defined as

$$\rho(r_{S_i}) = \begin{cases} \frac{1}{2} r_{S_i}^2, & \text{for } |r_{S_i}| < \tau \\ \tau |r_{S_i}| - \tau^2/2, & elsewhere \end{cases}, \tag{2.32}$$

where $\tau = 1.5$ leads to high statistical efficiency under Gaussian noise [37]. To minimize (5.21), one takes its partial derivative and sets it equal to zero, yielding

$$\frac{\partial J(\boldsymbol{x})}{\partial \boldsymbol{x}} = \sum_{i=1}^{m+n} -\frac{\varpi_i \boldsymbol{c}_i}{s} \psi(r_{S_i}) = \boldsymbol{0}, \tag{2.33}$$

where $\psi(r_{S_i}) = \partial \rho(r_{S_i}) / \partial r_{S_i}$ and $\boldsymbol{c}_i$ is the i-th column vector of the matrix $\boldsymbol{C}^T$, where $\boldsymbol{C} = {}^{\partial \boldsymbol{\varphi}}/_{\partial \boldsymbol{x}} |_{\boldsymbol{x}=\hat{\boldsymbol{x}}}$. Then, by dividing and multiplying by r_{S_i} on both sides of (5.23), we get

$$\boldsymbol{C}^T \boldsymbol{Q} (\boldsymbol{y} - \boldsymbol{\varphi}(\boldsymbol{x})) = \boldsymbol{0}, \tag{2.34}$$

where $\boldsymbol{Q}$ =diag$(q(r_{S_i}))$ and $q(r_{S_i}) = \psi(r_{S_i}) / r_{S_i}$.

Taking a first-order Taylor series expansion of $\boldsymbol{\varphi}(\boldsymbol{x})$ about $\widehat{\boldsymbol{x}}_{k|k}$ and using the IRLS algorithm [67], the state vector correction at the j-th iteration is calculated by

$$\Delta \widehat{\boldsymbol{x}}_{k|k}^{(j+1)} = \left(\boldsymbol{C}^T \boldsymbol{Q}^{(j)} \boldsymbol{C}\right)^{-1} \boldsymbol{C}^T \boldsymbol{Q}^{(j)} \left(\boldsymbol{y} - \boldsymbol{\varphi}\left(\widehat{\boldsymbol{x}}_{k|k}^{j}\right)\right), \tag{2.35}$$

where $\Delta \widehat{\boldsymbol{x}}_{k|k}^{(j+1)} = \widehat{\boldsymbol{x}}_{k|k}^{(j+1)} - \widehat{\boldsymbol{x}}_{k|k}^{(j)}$ and $\boldsymbol{C}$ is evaluated at $\widehat{\boldsymbol{x}}_{k|k}^{(j)}$. The algorithm converges when $\left\| \Delta \widehat{\boldsymbol{x}}_{k|k}^{(j+1)} \right\|_\infty \leq 10^{-2}$.

2.3.4 Robust IEKF versus Robust EKF

In this section, we further justify why the robust IEKF is preferred to the robust EKF. When a non-iterative EKF is considered, the measurement function given by (2.2) is linearized using a first-order Taylor series expansion about the predicted state vector $\widehat{\boldsymbol{x}}_{k|k-1}$, yielding

$$\boldsymbol{z}_k = \boldsymbol{h}\left(\widehat{\boldsymbol{x}}_{k|k-1}\right) + \boldsymbol{H}_k \left(\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k-1}\right) + \boldsymbol{v}_k, \tag{2.36}$$

where $\boldsymbol{H}_k = {}^{\partial \boldsymbol{h}}/_{\partial \boldsymbol{x}}\big|_{\boldsymbol{x}=\widehat{\boldsymbol{x}}_{k|k-1}}$. Then the regression model in (7.31) changes to

$$\begin{bmatrix} \widehat{\boldsymbol{x}}_{k|k-1} \\ \boldsymbol{z}_k - \boldsymbol{h}\left(\widehat{\boldsymbol{x}}_{k|k-1}\right) + \boldsymbol{H}_k \widehat{\boldsymbol{x}}_{k|k-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{H}_k \end{bmatrix} \boldsymbol{x}_k + \begin{bmatrix} -\boldsymbol{\eta}_k \\ \boldsymbol{v}_k \end{bmatrix}. \tag{2.37}$$

The PS of the row vectors of the revised matrix $\boldsymbol{Z}$ are then calculated and a statistical test is applied to identify observation outliers. However, because innovation outliers occur in the predicted state $\widehat{\boldsymbol{x}}_{k|k-1}$, their effects will spread out on the observation vector on the left side of (2.37) via the matrix $\boldsymbol{H}_k$, making the PS to break down. By contrast, only the innovation $\widehat{\boldsymbol{x}}_{k|k-1}$ corresponding to the outliers will be affected in (7.31). Thus, the PS will not break down and both observation and innovation outliers are suppressed.

2.3.5 Estimation Error Covariance Matrix Updating

After the convergence of the iterative process, the estimation error covariance matrix $\boldsymbol{\Sigma}_{k|k}$ needs to be updated, thus allowing us to predict the state at the next time step. The updating of $\boldsymbol{\Sigma}_{k|k}$ in the GM-IEKF is derived as the covariance matrix of the total influence function, $\boldsymbol{IF}$ [37, 66, 67], yielding

$$\begin{aligned} \boldsymbol{\Sigma}_{k|k} &= \mathbb{E}\left[\boldsymbol{IF} \cdot \boldsymbol{IF}^{\boldsymbol{T}}\right] \\ &= \frac{\mathbb{E}_\Phi\left[\psi^2\left(r_{S_i}\right)\right]}{\left\{\mathbb{E}_\Phi\left[\psi'\left(r_{S_i}\right)\right]\right\}^2} \left(\boldsymbol{C}_k^T \boldsymbol{C}_k\right)^{-1} \left(\boldsymbol{C}_k^T \boldsymbol{Q}_{\varpi} \boldsymbol{C}_k\right) \left(\boldsymbol{C}_k^T \boldsymbol{C}_k\right)^{-1}. \end{aligned} \tag{2.38}$$

This expression is derived as follows:

It has been shown by Hample [67] that the $\boldsymbol{IF}$ can be utilized to measure the sensitivity of an estimator to an infinitesimal contamination and to derive the error covariance matrix of an estimator under the assumed probability distribution, e.g., the Gaussian distribution. This approach is applied here for the GM-IEKF. Consider the ϵ-contamination model $G = (1-\epsilon)\,\Phi + \epsilon \Delta_r$, where Φ is the target distribution and Δ_r is the probability mass at r, and let the cumulative probability distribution of the residual vector $\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{\varphi}(\boldsymbol{x})$ calculated from (5.18) be $\Phi(\boldsymbol{r})$. The GM-estimator provides an estimate of the state by processing the redundant observation vector $\boldsymbol{y}$ and solving the following implicit equation:

$$\sum_{i=1}^{m+n} \boldsymbol{\lambda}_i\left(\boldsymbol{r}, \boldsymbol{x}\right) = \sum_{i=1}^{m+n} \varpi_i \frac{\partial \boldsymbol{\varphi}\left(\boldsymbol{x}\right)}{\partial \boldsymbol{x}} \psi\left(r_{S_i}\right) = \boldsymbol{0}, \tag{2.39}$$

which, by virtue of the Glivenko-Cantelli theorem [72], asymptotically tends to

$$\int \boldsymbol{\lambda}\left(\boldsymbol{r}, \boldsymbol{T}\right) dG = \boldsymbol{0}. \tag{2.40}$$

Here, the GM-estimator $\widehat{\boldsymbol{x}}_{k|k}$ at G has been put in a functional form, $\boldsymbol{T}(G)$. The asymptotic total influence function of $\boldsymbol{T}(G)$ is given by

$$\boldsymbol{IF}\left(\boldsymbol{r},\Phi\right)=\frac{\partial\boldsymbol{T}(G)}{\partial\epsilon}\bigg|_{\epsilon=0}=\lim_{\Delta\epsilon\to0}\frac{\boldsymbol{T}((1-\epsilon)\Phi+\epsilon\Delta r)-\boldsymbol{T}(\Phi)}{\epsilon}. \tag{2.41}$$

Substituting G into (2.40) yields

$$\int\boldsymbol{\lambda}(\boldsymbol{r},\boldsymbol{T}(G))d\Phi+\epsilon\int\boldsymbol{\lambda}(\boldsymbol{r},\boldsymbol{T}(G))d\left(\Delta_r-\Phi\right)=\boldsymbol{0}. \tag{2.42}$$

Taking the differentiation with respect to ϵ and evaluating it at $\epsilon=0$, assuming regularity conditions and Fisher consistency at Φ, given by $\int\boldsymbol{\lambda}(\boldsymbol{r},\boldsymbol{T}(\Phi))d\Phi=\boldsymbol{0}$, yields

$$\frac{\partial}{\partial\epsilon}\int\boldsymbol{\lambda}(\boldsymbol{r},\boldsymbol{T}(G))d\Phi\,\big|_{\epsilon=0}+\int\boldsymbol{\lambda}(\boldsymbol{r},\boldsymbol{T}(G))d(\Delta_r)|_{\epsilon=0}=\boldsymbol{0}. \tag{2.43}$$

After applying the sifting property of the Dirac impulse to the second term, we obtain

$$\frac{\partial}{\partial\epsilon}\int\boldsymbol{\lambda}(\boldsymbol{r},\boldsymbol{T}(G))d\Phi\,\big|_{\epsilon=0}+\boldsymbol{\lambda}(\boldsymbol{r},\boldsymbol{T}(\Phi))=\boldsymbol{0}. \tag{2.44}$$

Assuming $\boldsymbol{\lambda}(\cdot)$ is continuous and measurable and $\boldsymbol{\lambda}'(\cdot)$ is measurable, we can apply the interchangeability of differentiation and integration theorem to the first term in (2.44), yielding

$$\int\frac{\partial\boldsymbol{\lambda}(\boldsymbol{r},\boldsymbol{T}(G))}{\partial\epsilon}|_{\boldsymbol{T}(\Phi)}.\frac{\partial\boldsymbol{T}(G)}{\partial\epsilon}|_{\epsilon=0}d\Phi+\boldsymbol{\lambda}(\boldsymbol{r},\boldsymbol{T}(\Phi))=\boldsymbol{0}. \tag{2.45}$$

Thus, $\boldsymbol{IF}(\boldsymbol{r},\Phi)$ is expressed as

$$\begin{aligned}\boldsymbol{IF}(\boldsymbol{r},\Phi)&=\tfrac{\partial\boldsymbol{T}(G)}{\partial\epsilon}|_{\epsilon=0}\\&=-[\textstyle\int\frac{\partial\boldsymbol{\lambda}(\boldsymbol{r},\boldsymbol{T}(G))}{\partial\epsilon}|_{T(\Phi)}d\Phi]^{-1}\boldsymbol{\lambda}(\boldsymbol{r},\boldsymbol{T}(\Phi)).\end{aligned} \tag{2.46}$$

Taking the derivative of $\boldsymbol{\lambda}(\cdot)$ with respect to $\boldsymbol{x}$ and assuming that ϖ and s are independent of $\boldsymbol{x}$, we obtain

$$\frac{\partial\boldsymbol{\lambda}(\boldsymbol{r})}{\partial\boldsymbol{x}}=\varpi\psi(r_{S_i})[\frac{\partial^2\boldsymbol{\varphi}(\boldsymbol{x})}{\partial x_i\partial x_j}]+\varpi[\frac{\partial\psi(r_{S_i})}{\partial\boldsymbol{x}}][\frac{\partial\boldsymbol{\varphi}(\boldsymbol{x})}{\partial\boldsymbol{x}}]^T. \tag{2.47}$$

By neglecting the second-order term on the right-hand side of (2.47), we get

$$\frac{\partial\boldsymbol{\lambda}(\boldsymbol{r})}{\partial\boldsymbol{x}}=\varpi[\frac{\partial\psi(r_{S_i})}{\partial\boldsymbol{x}}][\frac{\partial\boldsymbol{\varphi}(\boldsymbol{x})}{\partial\boldsymbol{x}}]^T. \tag{2.48}$$

Applying the chain rule to the derivative of $\psi(r_{S_i})$ yields

$$\frac{\partial\boldsymbol{\lambda}(\boldsymbol{r})}{\partial\boldsymbol{x}}=-\frac{1}{s}\psi'\left(r_{S_i}\right)\left[\frac{\partial\boldsymbol{\varphi}(\boldsymbol{x})}{\partial\boldsymbol{x}}\right]\left[\frac{\partial\boldsymbol{\varphi}(\boldsymbol{x})}{\partial\boldsymbol{x}}\right]^T \tag{2.49}$$

By substituting (5.18), (2.39) and (2.49) into (2.46), we get

$$\boldsymbol{IF}\left(\boldsymbol{r},\Phi\right)=\left[\int\frac{1}{s}\psi'\left(r_{S_i}\right)\boldsymbol{CC^T}\Big|_{T(\Phi)}d\Phi\right]^{-1}\varpi\boldsymbol{C}\psi\left(r_{S_i}\right). \tag{2.50}$$

Thus, the state estimation error covariance matrix for the GM-IEKF at time step k is given by

$$\begin{aligned}\boldsymbol{\Sigma}_{k|k}&=\mathbb{E}\left[\boldsymbol{IF}\cdot\boldsymbol{IF^T}\right]\\&=\frac{\mathbb{E}_\Phi\left[\psi^2\left(r_{S_i}\right)\right]}{\left\{\mathbb{E}_\Phi\left[\psi'\left(r_{S_i}\right)\right]\right\}^2}\left(\boldsymbol{C}_k^T\boldsymbol{C}_k\right)^{-1}\left(\boldsymbol{C}_k^T\boldsymbol{Q}_\varpi\boldsymbol{C}_k\right)\left(\boldsymbol{C}_k^T\boldsymbol{C}_k\right)^{-1},\end{aligned} \tag{2.51}$$

where $\boldsymbol{Q}_\varpi=diag\left(\varpi_i^2\right)$ and $\boldsymbol{C}_k={}^{\partial\boldsymbol{\varphi}}/_{\partial\boldsymbol{x}}\Big|_{\boldsymbol{x}=\widehat{\boldsymbol{x}}_{k|k}}$.

Comment 1: In most related literature, the state estimation error covariance matrix of an M-estimator is updated using the same equations as for the standard least-squares-based EKF, e.g., [68, 69]. Obviously, this will lead to a degraded performance of the estimator, including poor tracking capabilities and lower statistical efficiency. The correct asymptotic covariance matrix of that estimator should be utilized, which may be derived from (2.38).

Comment 2: In the proposed GM-IEKF, the matrix $\boldsymbol{Q}$ and the weights $\boldsymbol{\varpi}$ are used to bound the influence of residual and of position, guaranteeing a robust state estimation, while the matrix $\boldsymbol{Q}_\varpi$ contributes to a robust estimation of the state estimation error covariance matrix and to a robust prewhitening in the next time step.

2.4 Numerical Results

The performance of the proposed estimator is tested on the IEEE 39-bus test system, whose data can be found in [60]. The standard EKF and UKF are implemented for comparisons. In the simulations, the real and reactive power injections and the bus voltage phasors are assumed to be metered using PMUs; a random Gaussian noise with zero mean and standard deviation 10^{-2} is assumed for either system or measurement noise; the diagonal elements of the initial error covariance matrix of the UKF are set to 10^{-4}; the initial values of the state vector are arbitrarily chosen for all three estimators. The damping ratio D is set equal to 0.005 for all generators and the PMU measurements are assumed to be received at a rate of 48 samples per second. At t=0.5s Line 15-16 is switched off, which causes a large disturbance to the system. The time domain simulation results are assumed to be the true state values; the parameters for GM-IEKF are set to τ=d=1.5, and the maximal number of iterations equals to 20.

2.4.1 Case 1: Small Process and Measurement Noise

We initialize each process noise with a very small variance of magnitude 10^{-4}. The error variance of the PMU measurements is set to 1%. Fig. 2.4 shows the tracking results of the

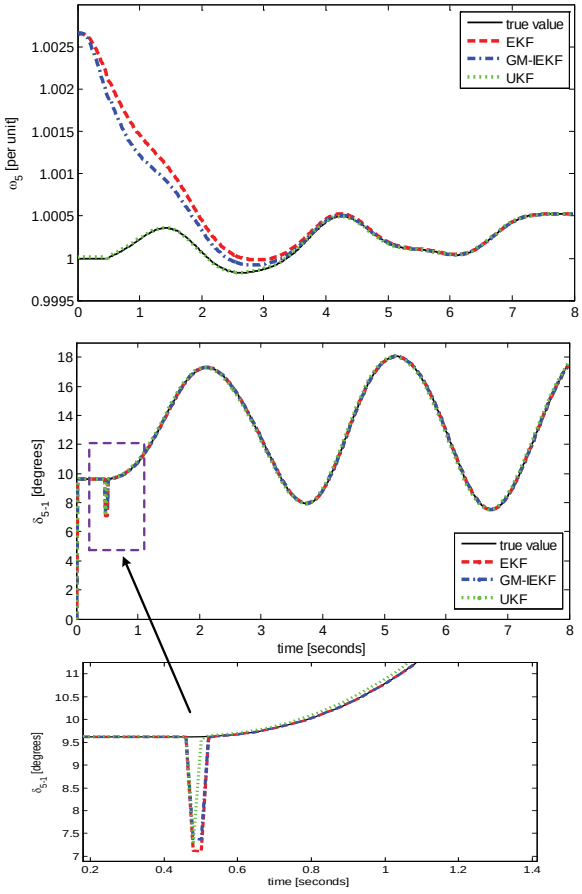


Figure 2.4: Estimated ω_5 and δ_{5-1} by the three methods in Case 1: small process and measurement noise.

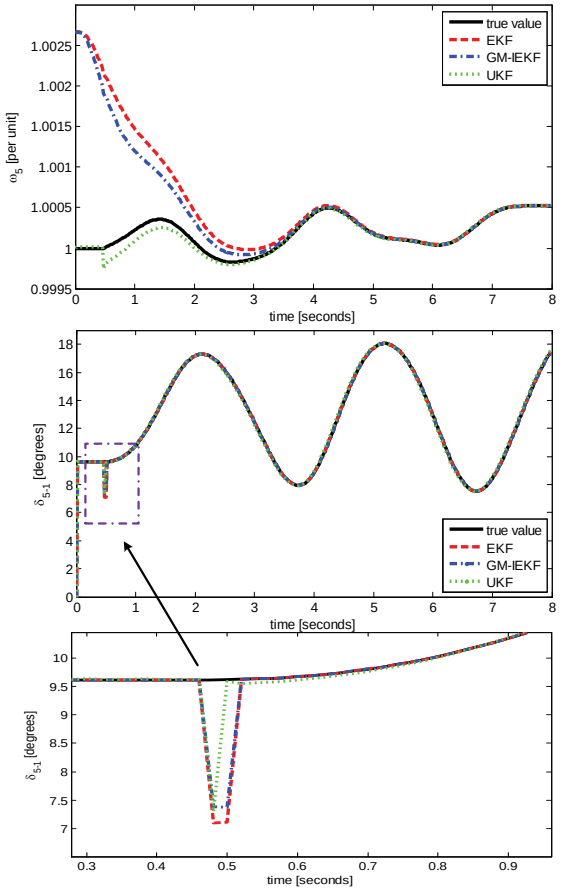


Figure 2.5: Estimated ω_5 and δ_{5-1} by the three methods in Case 2, where the variance of the process noise is increased from 10^{-4} to 10^{-2}.

three dynamic state estimators. We randomly pick Generator 5 as an example and omit the remaining plots since they do not bring any additional qualitative information. It can be seen from the figure that all the three methods accurately track the system dynamic states. The GM-IEKF outperforms the standard EKF in estimating both the rotor angle and the rotor speed. The UKF is the fastest method for tracking these state variables because it uses the *sigma points* to propagate the mean and the covariance matrix without relying on direct linearization, which improves the computational efficiency.

2.4.2 Case 2: Small Measurement Noise and Large Process Noise

In order to investigate the impacts of system process noise on the tracking performance, the variance of the process noise is increased from 10^{-4} to 10^{-2} while the measurement variance is kept the same as that in Case 1. Fig. 2.5 presents the tracking results. It can be seen from this figure that UKF is greatly affected by system process noise regarding the rotor speed estimation, while GM-IEKF and EKF are slightly affected. In this case, GM-IEKF is able to bound the influence of the process noise, resulting in the best tracking performance among the three methods.

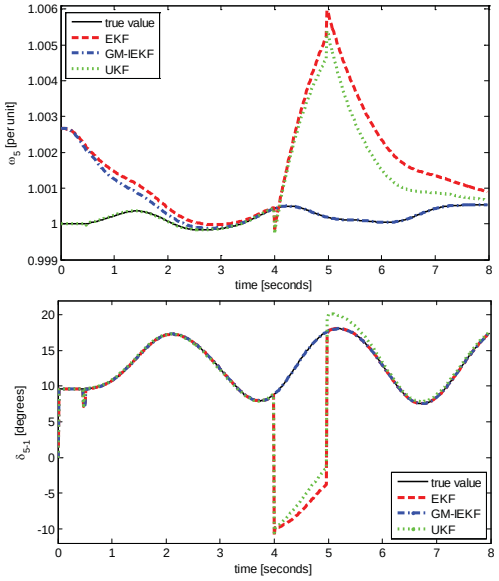


Figure 2.6: Estimated ω_5 and δ_{5-1} by the three methods in Case 3, where the communication link with the PMU placed at the Generator 5 is lost from t=4s to t=5s.

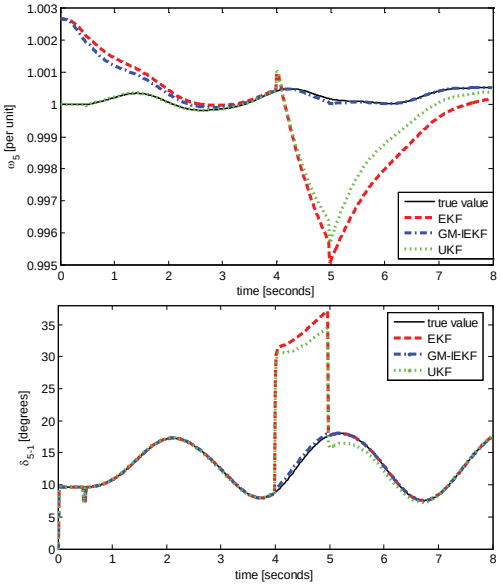


Figure 2.7: Estimated ω_5 and δ_{5-1} by the three methods in Case 4, where the real power measurement P_5 on Bus 34 is contaminated with gross errors from t=4s to t=5s.

2.4.3 Case 3: Momentary Loss of Communication Link

Some PMU devices may temporarily lose their communication links to the PDC due to device failures, cyber attacks, communication interruptions, to cite a few. In that case, the communication link with the PMU placed at Bus 34, where Generator 5 is connected, is assumed to be lost from t=4s to t=5s. Therefore, the measurement set $\{P_5, Q_5, V_5, \theta_5\}$ becomes unavailable during this time interval and their values are set equal to zero for simulation purpose. It is observed that UKF and EKF are not capable of tracking the trajectories of ω_5 and δ_{5-1}. By contrast, GM-IEKF exhibits good tracking capabilities by relying on the predicted state estimates.

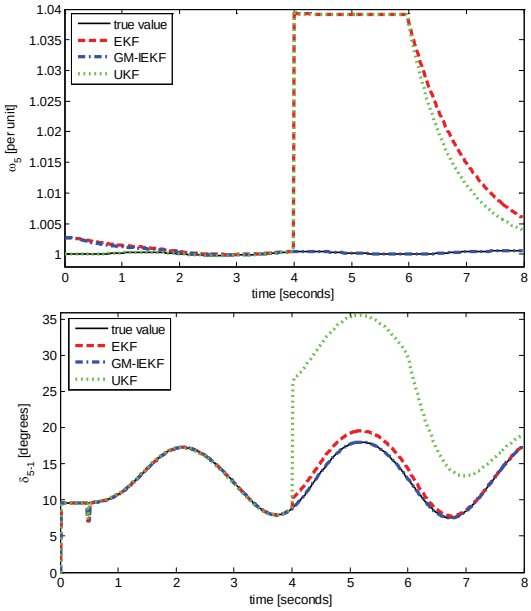


Figure 2.8: Estimated ω_5 and δ_{5-1} by the three methods in Case 5, where the predicted value of ω_5 is changed from 0.2072 to 2 pu from t=4s to t=6s.

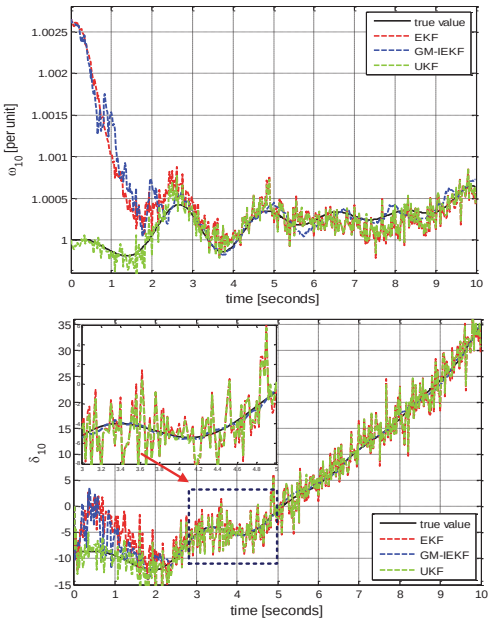


Figure 2.9: Estimated ω_{10} and δ_{10} by the three methods under non-Gaussian measurement noise.

2.4.4 Case 4: Occurrence of Observation Outliers

To investigate the effects of observation outliers on the tracking performances of the EKF, UKF, and GM-IEKF, P_5 on Bus 34 is contaminated with gross errors from t=4s to t=5s by changing its value to 8 pu, simulating a faulty synchronization or an impulsive communication noise. Note that by applying the PS to the Jacobian matrix $\boldsymbol{C}$, this outlier is flagged as bad leverage point. The test results are shown in Fig. 2.7. We note that when observation outliers occur, both EKF and UKF deviate far away from the true system states due to their

vulnerabilities to outliers. However, GM-IEKF significantly reduces its influence, which is reflected by a very small weight $\varpi_i \approx 10^{-4}$, keeping a good tracking performance.

2.4.5 Case 5: Occurrence of Innovation Outliers

As discussed before, the predictions may be unreliable due to the imperfect dynamical model or impulsive system process noise. To investigate the performances of the estimators under this condition, the predicted value of ω_5 is changed to 2 pu between t=4s and t=6s. Fig. 2.8 shows the simulation results. We observe that the UKF is greatly affected by innovation outliers, which makes its tracking trajectories unreliable. The EKF is less sensitive to the rotor angle estimation compared with the UKF, but its estimation for rotor speed is not acceptable. By contrast, the GM-IEKF effectively suppresses these innovation outliers, resulting in good states tracking performance.

2.4.6 Breakdown Point of the Proposed GM-IEKF

We demonstrate in the previous examples that the GM-IEKF can cope with observation and innovation outliers. The remaining question is how many outliers it can handle without giving unreliable estimation results, that is, what is its breakdown point? This concept provides a measure of the global robustness of an estimator and is formally given by

$$\varepsilon^* = \max \left\{ \varepsilon = \frac{N_f}{N_m}; b_{\max} \text{ finite} \right\}, \tag{2.52}$$

where N_f is the number of outliers; N_m is the total number of data points, which is the dimension of $\boldsymbol{y}_k$; b_{max} is the maximum possible bias induced by a given fraction of outliers. A definition of the finite-sample breakdown point of a robust estimator in nonlinear regression is given by Stromberg and Ruppert [48]. The expression of the maximum breakdown that any regression equivariant estimator may have in linear structured regression (e.g., in linearized power system state estimation model, which involves sparse Jacobian matrices) is derived by Mili and Coakley [47].

We carry out extensive simulations to determine the breakdown point of our GM-IEKF when applied to the IEEE 39-bus test system. By replacing an increasing number of data points by outliers in the vector $\boldsymbol{y}_k$ of dimension $m_t = m + n$, it is observed that the GM-IEKF can handle at least 25% of outliers among the data set, be they innovation or observation outliers, the worst case being clustered ones. Note that the exact value of the breakdown point of the GM-IEKF in power systems still needs to be determined; it will be investigated in a future work.

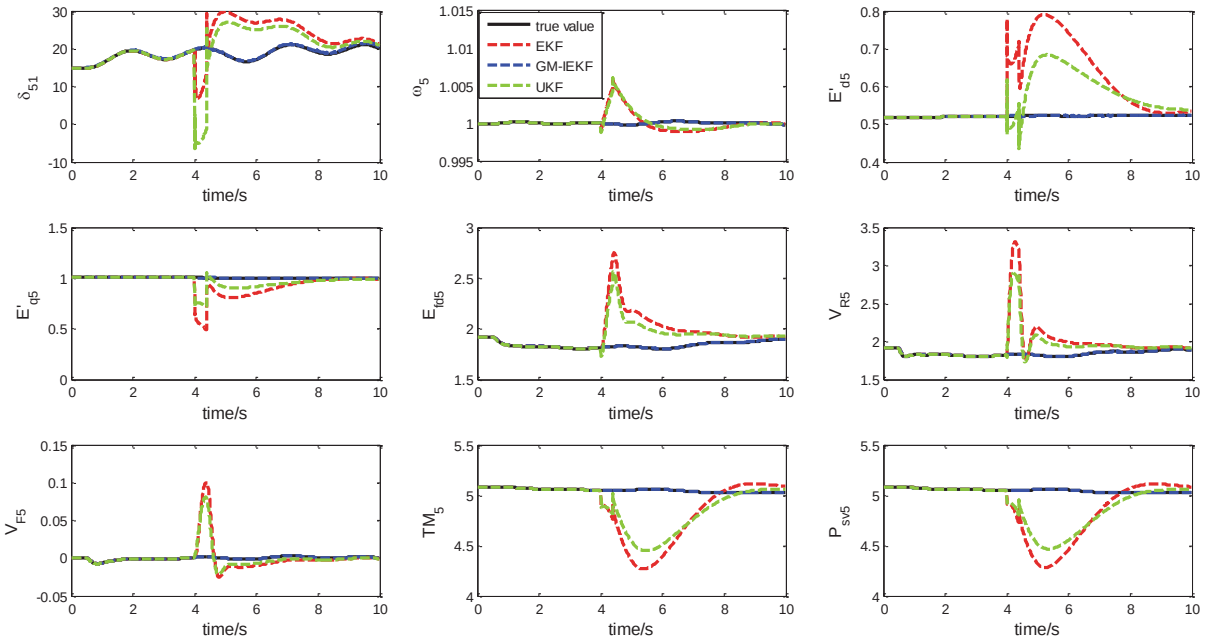


Figure 2.10: State tracking performance of the three methods in the presence of two observation outliers induced as follows: two pairs of measurements, P_5 and Q_5, on Bus 34 are contaminated with gross errors from t=4s to t=5s.

2.4.7 Robustness to Non-Gaussian Noise

To demonstrate the robustness of the proposed method under thick-tailed distributions, non-Gaussian PMU noise is simulated using the models proposed in a recent PNNL report [35,36]. Specifically, the bimodal Gaussian mixture model with zero mean, a variance of 10^{-4} and weights of 0.9 and 0.1 is assumed for the voltage magnitude measurement noise, whereas the Laplacian distribution with zero mean and a scale of 1 is assumed for the real and reactive power measurement noise. As an illustrative example, results for Generator 10 are displayed in Fig. 6.6. We observe that both EKF and UKF state estimates exhibit large oscillations while our robust GM-IEKF is closely following the true states.

2.4.8 Results Using the Detailed Model

Now, let us describe and analyze the test results obtained for all the three methods in presence of observation or innovation outliers using the two-axis machine model. In this test, the measurement settings and fault condition are the same as those given in the previous simulations. The time step for time domain simulation is 0.008s. The estimated states of Generator 5 are provided as examples. All the parameter values of the generators are taken from [70]. Figs. 2.10 and 2.11 present the results of the three methods with two observation outliers, i.e., P_5 and Q_5 changed to 0 pu, and one innovation outlier, i.e., δ_5 changed to 5 pu.

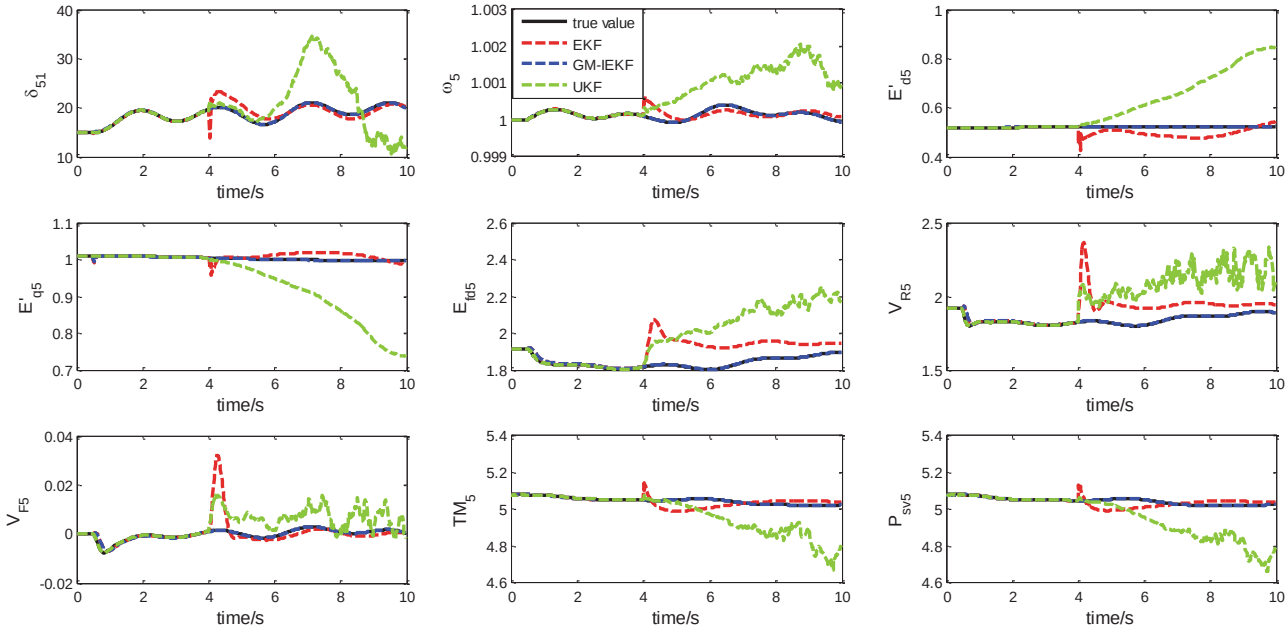


Figure 2.11: State tracking performance of the three methods in the presence of an innovation outlier induced as follows: the predicted rotor angle of Generator 5 is contaminated with gross errors from t=4s to t=5s.

From these two figures, we observe that even with high-order generator model, our GM-IEKF is able to effectively suppress both observation and innovation outliers. By contrast, EKF and UKF exhibit no robustness against any types of outliers, yielding unreliable estimation results.

2.4.9 GM-IEKF Computational Efficiency

To check whether the proposed GM-IEKF can keep up with PMUs at their scanned rate of 30 or 60 samples per second, its computational efficiency is analyzed and compared to that of the EKF and UKF under three different scenarios defined as follows: in Scenario 1, no outliers

Table 2.2: Average Computing Times of the Three Estimation Methods in Three Scenarios For Each Time Sample

Cases	EKF	UKF	GM-IEKF
Scenario 1	3.94ms	3.48ms	5.96ms
Scenario 2	3.95ms	3.49ms	5.99ms
Scenario 3	3.93ms	3.51ms or diverged	6.30ms

occur; in Scenario 2, observation outliers occur; in Scenario 3, innovation outliers occur. All the three scenarios are tested with the two-axis model of the synchronous generators. The testing environment consists of a PC configured with Intel Xeon(R), CPU E5-1650, 3.5GHz and 32GB RAM. The average computing times of all the three methods in the three scenarios for every time sample are displayed in Table 2.2. From this table, we observe that EKF and UKF have similar computing times that are much lower than the PMU sampling periods of 33.3ms and 16.7ms for 30 sample/s and 60 samples/s, respectively. As for GM-IEKF, although its execution time is larger than that of EKF, it is also lower than the PMU sampling period, enabling it to track system state dynamics in real-time. Furthermore, we notice that unlike EKF and GM-IEKF, UKF may exhibit numerical instability in presence of innovation outliers.

2.4.10 Centralized vs. Decentralized GM-IEKF

The proposed GM-IEKF is a centralized DSE, but it can be extended to a decentralized framework, typically employed for the real-time monitoring and control of very large-scale power systems. For instance, we may adopt the same approach as the one proposed in [26] by treating the generator terminal voltage and current phasors as system input and output variables, which will be both assumed to be metered by PMUs. By doing so, the dynamic equations for one generation unit can be decoupled from those of the other units. Therefore, GM-IEKF can be applied to estimate the system dynamic state by processing only local PMU measurements. However, with the decentralized GM-IEKF, the measurements processed by a DSE are only taken on machine terminal buses, yielding a very low local measurement redundancy. If the terminal measurements are lost due to cyber attacks or communication failures, or include either observation or innovation outliers, the decentralized GM-IEKF may lose its tracking ability. On the other hand, with a centralized GM-IEKF, all available measurements are transferred to the control center with some communication bandwidth utilization, resulting in an enhanced measurement redundancy. In this case, outliers and the loss of measurements can be effectively handled by the centralized GM-IEKF, thereby providing the control center with a complete set of state estimates, which enables it to take regional and global control actions.

Let us now compare the centralized and decentralized scheme from the viewpoint of the communication bandwidth requirement. Obviously, with a decentralized GM-IEKF, only local controls are implemented. If coordinated control is deployed between different local DSEs, additional communication bandwidth is required, making the comparison with the communication cost of the centralized GM-IEKF difficult to assess. In summary, choosing between the two schemes depends on the applications and the communication infrastructures being used.

2.5 Conclusion and Future Work

In this chapter, a robust GM-IEKF method is proposed for tracking the dynamic states of a power system. The estimation problem is formulated as a nonlinear regression problem by processing both state predictions and observations in a batch mode. Furthermore, projection statistics and a GM-estimator are applied to PMU metered values for bounding the influence of observation and innovation outliers, including the loss of measurement communication links. Simulations carried out on the IEEE 39-bus test system provide results that demonstrate the robustness and the statistical efficiency of GM-IEKF.

Despite evident advantages of our GM-IEKF, which include fast computations and good robustness to observation and innovation outliers, it suffers from several weaknesses that call for further research and developments. One weakness is its vulnerability to system parameter and topology errors, which require a very high level of local measurement redundancy [71]. The second weakness is that the iterative EKF may produce unreliable state estimates under strong nonlinearities of the power system model. Dynamic estimators that circumvent this difficulty include the unscented transformation based filters, the ensemble Kalman filter (EnKF), and the particle filter (PF), among others. These estimators make use of sigma/ensemble points or particles to approximate the true distribution of random state variables, resulting in good state predictions and filtering when the assumptions are satisfied. As a future work, we plan to develop robust versions of these filters for handling outliers under strong system nonlinearities. Finally, the proposed GM-IEKF will be extended to a decentralized framework for the real-time monitoring and control of a very large-scale power system.

Chapter 3

A Robust Generalized-Maximum Likelihood Unscented Kalman Filter: Theory and Its Application for Power System Dynamic State Estimation

3.1 Introduction

3.1.1 Motivation

The widespread deployment of Phasor Measurement Units (PMUs) on power transmission grids has made possible the real-time monitoring and control of power system dynamics. However, these functions are difficult to be achieved without the development of a fast and robust Dynamic State Estimator (DSE). Indeed, the dynamic state estimates of the synchronous machines can be utilized by various devices to enhance small signal stability and to initiate generation outages and load shedding during transient instabilities, such as power system stabilizers, automatic voltage regulators, and under-frequency relays [9, 73].

To date, a variety of dynamic state estimators have been proposed in the literature; they are based on the Extended Kalman Filter (EKF) [21, 22], the Iterated EKF (IEKF) [24, 38], the unscented Kalman filter (UKF) [25–27], to cite a few. However, all these methods suffer from several important shortcomings, precluding them from being adopted by power utilities for power system real-time applications. To be specific, they cannot handle *i) non-Gaussian process and observation noise of the system nonlinear dynamic models, ii) innovation, observation and structural outliers and iii) cyber attacks.*

There are several reasons for these shortcomings. Firstly, the current DSE approaches assume that both the process and the observation noise of the system nonlinear dynamic models are

Gaussian. However, two recent investigations conducted by PNNL [35, 36] revealed that the PMU measurement errors of the voltage and current magnitudes obey non-Gaussian probability distributions. Evidently, the presence of non-Gaussian noise calls for new research and development in robust power system DSE based on robust statistics. Secondly, three types of outliers associated with a given dynamical system model have been defined by Gandhi and Mili [37], namely observation outliers, which affect the metered values; innovation outliers, which corrupt the predicted state estimates; and structural outliers, which affect the system dynamic states and the observation functions. Observation outliers may result from large biases in PMU measurements due to infrequent calibration, or instrument failures, or impulsive communication noise [39, 40]. As for innovation outliers, they may occur in several different ways. For example, some of the generator models may not be well calibrated, resulting in highly inaccurate model outputs that are inconsistent with the measurements. This was precisely the case in the 1996 blackout, where the model being used predicted system stability while in reality the system was undergoing numerous cascading failures, which resulted in a rapid system collapse that occurred within minutes [42, 43]. Innovation outliers may also be induced by the approximations in the state prediction model or by a system process impulsive noise. By contrast, structural outliers are induced by wrong circuit breaker statuses or gross errors in the model parameters of the transmission lines, or of the automatic voltage regulators, or of the synchronous machines. In [44], it is reported that wrong estimates of the parameters of the synchronous machine models may result from the use of erroneous metered values. It turns out that the conventional filters, namely the EKF, the EnKF, the UKF, and the Particle Filter (PF) are not robust to any type of outliers. For instance, it is demonstrated in [45] that the performances of the EnKF, EKF, UKF and PF are significantly degraded in the presence of observation outliers. In addition, since the EKF, UKF and EnKF are derived based on Gaussian assumption [74], they may yield significantly biased estimates when non-Gaussian noise occurs. To address observation outliers, Rouhani and Abur [27] developed a robust UKF-based DSE using the Least-Absolute-Value (LAV) estimator. However, the authors do not address the vulnerability of the DSE to innovation outliers. In [38], a robust IEKF was proposed to handle observation and innovation outliers, but it may suffer from divergence problems if the nonlinearity of the system model is strong. In addition, both [27,38] do not address the non-Gaussianity of the measurement and system process noise. In signal processing and communication areas, Huber estimator-based robust UKF methods have been advocated to suppress observation outliers and bound the influence of non-Gaussian observation noise [75–78]. However, these methods are unable to handle innovations outliers, which can occur quite often in power systems due to controller parameter errors and model deficiency. Furthermore, they achieve very low statistical efficiency or may diverge when both process and observation noises are non-Gaussian.

3.1.2 Contributions and Paper Organization

To address the aforementioned challenges, this chapter resorts to robust statistics and develops a robust Generalized Maximum-Likelihood-type UKF (GM-UKF) for power system

dynamic state estimation [79]. Note that in [80], we have shown some preliminary results of our robust UKF, for instance that it is able to handle unknown Gaussian and non-Gaussian noise, and observation outliers. However, no theoretical proofs are provided to quantify its statistical efficiency and its breakdown point. This paper bridges these gaps by providing the corresponding theoretical justifications. Specifically, it includes the following new contributions:

- The statistical linearization approach is shown to yield the same solution as the sigma-points-based unscented transformation; this allows us to derive a redundancy linear batch-mode regression form by processing the predictions and observations simultaneously. Note that this overdetermined system of equations results in the data redundancy needed for the detection and suppression of the innovation and observation outliers.

- The application of the weighted least squares (WLS) estimator to the linear batch-mode regression form is proved to yield the same solution as the traditional UKF, making this filter vulnerable to non-Gaussian process and measurement noise and any type of outliers.

- A theorem is provided, which shows that the state estimates provided by our GM-UKF are asymptotically Gaussian even when the system process and measurement noise obey thick-tailed distributions; this is precisely the case when using PMU measurements. As a result, the means and the covariance matrices of the state estimates can be represented well by sigma points. Furthermore, this theorem guarantees a good performance of our GM-UKF for any distributions.

- Our GM-UKF is shown to be able to handle very strong system nonlinearity while other alternatives fail to converge or yield large biased state estimates.

- The robustness our GM-UKF to outliers and cyber attacks is investigated; it is shown that it can handle *at least* 25% of incorrect PMU measurements and predicted state variables.

- It is demonstrated both theoretically and numerically that our GM-UKF is able to achieve high statistical efficiency for a wide-range of system process and observation noise probability distributions.

The rest of the paper is organized as follows. Section 3.2 presents the problem formulation. Section 3.3 develops the theory of the proposed GM-UKF and Section 3.4 presents the application of GM-UKF to power system. Section 3.5 shows and analyzes the simulation results. Finally Section 3.6 concludes the paper.

3.2 Problem Formulation

3.2.1 Nonlinear Discrete-Time Dynamical System Model

A discrete-time state space representation of a general nonlinear dynamical system is expressed as

$$\boldsymbol{x}_k = \boldsymbol{f}\left(\boldsymbol{x}_{k-1}, \boldsymbol{u}_k\right) + \boldsymbol{w}_k, \tag{3.1}$$

$$\boldsymbol{z}_k = \boldsymbol{h}\left(\boldsymbol{x}_k, \boldsymbol{u}_k\right) + \boldsymbol{v}_k, \tag{3.2}$$

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

3.2.2 Dynamic State Estimation using UKF

The main idea underlying the UKF is the application of a deterministic sampling technique known as the unscented transformation, which allows us, under the Gaussian noise assumption, to choose a set of sample points, termed sigma points, that have the same mean and covariance matrix as those of the a priori state vector [52]. These sigma points are then propagated through the non-linear functions $\boldsymbol{f}$ and $\boldsymbol{h}$, yielding an estimation of the a posteriori state statistics by using the Kalman filter approach, i.e., the sample mean and the sample covariance matrix. Consequently, no calculation of Jacobian matrices is required, which can be by itself a difficult task to achieve in some cases or computationally costly.

To be specific, given a state estimate at time step k-1, $\widehat{\boldsymbol{x}}_{k-1|k-1} \in \mathbb{R}^{n\times 1}$, having a covariance matrix given by $\boldsymbol{P}^{xx}_{k-1|k-1}$, its statistics are captured by $2n$ weighted sigma points defined as

$$\boldsymbol{\chi}^{i}_{k-1|k-1} = \widehat{\boldsymbol{x}}_{k-1|k-1} \pm \left(\sqrt{n\boldsymbol{P}^{xx}_{k-1|k-1}}\right)_i, \tag{3.3}$$

with weights $w_i = \frac{1}{2n}, i = 1, ..., 2n$. Then, each sigma point is propagated through the nonlinear system process model (7.1), yielding a set of transformed samples expressed as

$$\boldsymbol{\chi}^{i}_{k|k-1} = \boldsymbol{f}\left(\boldsymbol{\chi}^{i}_{k-1|k-1}\right). \tag{3.4}$$

Next, the predicted sample mean and sample covariance matrix of the state vector are calculated by

$$\widehat{\boldsymbol{x}}_{k|k-1} = \sum_{i=1}^{2n} w_i \boldsymbol{\chi}^{i}_{k|k-1}, \tag{3.5}$$

$$\boldsymbol{P}^{xx}_{k|k-1} = \sum_{i=1}^{2n} w_i(\boldsymbol{\chi}^i_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})(\boldsymbol{\chi}^i_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})^T + \boldsymbol{Q}_k. \tag{3.6}$$

Finally, the measurement updating is performed and the filtered state $\widehat{\boldsymbol{x}}_{k|k}$ with the covariance matrix $\boldsymbol{P}^{xx}_{k|k}$ are calculated by

$$\boldsymbol{K}_k = \boldsymbol{P}^{xz}_{k|k-1}\left(\boldsymbol{P}^{zz}_{k|k-1}\right)^{-1}, \tag{3.7}$$

$$\widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\left(\boldsymbol{z}_k - \widehat{\boldsymbol{z}}_{k|k-1}\right), \tag{3.8}$$

$$\boldsymbol{P}^{xx}_{k|k} = \boldsymbol{P}^{xx}_{k|k-1} - \boldsymbol{K}_k\boldsymbol{P}^{zz}_{k|k-1}\boldsymbol{K}_k^T, \tag{3.9}$$

where $\widehat{\boldsymbol{z}}_{k|k-1} = \sum_{i=1}^{2n} w_i\boldsymbol{z}^i_{k|k-1}$ is the predicted measurement vector and $\boldsymbol{z}^i_{k|k-1} = \boldsymbol{h}(\boldsymbol{\chi}^i_{k|k-1})$; the self and cross-covariance matrices, $\boldsymbol{P}^{zz}_{k|k-1}$ and $\boldsymbol{P}^{xz}_{k|k-1}$, are respectively calculated by

$$\boldsymbol{P}^{zz}_{k|k-1} = \sum_{i=1}^{2n} w_i(\boldsymbol{z}^i_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1})(\boldsymbol{z}^i_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1})^T + \boldsymbol{R}_k, \tag{3.10}$$

$$\boldsymbol{P}^{xz}_{k|k-1} = \sum_{i=1}^{2n} w_i(\boldsymbol{\chi}^i_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})(\boldsymbol{z}^i_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1})^T. \tag{3.11}$$

3.2.3 Problem Statement

If the system process and measurement noise obey a Gaussian probability distribution, the filtered state, $\widehat{\boldsymbol{x}}_{k-1|k-1}$, will follow a Gaussian distribution as well. In that case, the sample mean and the sample covariance matrix of $\widehat{\boldsymbol{x}}_{k-1|k-1}$ will be captured by the sigma points and the UKF will produce reliable state estimates. However, the Gaussianity assumption may not hold true in practice. This is precisely the case in power systems; for instances, impulsive process noise may occur due to system model inaccuracy at a certain time window and the PMU measurement noise may not follow a Gaussian distribution as shown in Fig.1.1. This figure displays histograms and parametric probability density estimates of PMU errors on nodal voltage magnitudes and angles, line current magnitudes and angles, and line real and reactive powers. It is observed that except for the measurement errors on nodal voltage and line current angles, which are roughly Gaussian, the measurement errors on both nodal voltage and line current magnitudes obey a bimodal Gaussian mixture distribution. As for the measurement errors of line real and reactive powers calculated from voltage and current phasors, they follow a thick tailed distribution that may be approximated by either the Laplace or the Cauchy distribution. For more details about how to derive these error statistics, please refer to [81]. As a result, the sigma points may not capture the complete statistics of the state vector, resulting in poor or even diverging state estimates. Furthermore, since the UKF lacks statistical robustness, it is sensitive to any type of outliers, including

observation, innovation and structural outliers. When dealing with a general nonlinear dynamical system, three types of outliers, namely observation, innovation and structural outliers, are defined in [38]. In power system DSE, observation outliers refer to the phase biases and gross errors in PMU measurements [40]; innovation outliers may be induced by incorrect generator parameter values, failure of brushless exciter rotating diodes, or impulsive system process noise; and structural outliers may be caused by transmission parameter errors or topology errors. In the following section, we will propose a robust GM-UKF that is able to suppress observation and innovation outliers and to filter out various types of thick-tailed observation noises. Note that the problem of the identification and suppression of structural outliers is outside the scope of this paper since it requires a different formulation; it will be addressed in a future work.

3.3 Theory of the Proposed GM-UKF

The proposed GM-UKF is based on the following assumptions, which are satisfied by many engineering systems, including power systems:

- Assumption 1: The measurement redundancy is sufficient to make the state of the dynamical system observable. Recall that this condition is a prerequisite for state estimation.
- Assumption 2: The majority of the metered values are reliable. In other words, the percentage of bad data among the measurements does not exceed 25 percent, which is a sizable percentage that is rarely exceeded in practice.
- Assumption 3: The dynamical system model is sufficiently accurate. This assumption is acceptable since typically such a model is derived from physical laws and its parameters are known with reasonable precision.

Our filter consists of four major steps, namely a batch-mode regression form step, a robust pre-whitening step, a robust regression state estimation step, and a robust error covariance matrix updating step. They are described next.

3.3.1 Derivation of the Batch-Mode Regression Form

In this subsection, we first show the equivalence of statistical linearization and the unscented transformation using sigma points. We then derive the proposed batch-mode regression form. The former claim is presented in the following theorem:

Theorem 1. *Given the state estimate vector $\widehat{\boldsymbol{x}}_{k-1|k-1}$ and its associated covariance matrix $\boldsymbol{P}^{xx}_{k-1|k-1}$, statistical linear regression [54] applied to an arbitrary nonlinear function $\boldsymbol{g}(\boldsymbol{x})$ yields results that are equivalent to those of the unscented transformation using the sigma points generated according to (7.15).*

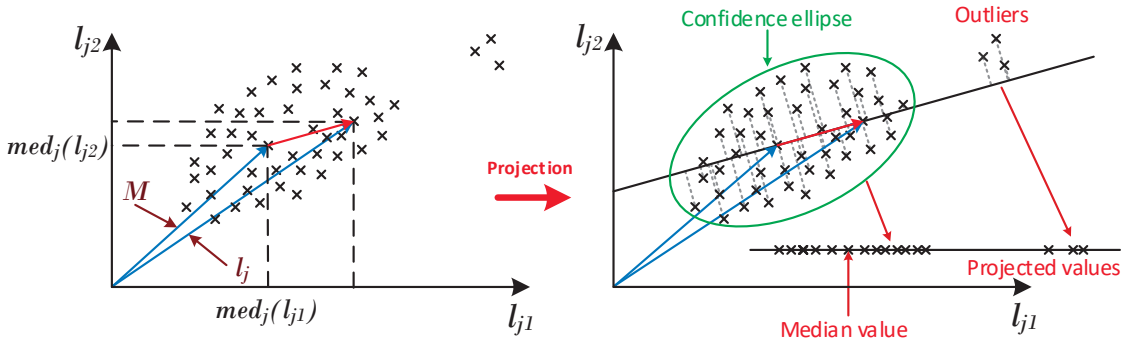


Figure 3.1: Application of the projection statistics to the matrix $\boldsymbol{Z}_k$ for detecting outliers in a two-dimensional dataset that consists of the innovation vectors and the predicted state vectors.

Proof. Consider a nonlinear function $\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x})$ evaluated in $2n$ points, i.e., $(\boldsymbol{\chi}_i, \boldsymbol{\gamma}_i)$, where $\boldsymbol{\gamma}_i = \boldsymbol{g}(\boldsymbol{\chi}_i)$ for i= 1,..., $2n$. Assuming that the nonlinear function is statistically linearized as $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b} + \boldsymbol{\zeta}$, the objective is to find $\widehat{\boldsymbol{A}}$ and $\widehat{\boldsymbol{b}}$ so that the point-wise linearization error $\boldsymbol{\zeta}_i$ is minimized, i.e.,

$$\left\{\widehat{\boldsymbol{A}}, \widehat{\boldsymbol{b}}\right\} = \arg\min \sum_{i=1}^{2n} w_i {\boldsymbol{\zeta}_i}^T \boldsymbol{\zeta}_i, \tag{3.12}$$

where $\boldsymbol{\zeta}_i = \boldsymbol{\gamma}_i - (\boldsymbol{A}\boldsymbol{\chi}_i + \boldsymbol{b})$. By taking the derivative of the objective function with respect to $\boldsymbol{A}$ and $\boldsymbol{b}$ and let them equal to zero, respectively, we obtain

$$\boldsymbol{b} = \overline{\boldsymbol{y}} - \widehat{\boldsymbol{A}}\overline{\boldsymbol{x}}, \tag{3.13}$$

$$\widehat{\boldsymbol{A}} = \boldsymbol{P}_{xy}^T \boldsymbol{P}_{xx}^{-1}, \tag{3.14}$$

where $\overline{\boldsymbol{x}} = \sum_{i=1}^{2n} w_i \boldsymbol{\chi}_i$; $\overline{\boldsymbol{y}} = \sum_{i=1}^{2n} w_i \boldsymbol{g}(\boldsymbol{\chi}_i) = \sum_{i=1}^{2n} w_i \boldsymbol{\gamma}_i$; $\boldsymbol{P}_{xx} = \sum_{i=1}^{2n} w_i (\boldsymbol{\chi}_i - \overline{\boldsymbol{x}})(\boldsymbol{\chi}_i - \overline{\boldsymbol{x}})^T$; $\boldsymbol{P}_{xy} = \sum_{i=1}^{2n} w_i (\boldsymbol{\chi}_i - \overline{\boldsymbol{x}})(\boldsymbol{\gamma}_i - \overline{\boldsymbol{y}})^T$. Then, the estimation error covariance matrix is calculated as

$$\begin{aligned} \boldsymbol{P}_{\zeta\zeta} &= \sum_{i=1}^{2n} w_i \widehat{\boldsymbol{\zeta}}_i \widehat{\boldsymbol{\zeta}}_i^T \\ &= \sum_{i=1}^{2n} w_i \left(\boldsymbol{\gamma}_i - \overline{\boldsymbol{y}} - \widehat{\boldsymbol{A}}(\boldsymbol{\chi}_i - \overline{\boldsymbol{x}})\right)\left(\boldsymbol{\gamma}_i - \overline{\boldsymbol{y}} - \widehat{\boldsymbol{A}}(\boldsymbol{\chi}_i - \overline{\boldsymbol{x}})\right)^T \\ &= \boldsymbol{P}_{yy} - \widehat{\boldsymbol{A}}\boldsymbol{P}_{xx}\widehat{\boldsymbol{A}}^T = \boldsymbol{P}_{yy} - \boldsymbol{P}_{xy}^T \boldsymbol{P}_{xx}^{-1} \boldsymbol{P}_{xy}, \end{aligned} \tag{3.15}$$

where $\boldsymbol{P}_{yy} = \sum_{i=1}^{2n} w_i (\boldsymbol{\gamma}_i - \overline{\boldsymbol{y}})(\boldsymbol{\gamma}_i - \overline{\boldsymbol{y}})^T$. Now, by taking the expectation and the outer product of the statistical linearized model, respectively, we obtain the posterior statistics given by

$$\widehat{\boldsymbol{y}} = \widehat{\boldsymbol{A}}\overline{\boldsymbol{x}} + \sum_{i=1}^{2n} w_i \boldsymbol{\chi}_i - \widehat{\boldsymbol{A}}\overline{\boldsymbol{x}} = \sum_{i=1}^{2n} w_i \boldsymbol{\chi}_i, \tag{3.16}$$

$$\begin{aligned}\boldsymbol{P}_{yy} &= \widehat{\boldsymbol{A}}\boldsymbol{P}_{xx}\widehat{\boldsymbol{A}}^T + \boldsymbol{P}_{\zeta\zeta} \\ &= \boldsymbol{P}_{xy}^T\boldsymbol{P}_{xx}^{-1}\boldsymbol{P}_{xy} + \sum_{i=1}^{2n} w_i\left(\boldsymbol{\gamma}_i - \overline{\boldsymbol{y}}\right)\left(\boldsymbol{\gamma}_i - \overline{\boldsymbol{y}}\right)^T - \boldsymbol{P}_{xy}^T\boldsymbol{P}_{xx}^{-1}\boldsymbol{P}_{xy} \\ &= \sum_{i=1}^{2n} w_i\left(\boldsymbol{\gamma}_i - \overline{\boldsymbol{y}}\right)\left(\boldsymbol{\gamma}_i - \overline{\boldsymbol{y}}\right)^T,\end{aligned} \tag{3.17}$$

which are the same expressions as those obtained by applying the unscented transformation to the nonlinear function $\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x})$, which completes the proof. □

By applying a statistical linearization to the nonlinear system process model, we obtain the predicted state vector $\widehat{\boldsymbol{x}}_{k|k-1}$ along with its covariance matrix $\boldsymbol{P}_{k|k-1}^{xx}$. We define $\widehat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{x}_k - \boldsymbol{\delta}_k$, where $\boldsymbol{x}_k$ is the true state vector; $\boldsymbol{\delta}_k$ is the prediction error; and $\mathbb{E}\left[\boldsymbol{\delta}_k\boldsymbol{\delta}_k^T\right] = \boldsymbol{P}_{k|k-1}^{xx}$. Then, statistical linearization can be applied to the nonlinear observation equation, yielding

$$\boldsymbol{z}_k = \boldsymbol{H}_k\left(\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k-1}\right) + \widehat{\boldsymbol{z}}_{k|k-1} + \boldsymbol{\nu}_k + \boldsymbol{v}_k, \tag{3.18}$$

where $\boldsymbol{H}_k = (\boldsymbol{P}_{k|k-1}^{xz})^T(\boldsymbol{P}_{k|k-1}^{xx})^{-1}$, which is no longer a Jacobian matrix. Here, the covariance of the statistical linearization error term is $\widetilde{\boldsymbol{R}}_k = \mathbb{E}\left[\boldsymbol{\nu}_k\boldsymbol{\nu}_k^T\right] = \boldsymbol{P}_{k|k-1}^{zz} - (\boldsymbol{P}_{k|k-1}^{xz})^T(\boldsymbol{P}_{k|k-1}^{xx})^{-1}\boldsymbol{P}_{k|k-1}^{xz}$, where $\boldsymbol{P}_{k|k-1}^{zz}$ and $\boldsymbol{P}_{k|k-1}^{xz}$ are two covariance matrices that are calculated by following the same steps as those of the UKF. By processing the predictions and the observations simultaneously, we get the following batch-mode regression form:

$$\begin{bmatrix} \boldsymbol{z}_k + \boldsymbol{H}_k\widehat{\boldsymbol{x}}_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1} \\ \widehat{\boldsymbol{x}}_{k|k-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_k \\ \boldsymbol{I} \end{bmatrix}\boldsymbol{x}_k + \begin{bmatrix} \boldsymbol{\nu}_k + \boldsymbol{v}_k \\ -\boldsymbol{\delta}_k \end{bmatrix} \tag{3.19}$$

which can be rewritten in a compact form as

$$\widetilde{\boldsymbol{z}}_k = \widetilde{\boldsymbol{H}}_k\boldsymbol{x}_k + \widetilde{\boldsymbol{e}}_k. \tag{3.20}$$

The error covariance matrix is given by

$$\boldsymbol{W}_k = \mathbb{E}\left[\widetilde{\boldsymbol{e}}_k\widetilde{\boldsymbol{e}}_k^T\right] = \begin{bmatrix} \boldsymbol{\Sigma}_{k|k-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}_{k|k-1}^{xx} \end{bmatrix} = \boldsymbol{S}_k\boldsymbol{S}_k^T, \tag{3.21}$$

where $\boldsymbol{\Sigma}_{k|k-1} = \mathbb{E}[(\boldsymbol{\nu}_k + \boldsymbol{v}_k)(\boldsymbol{\nu}_k + \boldsymbol{v}_k)^T] = \boldsymbol{R}_k + \widetilde{\boldsymbol{R}}_k$; $\boldsymbol{I}$ is an identity matrix; $\boldsymbol{S}_k$ is calculated by the Cholesky decomposition technique.

Theorem 2. *The weighted least squares estimator of the batch-mode regression form (7.32) yields an estimated state vector $\widehat{x}_{k|k}$ and its associated covariance matrix $\boldsymbol{P}_{k|k}^{xx}$ that are equivalent to those of the UKF.*

Proof. It is well-known that the state vector estimate based on the regression model given by (7.32) using the weighted least squares estimator is expressed as

$$\widehat{\boldsymbol{x}}_{k|k} = \left(\widetilde{\boldsymbol{H}}_k^T\boldsymbol{W}_k^{-1}\widetilde{\boldsymbol{H}}_k\right)^{-1}\widetilde{\boldsymbol{H}}_k^T\boldsymbol{W}_k^{-1}\widetilde{\boldsymbol{z}}_k, \tag{3.22}$$

with the covariance matrix $\boldsymbol{P}_{k|k}^{xx} = \left(\widetilde{\boldsymbol{H}}_k^T \boldsymbol{W}_k^{-1} \widetilde{\boldsymbol{H}}_k\right)^{-1}$. By applying an algebraic substitution and using the matrix inversion lemma, we get

$$\begin{aligned} \boldsymbol{P}_{k|k}^{xx} &= \left(\boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} \boldsymbol{H}_k + \left(\boldsymbol{P}_{k|k-1}^{xx}\right)^{-1}\right)^{-1} \\ &= \boldsymbol{P}_{k|k-1}^{xx} - \boldsymbol{P}_{k|k-1}^{xx} \boldsymbol{H}_k^T \left(\boldsymbol{H}_k \boldsymbol{P}_{k|k-1}^{xx} \boldsymbol{H}_k^T + \boldsymbol{R}_k\right)^{-1} \boldsymbol{H}_k \boldsymbol{P}_{k|k-1}^{xx} \\ &= \left(\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{H}_k\right) \boldsymbol{P}_{k|k-1}^{xx} = \boldsymbol{P}_{k|k-1}^{xx} - \boldsymbol{K}_k \boldsymbol{P}_{k|k-1}^{zz} \boldsymbol{K}_k^T, \end{aligned} \tag{3.23}$$

where the gain matrix is expressed as

$$\boldsymbol{K}_k = \boldsymbol{P}_{k|k-1}^{xx} \boldsymbol{H}_k^T (\boldsymbol{H}_k \boldsymbol{P}_{k|k-1}^{xx} \boldsymbol{H}_k^T + \boldsymbol{R}_k)^{-1} = \boldsymbol{P}_{k|k-1}^{xz} (\boldsymbol{P}_{k|k-1}^{zz})^{-1} \tag{3.24}$$

Thus, we can conclude that the estimation error covariance is identical to that of the UKF in (7.29). By applying similar substitutions and using the matrix inversion lemma, we can also show that the estimated state vector is given by

$$\widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k \left(\boldsymbol{z}_k - \widehat{\boldsymbol{z}}_{k|k-1}\right), \tag{3.25}$$

which completes the proof. □

Remark. *In the literature, a few Huber estimator-based robust UKF methods have been proposed and applied to various applications in signal processing, target tracking, to name a few [75–78]. However, in their developed regression models, $\boldsymbol{\nu}_k$ that compensates higher order Taylor-series expansion error terms is neglected. As a consequence, the estimation results are biased. In addition, although these estimators are able to suppress observation outliers, they fail to handle innovation outliers, which can happen quite frequently in power system due to controller parameter errors and model deficiency. Last but not the least, they cannot handle the cases when both process and observation noise are non-Gaussian.*

3.3.2 Robust Prewhitening

Before carrying out a robust regression, we uncorrelate the state prediction errors of the batch-mode regression form. This is achieved by pre-multiplying $\boldsymbol{S}_k^{-1}$ on both sides of (7.32), yielding

$$\boldsymbol{S}_k^{-1} \widetilde{\boldsymbol{z}}_k = \boldsymbol{S}_k^{-1} \widetilde{\boldsymbol{H}}_k \boldsymbol{x}_k + \boldsymbol{S}_k^{-1} \widetilde{\boldsymbol{e}}_k, \tag{3.26}$$

which can be put into the compact form given by

$$\boldsymbol{y}_k = \boldsymbol{C}_k \boldsymbol{x}_k + \boldsymbol{\xi}_k, \tag{3.27}$$

where $\mathbb{E}[\boldsymbol{\xi}_k \boldsymbol{\xi}_k{}^T] = \boldsymbol{I}$. However, if outliers occur, the application of $\boldsymbol{S}_k^{-1}$ will corrupt the prewhitening [37]. To overcome this problem, we first detect the outliers and calculate the weights using the projection statistics (PS) [37, 65]. Those weights will be incorporated in

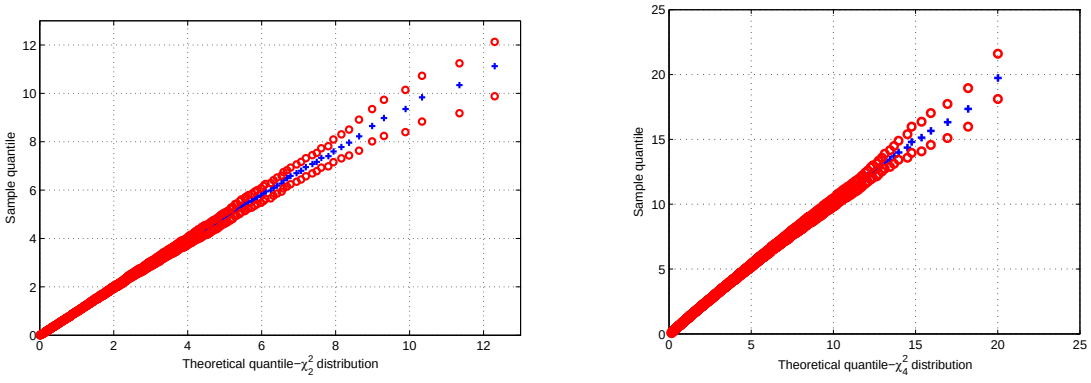


Figure 3.2: Q-Q plots of the sample quantiles of the PS vs. the corresponding quantiles of the χ^2_2 and χ^2_4 distributions, where (a) and (b) represent Q-Q plots of PS with Gaussian and Laplace noise, respectively.

the objective function that is defined in the proposed GM-estimator shown in Section III-C. Now, we describe the procedures used to calculate the weights. We apply the PS to a 2-dimensional matrix $\boldsymbol{Z}_k$ that contains serially correlated samples of the innovations and of the predicted state variables. Note that the innovation vector is defined as the difference between the observations and their associated predicted values at the previous step. Formally, we have

$$\boldsymbol{Z}_k = \begin{bmatrix} \boldsymbol{z}_{k-1} - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k-1|k-2}) & \boldsymbol{z}_k - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k|k-1}) \\ \widehat{\boldsymbol{x}}_{k-1|k-2} & \widehat{\boldsymbol{x}}_{k|k-1} \end{bmatrix}, \tag{3.28}$$

where $\boldsymbol{z}_{k-1} - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k-1|k-2})$ and $\boldsymbol{z}_k - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k|k-1})$ are the innovation vectors while $\widehat{x}_{k-1|k-2}$ and $\widehat{x}_{k|k-1}$ are the predicted state vectors at time instants k-1 and k, respectively. We may also apply the PS to higher dimensional samples, but we found that 2 dimensions are enough to identify outliers. Note that the PS values of the predictions and of the innovations are separately calculated because the values taken by the former and the latter are centered around different points.

The PS of the jth row vector, $\boldsymbol{l}_j^T$, of the predictions (respectively the innovations) in $\boldsymbol{Z}_k$ is defined as the maximum of the standardized projections of all the $\boldsymbol{l}_j^T$'s on every direction $\boldsymbol{\ell}$ that originates from the coordinatewise medians of the predictions (respectively the innovations) and that passes through every data point, and where the standardized projections are based on the sample median and the median-absolute-deviation [65]. The implementation of the PS to detect outliers in matrix $\boldsymbol{Z}_k$ is displayed in Fig. 3.1, while its mathematical expression is defined as [65]:

$$PS_j = \max_{\|\boldsymbol{\ell}\|=1} \frac{\left|\boldsymbol{l}_j^T\boldsymbol{\ell} - med_i\left(\boldsymbol{l}_i^T\boldsymbol{\ell}\right)\right|}{1.4826\ med_\kappa \left|\boldsymbol{l}_\kappa^T\boldsymbol{\ell} - med_i\left(\boldsymbol{l}_i^T\boldsymbol{\ell}\right)\right|}, \tag{3.29}$$

for $i, j, \kappa = 1, 2, ..., m+n$.

Once the PS values are calculated, they are compared to a statistical threshold to identify

outliers. Extensive Monte Carlo simulations and Q-Q plots reveal that the probability distributions of the PS applied to $\boldsymbol{Z}_k$, whose data points obey bivariate Gaussian and Laplace probability distributions, follow chi-square distributions with degree of freedom 2 and 4, respectively (See Fig. 3.2). This investigation allows us to apply statistical tests to the PS and to flag all the data points that satisfy $\mathrm{PS}_i > d$ as outliers. The latter are downweighted via

$$\varpi_i = \min\left(1, d^2/PS_i^2\right), \tag{3.30}$$

where the parameter d is set equal to 1.5 to yield good statistical efficiency at different distributions without increasing too much the bias induced by outliers. As an example, when the noise is assumed to be Laplacian, the PS obeys a chi-square distribution with 4 degrees of freedom. In that case, we can choose the statistical detection threshold as $\chi^2_{4,0.975}$ at a significance level of 97.5%.

3.3.3 Robust Filtering and Solution

To suppress the outliers and filter out thick-tailed non-Gaussian measurement noise, we develop a robust GM-estimator that minimizes the following objective function:

$$J\left(\boldsymbol{x}_k\right) = \sum_{i=1}^{m+n} \varpi_i^2 \rho\left(r_{S_i}\right), \tag{3.31}$$

where ϖ_i is calculated by (5.20); $r_{S_i} = r_i/s\varpi_i$ is the standardized residual; $r_i = y_i - \boldsymbol{c}_i^T\widehat{\boldsymbol{x}}$ is the residual, where $\boldsymbol{c}_i^T$ is the ith row vector of the matrix $\boldsymbol{C}_k$; $s = 1.4826 \cdot b_m \cdot \mathrm{median}_i\left|r_i\right|$ is the robust scale estimate; b_m is a correction factor to achieve unbiasedness for a finite sample of size $m+n$ at a given probability distribution [65]; $\rho(\cdot)$ is the nonlinear function of r_{S_i}. In this chapter, the convex Huber-ρ function [82] is adopted, that is

$$\rho\left(r_{S_i}\right) = \begin{cases} \frac{1}{2}r_{S_i}^2, & \text{for } \left|r_{S_i}\right| < \lambda \\ \lambda\left|r_{S_i}\right| - \lambda^2/2, & elsewhere \end{cases}, \tag{3.32}$$

where the parameter λ between the quadratic and the linear segment of $\rho(\cdot)$ is typically chosen between 1.5 to 3 in the literature.

To minimize (5.21), one takes its partial derivative with respect to $\boldsymbol{x}_k$ and sets it equal to zero, yielding

$$\frac{\partial J\left(\boldsymbol{x}_k\right)}{\partial \boldsymbol{x}_k} = \sum_{i=1}^{m+n} -\frac{\varpi_i \boldsymbol{c}_i}{s}\psi\left(r_{S_i}\right) = \boldsymbol{0}, \tag{3.33}$$

where $\psi\left(r_{S_i}\right) = \partial\rho\left(r_{S_i}\right)/\partial r_{S_i}$is the so-called ψ-function. By dividing and multiplying the standardized residual r_{S_i} to both sides of (5.23) and putting it in a matrix form, we get

$$\boldsymbol{C}_k^T\widehat{\boldsymbol{Q}}\left(\boldsymbol{y}_k - \boldsymbol{C}_k\boldsymbol{x}_k\right) = \boldsymbol{0}, \tag{3.34}$$

where $\widehat{\boldsymbol{Q}} = \text{diag}(q(r_{S_i}))$ and $q(r_{S_i}) = \psi(r_{S_i})/r_{S_i}$.

By using the IRLS algorithm [37,66], the state vector correction at the j iteration is calculated through

$$\Delta\widehat{\boldsymbol{x}}_{k|k}^{(j+1)} = \left(\boldsymbol{C}_k^T \widehat{\boldsymbol{Q}}^{(j)} \boldsymbol{C}_k\right)^{-1} \boldsymbol{C}_k^T \widehat{\boldsymbol{Q}}^{(j)} \boldsymbol{y}_k, \tag{3.35}$$

where $\Delta\widehat{\boldsymbol{x}}_{k|k}^{(j+1)} = \widehat{\boldsymbol{x}}_{k|k}^{(j+1)} - \widehat{\boldsymbol{x}}_{k|k}^{(j)}$. The algorithm converges when $\left\|\Delta\widehat{\boldsymbol{x}}_{k|k}^{(j+1)}\right\|_\infty \leq 10^{-2}$.

3.3.4 Asymptotic Error Covariance Matrix of the GM-UKF State Estimates

Upon convergence of the iterative algorithm, the error covariance matrix $\boldsymbol{P}_{k|k}^{xx}$ is updated so that the state prediction for the next step can be performed. To this end, consider the ϵ-contamination model $G = (1-\epsilon)\,\Phi + \epsilon\Delta_r$, where G and Φ are the contaminated and the true cumulative probability distribution function of the residual, respectively; Δ_r is the point mass to model outliers. Inspired by Fernholtz [83] and Hampel *et al.* [67], the error covariance matrix is updated based on the following theorem:

Theorem 3. *Let $\boldsymbol{T}(\cdot)$ be the functional form of the GM-estimator with a bounded $\psi(\cdot)$ function and Φ_α be the empirical cumulative probability distribution function, then*

$$\sqrt{\alpha}(\boldsymbol{T}(\Phi_\alpha) - \boldsymbol{T}(\Phi)) \xrightarrow{d} \mathcal{N}(\boldsymbol{0}, \boldsymbol{P}_{k|k}^{xx}), \tag{3.36}$$

where $\alpha = m + n$; $\xrightarrow{d}$ means convergence in probability distribution; the estimation error covariance matrix is $\boldsymbol{P}_{k|k}^{xx} = \mathbb{E}[\boldsymbol{IF}(\boldsymbol{x}; \Phi, \boldsymbol{T}) \cdot \boldsymbol{IF}(\boldsymbol{x}; \Phi, \boldsymbol{T})^T]$ with the total influence function $\boldsymbol{IF}(\boldsymbol{x}; \Phi, \boldsymbol{T})$ evaluated at Φ.

Proof. By taking a von Mises expansion of the functional form of the estimator $\boldsymbol{T}$ with respect to Φ, we get

$$\boldsymbol{T}(\Phi_\alpha) = \boldsymbol{T}(\Phi) + \boldsymbol{T}'(\Phi_\alpha - \Phi) + \text{Rem}(\Phi_\alpha - \Phi), \tag{3.37}$$

which can be reorganized into the following form by multiplying $\sqrt{\alpha}$ on both sides of the equality:

$$\begin{aligned}
\sqrt{\alpha}\left(\boldsymbol{T}\left(\Phi_\alpha\right) - \boldsymbol{T}\left(\Phi\right)\right) &= \sqrt{\alpha}\boldsymbol{T}'\left(\Phi_\alpha - \Phi\right) + \sqrt{\alpha}\text{Rem}\left(\Phi_\alpha - \Phi\right) && (3.38)\\
&= \sqrt{\alpha}\int \boldsymbol{IF}(\boldsymbol{x}; \Phi, \boldsymbol{T}) d(\Phi_\alpha - \Phi) + \sqrt{\alpha}\text{Rem}(\Phi_\alpha - \Phi) && (3.39)\\
&= \sqrt{\alpha}\int \boldsymbol{IF}(\boldsymbol{x}; \Phi, \boldsymbol{T}) d\Phi_\alpha + \sqrt{\alpha}\text{Rem}(\Phi_\alpha - \Phi) && (3.40)\\
&= \frac{1}{\sqrt{\alpha}}\sum_{i=1}^{\alpha} \boldsymbol{IF}\left(x_i; \Phi, \boldsymbol{T}\right) + \sqrt{\alpha}\text{Rem}\left(\Phi_\alpha - \Phi\right), && (3.41)
\end{aligned}$$

where the definition of the influence function is applied to yield (3.38) to (3.39). By virtues of Fisher consistency at the distribution Φ, that is, $\int \boldsymbol{IF}(\boldsymbol{x};\Phi,\boldsymbol{T})d\Phi = \boldsymbol{0}$, (39) reduces to (40). Finally, by using the property of the empirical cumulative probability distribution function, we have

$$\int \boldsymbol{IF}(\boldsymbol{x};\Phi,\boldsymbol{T})d\Phi_\alpha = \frac{1}{\alpha}\sum_{i=1}^{\alpha} \boldsymbol{IF}\left(x_i;\Phi,\boldsymbol{T}\right), \tag{3.42}$$

yielding (3.40) to (3.41).

Following the work of Fernholz [83], we can show that

$$\sqrt{\alpha}\mathrm{R}em\left(\Phi_\alpha - \Phi\right) \xrightarrow{p} 0, \tag{3.43}$$

where $\xrightarrow{p}$ means probability convergence. Therefore, by applying the central limit theorem and Slutsky's lemma to (3.41), it follows that

$$\sqrt{\alpha}(\boldsymbol{T}(\Phi_\alpha) - \boldsymbol{T}(\Phi)) \xrightarrow{d} \mathcal{N}(\boldsymbol{0}, \boldsymbol{P}_{k|k}^{xx}), \tag{3.44}$$

where $\boldsymbol{P}_{k|k}^{xx} = \mathbb{E}[\boldsymbol{IF}(\boldsymbol{x};\Phi,\boldsymbol{T}) \cdot \boldsymbol{IF}(\boldsymbol{x};\Phi,\boldsymbol{T})^T]$. □

Discussion: The UKF is able to provide good results only when the process and observation noises obey a Gaussian distribution [52]. Indeed, in that case the filtered state vector $\widehat{\boldsymbol{x}}_{k|k}$ is Gaussian and the mean and covariance matrix of $\widehat{\boldsymbol{x}}_{k|k}$ can be accurately estimated by the sample mean and the sample covariance matrix of the sigma points. However, this property no longer holds true if the Gaussianity assumption of the noises is violated. In that case, the state estimate vector $\widehat{\boldsymbol{x}}_{k|k}$ obtained from the UKF is significantly biased due to the lack of statistical robustness of the filter to thick-tailed non-Gaussian noise. By contrast, our GM-UKF guarantees the asymptotic Gaussianity of $\widehat{\boldsymbol{x}}_{k|k}$ for thick-tailed non-Gaussian noises and yields reliable state estimates with good statistical efficiency. It is worth pointing out that the robustness of the Huber M-estimator-based UKF to outliers and to non-Gaussian noises reported in [75–77] has not been confirmed by the simulations carried out on the IEE New-England 39-bus 10-machine system. In fact, it is found that this estimator provides strongly biased state estimates for that system as shown in Section V.

Let us now derive the total influence function, $\boldsymbol{IF}(\boldsymbol{x};\Phi,\boldsymbol{T})$, of our GM-UKF at the cumulative probability distribution Φ.

Corollary 3.1. *The total influence function of the GM-UKF defined by (5.23) using the regression model (5.18) is expressed as*

$$\boldsymbol{IF}(\boldsymbol{x};\Phi,\boldsymbol{T}) = \left[\int \frac{1}{s}\psi'\left(r_{S_i}\right)\boldsymbol{C}\boldsymbol{C}^{\boldsymbol{T}}\Big|_{T(\Phi)} d\Phi\right]^{-1} \varpi \boldsymbol{C}\psi\left(r_{S_i}\right). \tag{3.45}$$

Proof. In our previous work [38], the total influence function of a GM-estimator based on a nonlinear regression model given by $\boldsymbol{y} = \boldsymbol{\varphi}(\boldsymbol{x}) + \boldsymbol{\xi}$ is expressed as

$$\begin{aligned} &\boldsymbol{IF}(\boldsymbol{x}; \Phi, \boldsymbol{T}) \\ &= \left(\int \left\{ \frac{\psi'\left(\boldsymbol{r}_{S_i}\right)}{s} \frac{\partial \boldsymbol{\varphi}(\boldsymbol{x})}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{\varphi}(\boldsymbol{x})}{\partial \boldsymbol{x}}^T - \varpi \psi\left(\boldsymbol{r}_{S_i}\right) \boldsymbol{D} \right\} \big|_{T(\Phi)} d\Phi \right)^{-1} \\ &\quad \cdot \varpi \frac{\partial \boldsymbol{\varphi}(\boldsymbol{x})}{\partial \boldsymbol{x}} \psi\left(\boldsymbol{r}_{S_i}\right), \end{aligned} \tag{3.46}$$

where $\boldsymbol{D} = \frac{\partial^2 \boldsymbol{\varphi}(\boldsymbol{x})}{\partial x_i \partial x_j}$ is the Hessian matrix of $\boldsymbol{\varphi}(\boldsymbol{x})$. Since we have $\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{C}\boldsymbol{x}$ for the GM-UKF, (3.46) reduces to

$$\boldsymbol{IF}(\boldsymbol{x}; \Phi, \boldsymbol{T}) = \left[\int \frac{1}{s} \psi'\left(r_{S_i}\right) \boldsymbol{C}\boldsymbol{C}^{\boldsymbol{T}} \Big|_{T(\Phi)} d\Phi \right]^{-1} \varpi \boldsymbol{C} \psi\left(r_{S_i}\right). \tag{3.47}$$

□

Now, we are in a position to derive the covariance matrix $\boldsymbol{P}_{k|k}^{xx}$ from (3.44). First, let us prove the following theorem:

Theorem 4. *The sample variance of the robust scale estimator s of the GM-standardized residuals tends to one as the number of observation tends to infinity.*

Proof. By virtue of the law of large numbers, the distribution of the residuals tends to the Gaussian distribution, i.e., $\Phi \sim \mathcal{N}(\mu, \sigma^2)$. Since the median absolute deviation (MAD=$\frac{s}{1.4826 \cdot b_m}$) is a consistent estimator for the standard deviation σ of a Gaussian distribution, we get

$$\begin{aligned} \tfrac{1}{2} &= P\left(|X - \mu| \leq \tfrac{s}{1.4826 \cdot b_m} \right) = P\left(\tfrac{|X-\mu|}{\sigma} \leq \tfrac{s}{1.4826 \cdot b_m \cdot \sigma} \right) \\ &= 2\Phi\left(\tfrac{s}{1.4826 \cdot b_m \cdot \sigma} \right) - 1. \end{aligned} \tag{3.48}$$

Therefore, we obtain $s/\sigma = 1.4826 \cdot b_m \cdot \Phi^{-1}\left(\frac{3}{4}\right) = b_m \to 1$ as m tends to infinity, where Φ is the cumulative probability function of the standard Gaussian distribution. On the other hand, from (5.18) and $\mathbb{E}[\boldsymbol{\xi}_k \boldsymbol{\xi}_k{}^T] = \boldsymbol{I}$, the residuals can be shown to follow the standard Gaussian distribution. Therefore, $\mathbb{E}_F[s^2] = s^2 \to 1$. □

Finally, by virtue of Theorems 3 and 4, the asymptotic error covariance matrix of our GM-UKF at time sample k is updated by

$$\begin{aligned} \boldsymbol{P}_{k|k}^{xx} &= \mathbb{E}[\boldsymbol{IF}(\boldsymbol{x}; \Phi, \boldsymbol{T}) \cdot \boldsymbol{IF}(\boldsymbol{x}; \Phi, \boldsymbol{T})^T] \\ &= \frac{\mathbb{E}_\Phi\left[\psi^2\left(r_{S_i}\right)\right]}{\left\{\mathbb{E}_\Phi\left[\psi'\left(r_{S_i}\right)\right]\right\}^2} \left(\boldsymbol{C}_k^T \boldsymbol{C}_k\right)^{-1} \left(\boldsymbol{C}_k^T \boldsymbol{Q}_\varpi \boldsymbol{C}_k\right) \left(\boldsymbol{C}_k^T \boldsymbol{C}_k\right)^{-1} \end{aligned} \tag{3.49}$$

where $\boldsymbol{Q}_\varpi = diag\left(\varpi_i^2\right)$.

It should be noted that thanks to the Gaussian normality of our GM-UKF state estimates and residuals, the bulk of the data, including forecasted state variables and received observations, obey Gaussian distribution. As a result, our GM-UKF always achieves a high statistical efficiency for a broad range of noise probability distributions. This will be demonstrated on the IEEE 10-machine system in Section V-G.

Remark: To achieve robustness to outliers or model deviations, an estimator has to sacrifice some estimation efficiency. Under ideal conditions, including perfect linear system model, known Gaussian system process and measurement noise and reliable data without any outliers, the Kalman filter is the maximum likelihood estimator, yielding 100% estimation efficiency. Under these assumptions, our proposed GM-UKF achieves at least 95% statistical efficiency as demonstrated in Section V-G. However, if any of these assumptions is violated, the UKF achieves quite poor estimation efficiency and may provide strongly biased state estimates. By contrast, our GM-UKF is able to withstand these deviations and yield a slightly decreased estimation efficiency. Since the aforementioned assumptions are usually violated in practice, our GM-UKF is guaranteed to provide reliable state estimates and to yield better efficiency than the UKF under the foregoing conditions.

3.4 Application to Power System

The developed robust GM-UKF is a generic technique for online monitoring of many dynamical cyber-physical systems, including smart grids, autonomous vehicles, aircraft tracking, GPS tracking and navigation, radar systems, to name a few. In this chapter, we take power system dynamic state estimation as an illustrative example to demonstrate the capabilities of our GM-UKF for suppressing observation and innovation outliers while being able to filter out non-Gaussian noise.

For an electric power system, its discrete-time state space representation can be expressed as in (5.8) and (5.9) as follows:

$$\boldsymbol{x}_k = \boldsymbol{f}\left(\boldsymbol{x}_{k-1}, \boldsymbol{u}_k\right) + \boldsymbol{w}_k, \tag{3.50}$$

$$\boldsymbol{z}_k = \boldsymbol{h}\left(\boldsymbol{x}_k, \boldsymbol{u}_k\right) + \boldsymbol{v}_k, \tag{3.51}$$

where the state vector $\boldsymbol{x}_k$ contains the rotor angle, rotor speed, the d- and q- axis state variables of the synchronous generator, the exciter, the voltage regulator, and the governor. $\boldsymbol{z}_k$ is the measurement vector that contains a collection of voltage phasors, current phasors, real and reactive power flows and power injections. The noise $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$, which are not necessary to be Gaussian noise, are assumed to be white and independent of each other. $\boldsymbol{u}_k$ represents the input vector; $\boldsymbol{h}(\cdot)$ is the vector-valued measurement function while $\boldsymbol{f}(\cdot)$ is the vector-valued function that relates $\boldsymbol{x}_k$ to $\boldsymbol{x}_{k-1}$. In this chapter, the detailed 9th order two-axis model with IEEE-DC1A exciter and TGOV1 turbine-governor are implemented and tested. It can be represented by the following differential and algebraic equations [84]:

Differential equations:

$$T'_{do}\frac{dE'_q}{dt} = -E'_q - (X_d - X'_d)\, I_d + E_{fd}, \tag{3.52}$$

$$T'_{qo}\frac{dE'_d}{dt} = -E'_d - \left(X_q - X'_q\right) I_q, \tag{3.53}$$

$$\frac{d\delta}{dt} = \omega - \omega_s, \tag{3.54}$$

$$\frac{2H}{\omega_s}\frac{d\omega}{dt} = T_M - P_e - D\left(\omega - \omega_s\right), \tag{3.55}$$

$$T_E\frac{dE_{fd}}{dt} = -\left(K_E + S_E\left(E_{fd}\right)\right) E_{fd} + V_R, \tag{3.56}$$

$$T_{\mathrm{F}}\frac{dV_F}{dt} = -V_F + \frac{K_F}{T_E}V_R - \frac{K_F}{T_E}\left(K_E + S_E\left(E_{fd}\right)\right) E_{fd}, \tag{3.57}$$

$$T_A\frac{dV_R}{dt} = -V_R + K_A\left(V_{ref} - V_F - V\right), \tag{3.58}$$

$$T_{CH}\frac{dT_M}{dt} = -T_M + P_{SV}, \tag{3.59}$$

$$T_{SV}\frac{dP_{SV}}{dt} = -P_{SV} + P_C - \frac{1}{R_D}\left(\frac{\omega}{\omega_s} - 1\right), \tag{3.60}$$

Algebraic equations:

$$V_d = V\sin\left(\delta - \theta\right), V_q = V\cos\left(\delta - \theta\right), \tag{3.61}$$

$$I_d = \frac{E'_q - V_q}{X'_d}, I_q = \frac{V_d - E'_d}{X'_q}, \tag{3.62}$$

$$P_e = V_d I_d + V_q I_q, Q_e = -V_d I_q + V_q I_d, \tag{3.63}$$

where T'_{do}, T'_{qo}, T_E, T_F, T_A, T_{CH} and T_{SV} are time constants, in seconds; K_E, K_F and K_A are controller gains; V_{ref} and P_C are known control inputs; E'_q, E'_d, E_{fd}, V_F, V_R, T_M and P_{SV} are the q-axis and d-axis transient voltages, field voltage, scaled output of the stabilizing transformer and scaled output of the amplifier, synchronous machine mechanical torque and steam valve position, respectively; X_d, X'_d, X_q and X'_q are generator parameters; V and θ are the terminal bus voltage magnitude and phase angle, respectively; P_e and Q_e are the active and reactive electrical power outputs; I_d and I_q are the d and q axis currents, respectively.

By applying time discretization to (7.3)-(7.14) using the 4th-order Ruger-Kutta method, we get (5.8)-(5.9), yielding the state vector given by $\boldsymbol{x}_k = [\delta\ \omega\ E'_d\ E'_q\ E_{fd}\ V_F\ V_R\ T_M\ P_{SV}]$. The relationships given by (7.3)-(7.11) and (7.12)-(7.14) are represented in compact forms by the vector-valued function $\boldsymbol{f}(\cdot)$, which relates $\boldsymbol{x}_k$ to $\boldsymbol{x}_{k-1}$, and by $\boldsymbol{h}(\cdot)$, respectively. The system input vector is denoted by $\boldsymbol{u}_k = [V_{ref}\ P_C]^T$. The measurement vector $\boldsymbol{z}_k$ contains a collection of voltage phasor measurements $V\angle\theta$ and real and reactive power injections P_e and Q_e. They are measured by the PMUs.

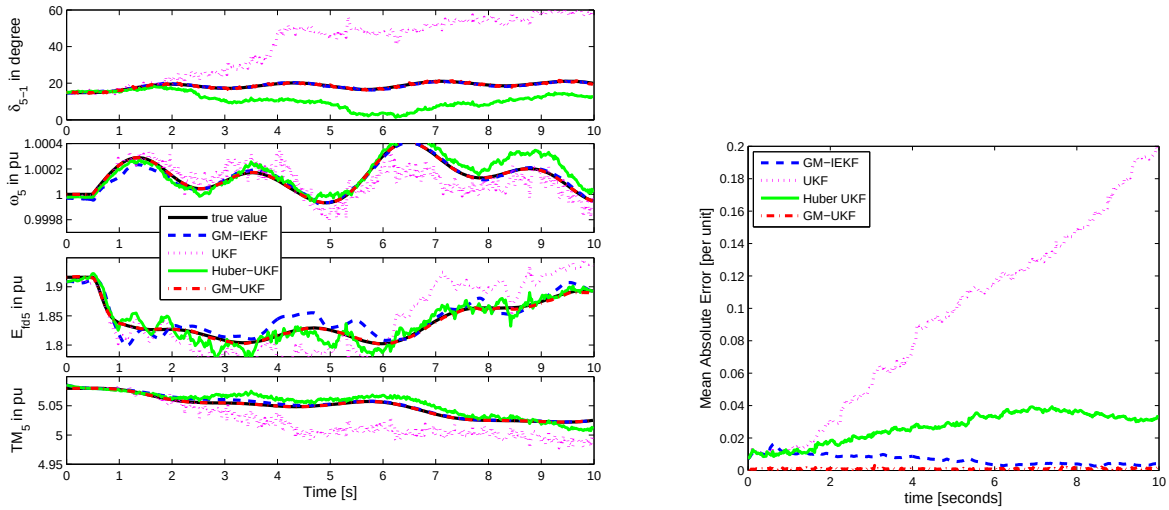
Remark: There exist two ways of implementing a DSE to estimate power system dynamic state variables: centralized estimation vs. decentralized estimation. The centralized DSE approach requires accurate knowledge of each system component and all real-time wide-area PMU measurements. Although the estimation results are helpful for global control applications, it requires large computing times that may not be compatible with real-time applications in large-scale interconnected power systems. By contrast, the decentralized DSE approach is implemented at each local synchronous generator, and therefore, is fast to execute. The proposed GM-UKF can be implemented in both ways but only the decentralized GM-UKF is used for demonstration in this chapter.

3.5 Numerical Results

The performances of our GM-UKF to handle thick-tailed non-Gaussian noise along with innovation and observation outliers are assessed on the IEEE 39-bus system. The UKF, the robust Huber-UKF [75–77] and the GM-IEKF [38] are included for comparisons. The time-domain simulation results are used to generate a collection of samples of the nodal voltage magnitudes and phase angles as well as of the real and reactive power injections at the terminal buses of all the generators. A sampling rate of 50 samples/second is assumed. A synthetic noise is added to the true values following the probability distributions displayed in Fig. 1. Specifically, a zero mean Gaussian noise is assumed for the voltage angles, a bimodal Gaussian mixture distribution is assumed for the noise of the voltage magnitudes, and either a Laplace or a Cauchy distribution is assumed for the noise of the real and reactive power measurements. The two-axis generator model is assumed and tested, whose parameters are taken from [70]. A disturbance is applied at t=0.5s by opening the transmission line between Buses 15 and 16. The maximal number of iterations allowed for the IRLS algorithm is 20. The parameters λ and d are set to 1.5. The convergence tolerance threshold of the IRLS algorithm is 0.01. For the state initialization, the steady-state values with 10% errors are used. Due to space limitation, not all the 9 state variables of each generator are shown; instead estimated values of the rotor angle and speed, the field voltage and the mechanical power of Generator 5 are utilized for illustration purposes. The mean absolute error (MAE) is utilized to evaluate the overall performance of each method.

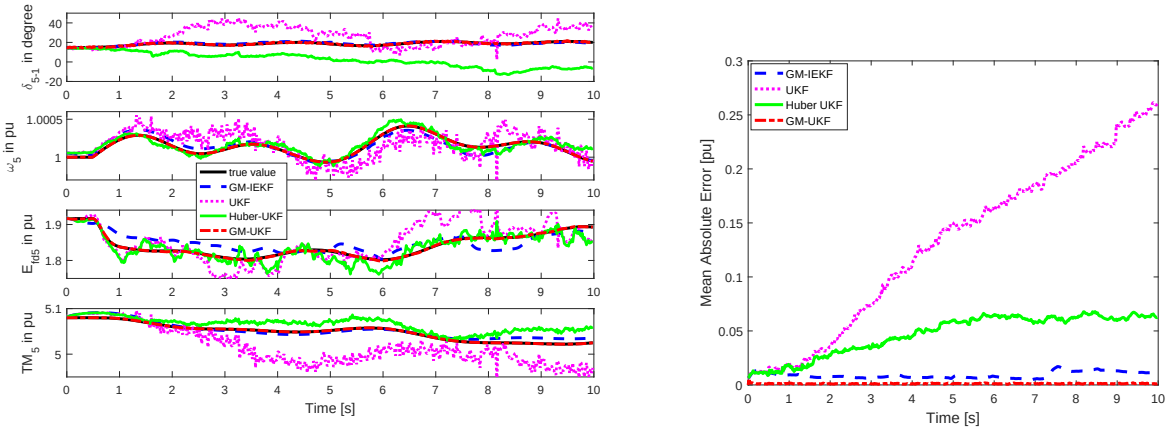
3.5.1 Case 1: Thick-tailed Non-Gaussian Measurement Noise without Outliers

In this section, we first evaluate the performance of the GM-IEKF, the UKF, the Huber-UKF and the GM-UKF under normal conditions. Specifically, a zero mean Gaussian noise with covariance matrices of $10^{-4}\boldsymbol{I}$ with the appropriate dimensions are assumed for the process noise and the noise of voltage and current angles; the noise of the voltage magnitudes follows a bimodal Gaussian mixture with zero mean, variances of 10^{-4} and 10^{-3} and weights of 0.9 and 0.1, respectively; a Laplace noise with zero mean and scale 0.02 is added to the real



(a) (b)

Figure 3.3: Case 1: Tracking performance of the GM-IEKF, UKF, the Huber-UKF and GM-UKF without outliers; (a) the estimated rotor angle and speed, field voltage and mechanical power of Generator 5 are used for illustration purposes; (b) mean absolute error of each of the four filters.



(a) (b)

Figure 3.4: Tracking performance of the GM-IEKF, UKF, the Huber-UKF and GM-UKF when the variances of the process and measurement noise change over time; (a) the estimated rotor angle and speed, field voltage and mechanical power of Generator 5 are used for illustration purposes; (b) the mean absolute error of each of the four filters.

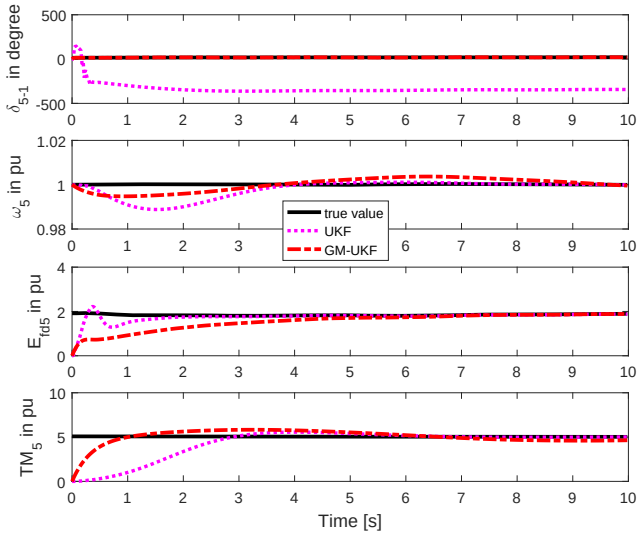
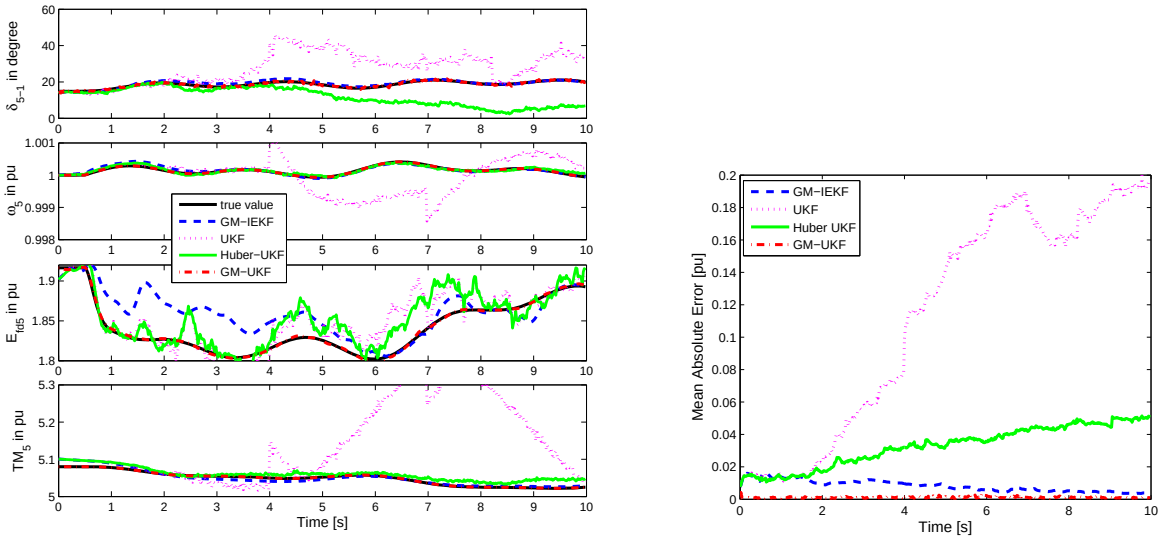


Figure 3.5: Tracking performance of the GM-IEKF, UKF, the Huber-UKF and GM-UKF when all state variables are initialized with zeros.

and reactive power injections measurements. The test results are displayed in Fig. 3.3. It is observed that the UKF is not able to cope with a Laplacian noise even in the absence of outliers. By contrast, the GM-IEKF and the Huber-UKF can filter out such noise to a certain degree. However, their statistical efficiencies are much lower than our GM-UKF. In particular, the GM-IEKF poorly estimates the field voltage and the rotor speed while the Huber-UKF tracks the rotor angle and the field voltage poorly. By observing Fig. 3.3, it is interesting to note that the turbine mechanical power is changing during the transient process. According to the CIGRE report [85], it can significantly vary when control features such as fast valving or special protection schemes are used to limit the output of the steam driven generator during transients. Consequently, it is of vital importance to not assume it to be fixed at a constant steady-state value as commonly done in most of the literature, but to obtain accurate dynamic state variables of the governor for controls and stability analysis.

In practical power systems, the process and measurement noise may change over time, yielding unknown noise statistics. To further validate the performance of the proposed GM-UKF under these conditions, we consider the following scenario: the variance of the system process noise is time varying and follows a Gaussian distribution with mean 10^{-2} and standard deviation 2×10^{-3}; the variances of the Gaussian mixture model are assumed to be time varying as well; in particular, their variances are assumed to follow Gaussian distributions with means 10^{-4} and 10^{-3}, and standard deviations 4×10^{-5} and 3×10^{-4}, respectively; the scales of the Laplace distributions are assumed to follow Gaussian distributions with means 0.25 and 0.3, and standard deviations 2×10^{-3} and 3×10^{-3}, respectively. The test results are shown in Fig. 3.4. It is observed from the figure that due to the time-varying pro-

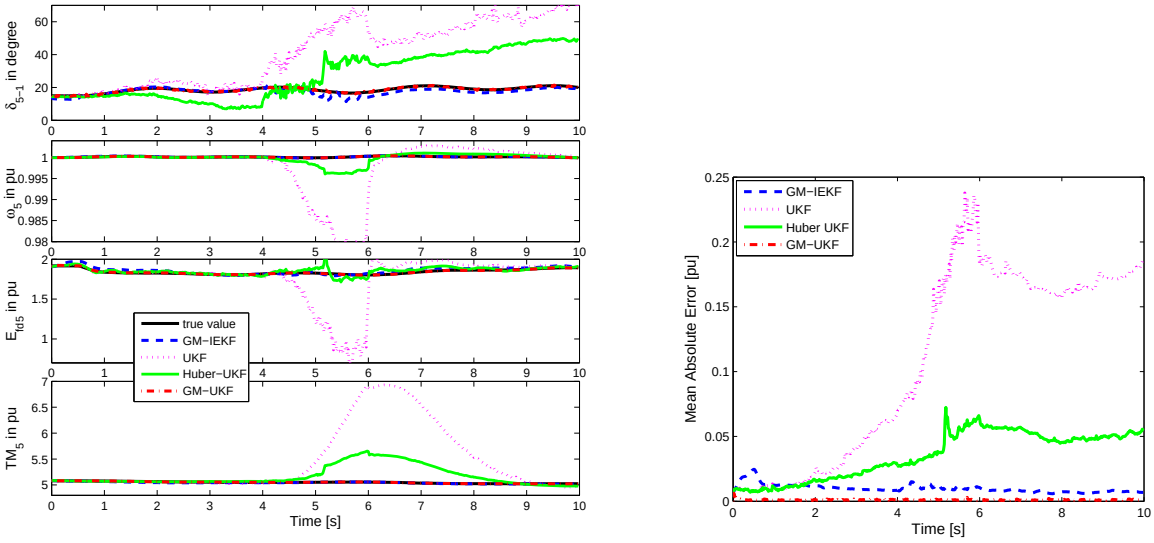


(a) (b)

Figure 3.6: Case 2: Tracking performance of the GM-IEKF, the UKF, the Huber-UKF and the GM-UKF in the presence of observation outliers from t=4s to t=6s, where the real and reactive power measurements of Generator 5 are corrupted with 20% errors; (a) the estimated rotor angle and speed, field voltage and mechanical power of Generator 5 are used for illustration purposes; (b) mean absolute error of each of the four filters.

cess and observation noise, the estimates of UKF and Huber-UKF show a lot of oscillations and impulses. In addition, UKF, Huber-UKF and GM-IEKF obtain increased estimation errors. By contrast, our proposed GM-UKF is able to filter out time varying process and measurement noise, yielding much better results than the other alternatives.

We also test the case when all the state variables are initialized to zeros to show the convergence speed of the proposed GM-UKF. It is worth pointing out that this is an extreme case because for practical power systems, we can always infer a reasonable initialization for the state variables from the power flow solutions. The test results are displayed in Fig. 3.5. Note that due to the incorrect initialization and the non-Gaussian noise, both the GM-IEKF and the Huber-UKF diverge and their results are not shown. In fact, the UKF diverges as well since it does not estimate the rotor angle correctly. By contrast, our proposed GM-UKF is able to track the true state variables although it needs some time to converge due to incorrect state initialization.

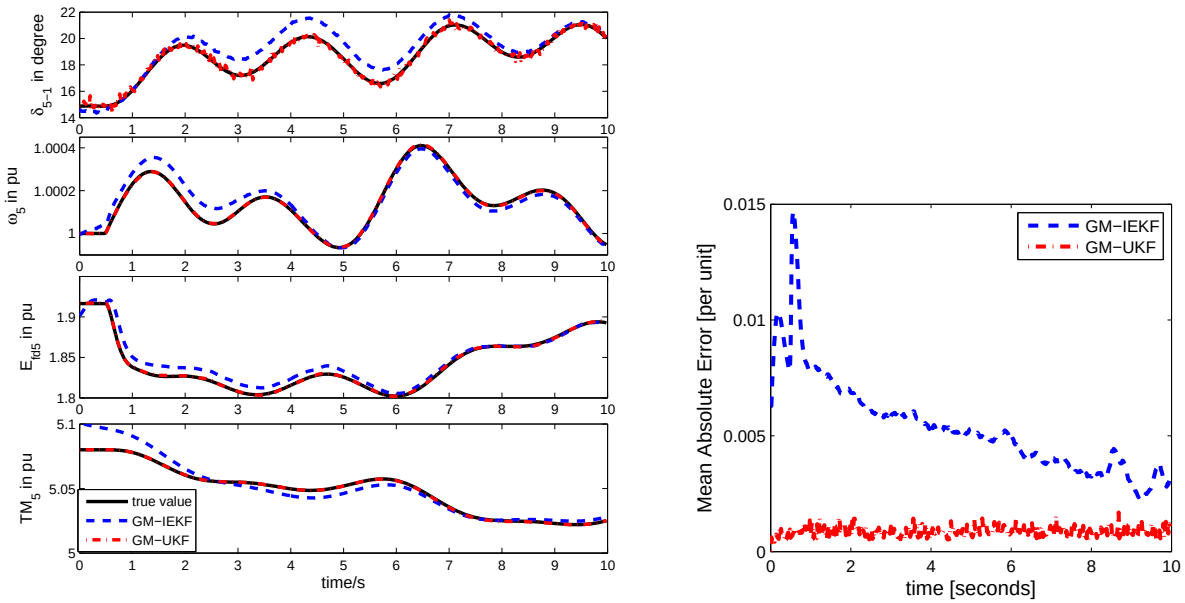


(a) (b)

Figure 3.7: Case 3: Tracking performance of the GM-IEKF, the UKF, the Huber-UKF and the GM-UKF in the presence of innovation outliers from t=4s to t=6s, where the predicted rotor angle of Generator 5 is corrupted with 20% errors; (a) the estimated rotor angle and speed, field voltage and mechanical power of Generator 5 are used for illustration purposes; (b) mean absolute error of each of the four filters.

3.5.2 Case 2: Thick-tailed Non-Gaussian Measurement Noise with Observation Outliers

The settings are the same as those of Case 1 except for the presence of observation outliers from t=4s to t=6s. The latter are simulated by adding 20% errors to the real and the reactive power measurements of Generator 5. The results are presented in Fig. 3.6. From this figure, we observe that the UKF is not robust to observation outliers since it yields significantly biased results. Although the GM-IEKF and Huber-UKF can handle them, it produces increased biases on the estimates at the time when observation outliers occur. By contrast, the GM-UKF suppresses the outliers and produces much less bias than the GM-IEKF and Huber-UKF. Note that the Gaussianity of the GM-estimator used in the estimation step of the GM-UKF allows that method to filter out thick-tailed noise while its statistical robustness enables it to suppress the outliers, hence achieving very good estimates.

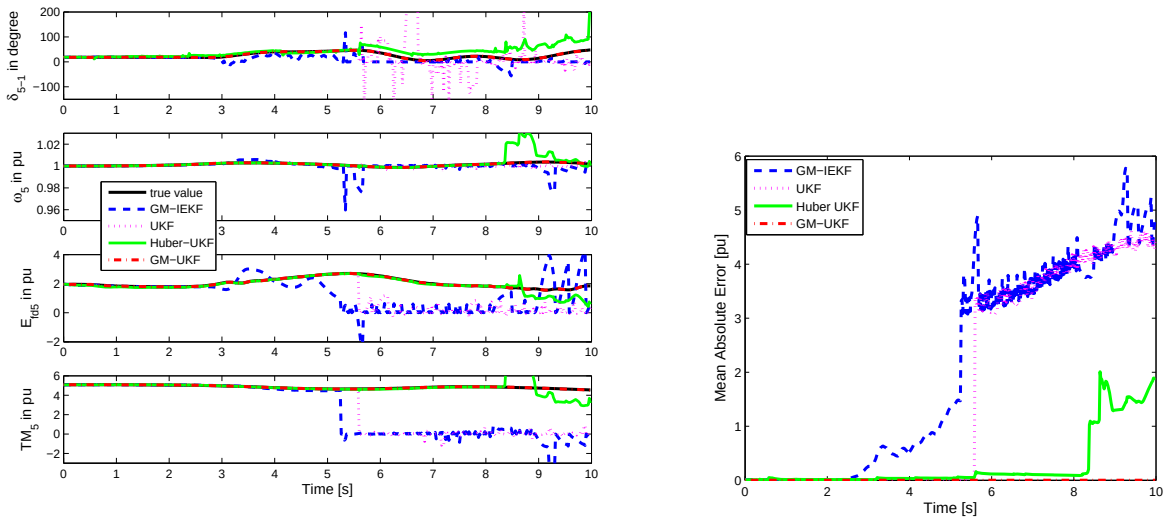


(a) (b)

Figure 3.8: Case 4: Tracking performance of the GM-IEKF, the UKF, the Huber-UKF and the GM-UKF in the presence of Cauchy power measurement noise. The estimated rotor angle and speed, field voltage and mechanical power of Generator 5 are used for illustration purposes; (b) mean absolute error of each of the four filters. Since the UKF and Huber-UKF diverge, their results are not shown in the figure.

3.5.3 Case 3: Thick-tailed Non-Gaussian Measurement Noise with Innovation Outliers

The settings are the same as those of Case 1 except for the presence of innovation outliers from t=4s to t=6s. They are simulated by adding 20% errors to the predicted rotor angle of Generator 5. This innovation outlier is induced by a gross parameter value in the model. The comparison results are shown in Fig. 3.7. As expected, due to the non-robustness of the UKF, it is unable to handle innovation outliers. Furthermore, Huber-UKF fails to suppress that as well since it cannot distinguish outliers from observations and predicted state variables. On the other hand, the GM-IEKF can handle them, but it produces larger biases compared with Case 2. This can be explained by the fact that the model errors will not only affect the predicted state vector but also will produce a smearing effect throughout the Jacobian matrix. As a result, it downweights several good measurements. By contrast, our GM-UKF is capable of handling both observation and innovation outliers, yielding comparable performances. Because only the sigma points associated with the model errors will be affected and downweighted, this filter obtains better estimates than the GM-IEKF in presence of model errors.



(a) (b)

Figure 3.9: Case 5: Tracking performance of the GM-IEKF, the UKF, the Huber-UKF and the GM-UKF in the presence of strong system nonlinearity. (a) The estimated rotor angle and speed, field voltage and mechanical power of Generator 5 are used for illustration purposes; (b) mean absolute error of each of the four filters.

3.5.4 Case 4: Long Tail Cauchy Measurement Noise

As discussed before, the real and the reactive power measurement noises may obey a Cauchy distribution as well, which is a very thick tailed distribution with no moments being defined. To test the capability of our GM-UKF to handle that case, we assume that the simulation settings are the same as those of Case 1 except that the Cauchy noises with zero median and a scale of 0.005 is added to the real and the reactive power injection measurements. The obtained test results are displayed in Fig. 3.8. Note that in presence of Cauchy measurement noise, the UKF and Huber-UKF have a non-positive definite covariance matrix, resulting in their divergences. By contrast, thanks to the Gaussian normality and robustness of our GM-UKF, the total influence function-based covariance matrix updating approach can always guarantee its positive-definiteness. On the other hand, compared with the results obtained when using Laplacian power measurement noises, the GM-IEKF produces larger biases of the state estimates and takes much longer time to approach the true system states. This is not the case for our GM-UKF since it achieves a comparable performance and tracks the system state at the very beginning of the transient process.

3.5.5 Case 5: Robustness to Strong System Nonlinearity

Practical systems may be heavily loaded, resulting in strong nonlinear dynamics. To illustrate the capability of our GM-UKF to handle that case, we assume that the load at Bus 7 is increased from 233.8 MW to 1500 MW to stress the system before switching Line 15-16 while the other simulation settings are the same as those of Case 1. Note that the steady-state maximum loadability at Bus 7 is around 2000 MW. After the line switching, the system operates under even greater stressed conditions. The test results are displayed in Fig. 3.9. It is observed from these two figures that the GM-IEKF and UKF fail to converge at the very beginning of the transient process while the Huber-UKF diverges latter on. By contrast, our GM-UKF is able to handle this scenario while achieving excellent tracking performance. The underlying reasons are as follows:

- Under stressed system operating conditions, the first-order Taylor series expansion used in the EKF and the GM-IEKF is too approximate and is unable to account for strong system nonlinearities. As a result, these two filters produce very large approximation errors and eventually diverge;
- Thanks to the sigma-points-based unscented transformation and its approximation accuracy up to at least third-order Taylor series expansion, the UKF, Huber-UKF and the GM-UKF are able to handle strong system nonlinearities. However, due to the accumulative estimation error induced by the non-Gaussian measurement error, the estimation error covariance matrices of the UKF and Huber-UKF are close to non-positive definitiveness. Therefore, they produce large estimation errors and finally diverge. By contrast, our GM-UKF leverages the strength of the unscented transformation to handle system nonlinearities while the robustness of the GM-estimator allows it to filter out thick-tailed non-Gaussian measurement noise, yielding good estimation results.

3.5.6 Breakdown Point of the GM-UKF to Cyber Attacks

To investigate the breakdown point of the GM-UKF to cyber attacks, which is defined as the maximum number of outliers that the filter can handle without yielding unreliable estimates, we carry out extensive simulations on the IEEE 39-bus test system using the concept of finite sample breakdown in nonlinear regression introduced by Stromberg and Ruppert [48]. By replacing a varying percentage of observations by outliers in the vector $\boldsymbol{y}_k$, it is observed that the GM-UKF can handle at least 25% of corrupted observations. It is worth noting that the breakdown point of the GM-estimator in nonlinear regression is still unknown. This problem will be investigated as a future work. Another interesting problem is the determination of the maximum breakdown point that any regression estimator may have in structured nonlinear regression such as power system state estimation problems; this will be an interesting extension of the results proved in Mili and Coakley [47] in the linear case.

Table 3.1: Average Computing Times of the Four DSE Methods For Every PMU Sample, where NA represents not applicable.

Cases	UKF	GM-IEKF	Huber-UKF	GM-UKF
Case 1	6.28ms	9.64ms	8.56ms	9.52ms
Case 2	6.31ms	9.68ms	8.62ms	9.55ms
Case 3	6.38ms	9.72ms	8.74ms	9.63ms
Case 4	NA	9.70ms	NA	9.59ms
Case 5	6.42ms	9.80ms	8.78ms	9.70ms

3.5.7 Discussions about the Statistical Efficiency of the GM-UKF

Let us now show that our proposed GM-UKF with λ=1.5 has a high efficiency under various probability distributions of the measurement noise as depicted in Fig. 1.1. This can be done by analyzing the derived asymptotic variance matrix shown in (5.26). Under Gaussian noise assumed for all the measurements in the system, our filter exhibits a high statistical efficiency of 96.4%. Under the Laplace distribution for the real and reactive power measurement noises while the measurement noises of the voltage magnitudes and phase angles respectively obey a Gaussian mixture and the Gaussian distribution as indicated in Fig. 1.1, our filter exhibits an average efficiency of 96.4% for 500 samples (10 s). These high efficiencies do not come as a surprise since for λ tending to infinity or to zero, the GM-estimator becomes a maximum likelihood estimator at the Gaussian or at the Laplace distribution, respectively. Finally, under the Cauchy distribution of the real and reactive power measurement noises, our GM-UKF has an average efficiency of 83.3% for 500 samples. A lower efficiency is expected here because the Cauchy distribution has much thicker tails than the two others. Recall that the Cauchy tails are so thick that no moment exists. Note that when the estimation error covariance matrix is updated, all the outliers with respect to the Gaussian distribution, which include the measurements associated with the tails of the Laplace or the Cauchy distribution or other thick-tailed distributions, will be heavily downweighted through the matrix $\boldsymbol{Q}_{\varpi}$, yielding bounded biases and variances in the state estimates.

3.5.8 Computational Efficiency

To validate the applicability of the proposed GM-UKF to online estimation with a PMU sampling rate of 30 or 60 samples per second, its computational efficiency is analyzed and compared to that of the UKF, the Huber-UKF and the GM-IEKF in Cases 1-5. The test is performed on a PC with Intel Core i5, 2.50 GHz, 8GB of RAM. The average computing time of each method for every PMU sample is displayed in Table 3.1. We observe from this table that the UKF has the best computational efficiency, exhibiting computing times much lower than the PMU sampling period, which are 33.3ms and 16.7ms for 30 sample/s and

60 samples/s, respectively. Although the execution times of the GM-UKF are a little bit longer, they are still smaller than the PMU sampling period, demonstrating its ability to track system real-time dynamic states.

3.6 Conclusion

In this chapter, the robust GM-UKF is proposed to handle non-Gaussian noise and outliers. We show first that the UKF estimates the state vector via a weighted least squares estimator under the Gaussianity assumption of the system process or measurement noises; consequently, it yields strongly biased state estimates when the noises follow non-Gaussian probability distributions, which is precisely the case when processing PMU measurements. By contrast, the state estimates and residuals of our GM-UKF are proved to be asymptotically Gaussian, allowing the sigma points to reliably approximate the mean and the covariance matrices of the predicted and corrected state vectors. Furthermore, by relying on the projection statistics and the GM-estimator, the proposed GM-UKF is able to suppress observation and innovation outliers while exhibiting high statistical efficiency of the state estimates. In addition, we derive the expression of the asymptotic error covariance matrix of the GM-UKF state estimates from the total influence function of the GM-estimator. Numerical results carried out on the IEEE 39-bus 10-machine system demonstrate the effectiveness and robustness of the proposed method.

Chapter 4

Power System Robust Decentralized Dynamic State Estimation Against Outliers Based on Multiple Hypothesis Testing

4.1 Introduction

The wide-area deployment of phasor measurement units (PMUs) on power transmission grids has made possible the real-time monitoring and control of power system dynamics. However, these functions may not be reliably achieved without the development of a fast and robust dynamic state estimator (DSE). Indeed, the benefits of using a DSE are an improved dynamic security assessment scheme and an enhanced local and global system control, to cite a few.

To enable the implementation of a power system DSE, several types of Kalman filters are advocated, including the extended Kalman filter (EKF), the unscented Kalman filter (UKF), the ensemble Kalman filter (EnKF) [21, 23, 24, 26, 38, 45, 86, 87]. In [21, 38, 45], centralized DSEs using EKF, UKF and EnKF have been proposed. They all require accurate knowledge of the status of each system component as well as the real-time wide-area PMU measurements. Although a global dynamic state estimation may be implemented to achieve global control applications, they may require large computing times that are incompatible for applications in large-scale interconnected power systems. This motivates the development of decentralized DSE that is implemented at each local synchronous generator, and therefore, is fast to execute [23, 24, 26, 86, 87]. In [23, 24, 86], EKF-based DSEs are proposed to estimate the state variables of classical and fourth-order synchronous generators. Since the power system can have strong nonlinearities when operating under stressed conditions or subject to severe disturbances, the first-order Taylor series expansion-based EKF may induce large

estimation errors. To circumvent this difficulty, UKF-based DSEs are advocated [26, 87]. Note that the model decoupling approach of [23, 24, 26, 86, 87] to enable the decentralized implementation of a DSE is achieved by treating metered generator terminal voltage phasors as inputs and real and reactive power injections or current phasors as outputs. On the other hand, although decentralized DSE is faster to execute compared with the centralized approaches, its global measurement redundancy decreases significantly, and as a result, they are more vulnerable to outliers. For a general cyber-physical system that involves dynamic model and measurements, three types of outliers, namely the observation, innovation and structural outliers, are defined in [37]. In power system DSE problem, they are defined specifically as follows: observation outliers refer to the received PMU measurements providing unreliable metered values due to gross errors, cyber attacks, and measurement losses; innovation outliers are typically caused by impulsive system process noise due to unknown disturbances and sudden changes of the model inputs while structural outliers are induced by incorrect parameters of the generators or its associated controllers, such as exciters, speed governors. In the literature [23, 24, 26, 86, 87], only observation outliers are considered [26], where the normalized innovation vector-based statistical test is advocated to detect them. However, the detection threshold of this approach is system-dependent and sensitive to the change of process and measurement noises, which limits its practical value. In addition, this approach may fail to work in presence of innovation or structural outliers.

In this chapter, a fast and robust unscented Kalman filter-based decentralized DSE is proposed. Our robust DSE is able to detect, identify, and suppress observation, innovation and structural outliers [88]. To enable the fast and decentralized estimation of generator states of large-scale power systems, two model decoupling approaches are presented and compared. The first one is achieved by treating the metered generator terminal voltage phasors as inputs and the current phasors and frequency as outputs. By contrast, the second one takes measured current phasors as inputs while treating voltage phasors and frequency as outputs. It is shown that the latter one has higher statistical efficiency than the former one in the presence of both small and large measurement noise. To detect and distinguish the three types of outliers, projection statistics-based multiple hypothesis testing approach is proposed. Specifically, three hypotheses corresponding to the occurrence of three types of outliers are tested using three different innovation matrices; these matrices are made up of time-correlated innovation vectors and/or predicted states and/or measurements; then projection statistics are applied to each of the innovation matrix and its calculated projection values are checked via a statistical test to validate the assumed hypothesis. Then, a generalized maximum-likelihood type (GM)-estimator that uses the convex Huber score function is advocated to suppress the identified outliers, yielding robust state estimates. Finally, the estimation error covariance matrix is updated through the total influence function-based approach.

The remainder of the paper is organized as follows. Section 4.2 presents the problem formulation. Section 4.3 describes the proposed robust DSE and Section 4.4 shows and analyzes the simulation results. Finally, Section 4.5 concludes the paper.

4.2 Problem Formulation

Decentralized DSE is implemented at each synchronous generator using local PMU measurements. Consequently, synchronous generator model and measurement functions are derived first. In this chapter, the two-axis model with IEEE-DC1A exciter and TGOV1 turbine-governor is considered. It is represented by the following differential and algebraic equations [84],

Differential equations of generator:

$$T'_{do}\frac{dE'_q}{dt} = -E'_q - \left(X_d - X'_d\right) I_d + E_{fd}, \tag{4.1}$$

$$T'_{qo}\frac{dE'_d}{dt} = -E'_d - \left(X_q - X'_q\right) I_q, \tag{4.2}$$

$$\frac{d\delta}{dt} = \omega - \omega_s, \tag{4.3}$$

$$\frac{2H}{\omega_s}\frac{d\omega}{dt} = T_M - P_e - D\left(\omega - \omega_s\right), \tag{4.4}$$

Differential equations of IEEE-DC1A exciter:

$$T_E\frac{dE_{fd}}{dt} = -\left(K_E + S_E\left(E_{fd}\right)\right) E_{fd} + V_R, \tag{4.5}$$

$$T_{\mathrm{F}}\frac{dV_F}{dt} = -V_F + \frac{K_F}{T_E}V_R - \frac{K_F}{T_E}\left(K_E + S_E\left(E_{fd}\right)\right) E_{fd}, \tag{4.6}$$

$$T_A\frac{dV_R}{dt} = -V_R + K_A\left(V_{ref} - V_F - V\right), \tag{4.7}$$

Differential equations of TGOV1 turbine-governor:

$$T_{CH}\frac{dT_M}{dt} = -T_M + P_{SV}, \tag{4.8}$$

$$T_{SV}\frac{dP_{SV}}{dt} = -P_{SV} + P_C - \frac{1}{R_D}\left(\frac{\omega}{\omega_s} - 1\right), \tag{4.9}$$

Algebraic equations:

$$V_d = V\sin\left(\delta - \theta\right), V_q = V\cos\left(\delta - \theta\right), \tag{4.10}$$

$$I_d = \frac{E'_q - V_q}{X'_d}, I_q = \frac{V_d - E'_d}{X'_q}, \tag{4.11}$$

$$P_e = V_d I_d + V_q I_q, Q_e = -V_d I_q + V_q I_d, \tag{4.12}$$

where T'_{do}, T'_{qo}, T_E, T_F, T_A, T_{CH} and T_{SV} are time constants, in seconds; K_E, K_F and K_A are controller gains; V_{ref} and P_C are known control inputs; E'_q, E'_d, E_{fd}, V_F, V_R, T_M and P_{SV}

are the q-axis and d-axis transient voltages, field voltage, scaled output of the stabilizing transformer and scaled output of the amplifier, synchronous machine mechanical torque and steam valve position, respectively; X_d, X'_d, X_q and X'_q are generator parameters; V and θ are the terminal bus voltage magnitude and phase angle, respectively; P_e and Q_e are the active and reactive electrical power outputs; I_d and I_q are the d and q axis currents, respectively.

The above differential and algebraic equations can be put into the following discrete-time state space form:

$$\boldsymbol{x}_k = \boldsymbol{f}\left(\boldsymbol{x}_{k-1}, \boldsymbol{u}_k\right) + \boldsymbol{w}_k, \tag{4.13}$$

$$\boldsymbol{z}_k = \boldsymbol{h}\left(\boldsymbol{x}_k\right) + \boldsymbol{v}_k, \tag{4.14}$$

where (6.12) and (6.13) correspond to equations (7.3)-(7.11) and (7.12)-(7.14), respectively; $\boldsymbol{x}_k$ is the state vector that includes state variables of the synchronous generator, the exciter and the governor; $\boldsymbol{z}_k$ is the measurement vector that contains a collection of generator terminal voltage phasors, current phasors and frequency; the noises $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are assumed to be white and with covariance matrices $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$, respectively; $\boldsymbol{u}_k$ represents the input vector; $\boldsymbol{f}(\cdot)$ and $\boldsymbol{h}(\cdot)$ are the vector-valued functions. It is worth pointing out that V_{ref} and P_C are the voltage and power references of the exciter and speed governor, respectively, whose values are adjusted by the local generator operator. Hence, it is reasonable to consider them as the known generator parameters. Therefore, only local PMU measurements and generator parameters are needed to perform decentralized DSE.

Problem statement: the decentralized DSEs proposed in [23, 24, 26, 86, 87] are vulnerable to outliers and do not achieve high statistical efficiency in presence of thick tailed-probability distribution of the measurement noise. This paper will develop a robust DSE to address these problems while satisfying the following constraints: given a limited number of PMUs installed at the terminal bus of each generator, i) decouple the generator from each other to enable the decentralized state estimation and thus make the algorithm suitable for large-scale power system applications; ii) detect, distinguish and suppress three types of outliers; iii) if generator parameter errors occur, locate them in order to perform appropriate model calibration and validation.

4.3 The Proposed Robust Decentralized DSE

We first present two generator decoupling approaches to enable the decentralized implementation of a DSE. Then, we derive the proposed robust decentralized DSE using multiple hypothesis testing to detect, identify, and suppress the three types of outliers.

4.3.1 Generator Model Decoupling Approaches

Once a disturbance takes place at one point of the power system, synchronous generators will response to it. These responses reveal themselves in their terminal voltage or current

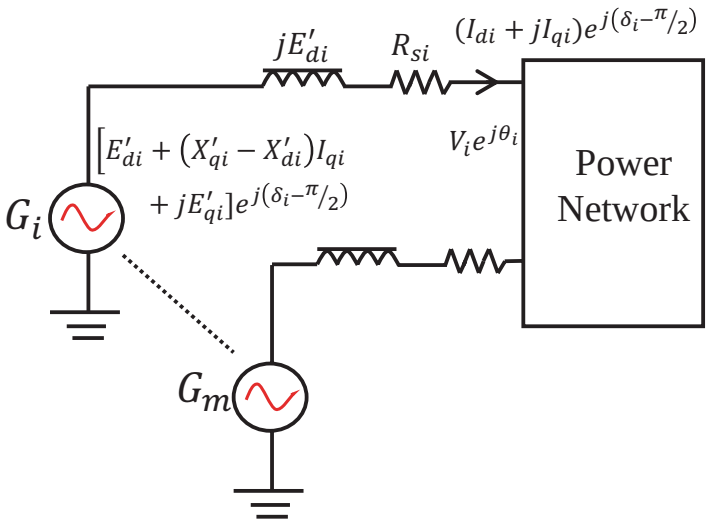


Figure 4.1: Interconnected system with m synchronous generators

phasors [26]. In other words, a generator is coupled with the power system at the point of connection, and its interactions with the rest of the system are through the terminal voltage and current phasors. If the generator terminal voltage or current phasors are measured by PMUs, its responses to the disturbance are captured completely and no other system information is required. By doing so, a generator model can be decoupled from the rest of the system from a DSE prospective and its dynamic states can be estimated using only local measurements [26]. Based on the way of treating these measurement pairs, two decoupling approaches can be achieved: A1) taking the metered terminal voltage phasors as inputs and the current phasors and frequency as outputs; A2) taking the metered terminal current phasors as inputs and the voltage phasors and frequency as outputs. It should be noted that at the model decoupling stage the generator terminal real and reactive powers are not taken as model outputs as advocated by several references [23,24,86]. This is due to the fact that gross error on single voltage or current phasor will propagate to both real and reactive powers, causing a smearing effect. On the other hand, as shown by [89], the terminal real and reactive powers, when measured, enable a better observability of the generator states than either the voltage phasors or the current phasors. Therefore, in this chapter if both the voltage phasors and current phasors are not detected as outliers by the proposed multiple hypothesis testing shown in the next section, the terminal real and reactive power will be taken as outputs. Otherwise, if any voltage phasors and current phasors are flagged as outliers, the outputs in the two model decoupling approaches will be preferred. We illustrate the two decoupling approaches A1) and A2) on an interconnected power system with m synchronous generators shown in Fig. 4.1, and the ith generator is taken as an example.

A1: Taking the current phasor $I_i\angle\phi_i$ as known inputs and the voltage phasor $V_i\angle\theta_i$ and frequency f_i as outputs, we have $\boldsymbol{u}_i = [I_i\ \phi_i]^T$, $I_{di} = \text{real}[I_i e^{j(\phi_i-\delta_i+\frac{\pi}{2})}]$, $I_{qi} = \text{Imag}[I_i e^{j(\phi_i-\delta_i+\frac{\pi}{2})}]$, $V_{di} = E'_{di} + I_{qi}X'_{qi}$ and $V_{qi} = E'_{qi} - I_{di}X'_{di}$. Define $\boldsymbol{z}_i = [z_1\ z_2\ z_3]^T$, the measurement function

can be expressed as

$$z_1 = V_i = \sqrt{V_{di}^2 + V_{qi}^2} + v_{Vi}, \tag{4.15}$$

$$z_2 = \theta_i = \delta_i - arctan(\frac{V_{di}}{V_{qi}}) + v_{\theta i}, \tag{4.16}$$

$$z_3 = f_i = f_0(\Delta\omega_i + 1) + v_{fi}, \tag{4.17}$$

where $\boldsymbol{v}_i = [v_{Vi}\ v_{\theta i}\ v_{fi}]^T$ is the measurement noise vector.

A2: Taking the voltage phasor $V_i\angle\theta_i$ as known inputs and the current phasor $I_i\angle\phi_i$ and frequency f_i as outputs, we have $\boldsymbol{u}_i = [V_i\ \theta_i]^T$, $V_{di} = V_i\sin(\delta_i - \theta_i)$, $V_{qi} = V_i\cos(\delta_i - \theta_i)$, $I_{di} = (E'_{qi} - V_{qi})/X'_{di}$ and $I_{qi} = (-E'_{di} + V_{di})/X'_{qi}$. Define $\boldsymbol{z}_i = [z_1\ z_2\ z_3]^T$, the measurement function can be expressed as

$$z_1 = I_i = \sqrt{I_{di}^2 + I_{qi}^2} + v_{Ii}, \tag{4.18}$$

$$z_2 = \phi_i = \arg(I_{qi} - jI_{di}) + \delta_i + v_{\phi i}, \tag{4.19}$$

$$z_3 = f_i = f_0(\Delta\omega_i + 1) + v_{fi}, \tag{4.20}$$

where $\boldsymbol{v}_i = [v_{Ii}\ v_{\phi i}\ v_{fi}]^T$ is the measurement noise vector.

Note that if our proposed multiple hypothesis testing method does not flag any voltage phasors and current phasors as outliers, we advocate the use of the terminal real and reactive power as outputs. As a result, z_1 and z_2 of the two decoupling approaches A1) and A2) are rewritten as

$$P_{ei} = V_{di}I_{di} + V_{qi}I_{qi} + v_{Pi}, \tag{4.21}$$

$$Q_{ei} = -V_{di}I_{qi} + V_{qi}I_{di} + v_{Qi}, \tag{4.22}$$

where $\boldsymbol{v}_i = [v_{Pi}\ v_{Qi}\ v_{fi}]^T$ is the measurement noise vector. Therefore, the dynamical system model of the ith generator at the time sample k can be expressed as

$$\boldsymbol{x}_k^i = \boldsymbol{f}_i\left(\boldsymbol{x}_{k-1}^i, \boldsymbol{u}_k^i\right) + \boldsymbol{w}_k^i, \tag{4.23}$$

$$\boldsymbol{z}_k^i = \boldsymbol{h}_i\left(\boldsymbol{x}_k^i\right) + \boldsymbol{v}_k^i, \tag{4.24}$$

where $\boldsymbol{w}_k^i$ is the system process noise that includes the noise associated with the input vector $\boldsymbol{u}_i$ and the model approximation error.

Remark: In a multi-machine power system, a common reference is required to define the voltage and rotor angles. Usually, the angle of the synchronous machine that has the largest installed capacity is used as the reference [84]. Interestingly, North American Electric Reliability Corporation (NERC) has a regulation rule of installing PMUs to monitor and calibrate the machine with the largest generation capacity [90]. Therefore, it is not difficult to obtain the reference angle. On the other hand, although [87, 91] presented an approach to avoid the definition of the reference angle for decentralized DSE by redefining the i-th rotor angle with $\alpha_i = \delta_i - \theta_i$, it may suffer from the following two drawbacks:

- After the dynamic state estimation, the estimated α_i of the ith local generator can be used to estimate its internal rotor angle through $\delta_i = \alpha_i + \theta_i$. However, θ_i is always subject to noise or even gross errors. As a result, the estimate of δ_i is biased, which may subsequently deteriorate the performance of the control actions;

- To damp out inter-area oscillations, the knowledge of just a few local generator dynamic state variables may not be sufficient. Coordinated controls are usually required and in this case, the reference angle can be easily defined at the coordination level. Interestingly, this has been implicitly shown in [26]. Note that inappropriate local controls of inter-area oscillations can produce adverse effects.

Based on the above discussions, we conclude that the choice of a reference angle depends on specific applications.

4.3.2 The Proposed Robust DSE

The model decoupling approach enables a generator to be decoupled from the rest of the system model, which in turn allows us to rely only on local measurements to estimate the state variables of a generator through (6.20) and (6.21). To this end, this paper presents a robust decentralized UKF-based DSE, termed the RDUKFI. Our choice of the UKF as the basic filter is motivated by the fact that it achieves a more balanced performance between computational efficiency and ability to cope with strong system nonlinearities than other filters [113]. The RDUKFI consists of four major steps, namely a batch-mode regression form step, an outlier detection and processing step using multiple hypothesis testing, a robust state estimation step, and a robust error covariance matrix updating step. In the following subsections, we will discuss them in detail. Note that the index i associated with the generator is neglected for simplicity but without loss of generality.

Batch-Mode Regression Form

The main idea of the UKF is to use a set of chosen deterministic sigma points for statistical information propagation. For instance, if a state estimate with mean $\widehat{\boldsymbol{x}}_{k-1|k-1} \in \mathbb{R}^{n\times 1}$ and covariance matrix $\boldsymbol{\Sigma}^{xx}_{k-1|k-1}$ is given at time step k-1, its statistics can be captured by $2n$ weighted sigma points defined as [113]

$$\boldsymbol{\chi}^{j}_{k-1|k-1} = \widehat{\boldsymbol{x}}_{k-1|k-1} \pm \left(\sqrt{n\boldsymbol{\Sigma}^{xx}_{k-1|k-1}}\right)_j, \tag{4.25}$$

with weights $w_j = 1/2n, j = 1, ..., 2n$. Then, each sigma point is propagated through the nonlinear system process model (6.20), yielding the following transformed samples

$$\boldsymbol{\chi}^{j}_{k|k-1} = \boldsymbol{f}\left(\boldsymbol{\chi}^{j}_{k-1|k-1}\right). \tag{4.26}$$

Then, the predicted state $\widehat{\boldsymbol{x}}_{k|k-1}$ and its covariance matrix $\boldsymbol{\Sigma}^{xx}_{k|k-1}$ are calculated by the weighted sample mean and sample covariance matrix of the transformed sigma points. Formally, we have

$$\widehat{\boldsymbol{x}}_{k|k-1} = \sum_{j=1}^{2n} w_j \boldsymbol{\chi}^j_{k|k-1},$$

$$\boldsymbol{\Sigma}^{xx}_{k|k-1} = \sum_{j=1}^{2n} w_j (\boldsymbol{\chi}^j_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})(\boldsymbol{\chi}^j_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})^T + \boldsymbol{Q}_k.$$

To derive the regression model, we will apply statistical linerization [54, 92] to the nonlinear measurement function around $\widehat{\boldsymbol{x}}_{k|k-1}$. The main idea of statistical linerization is to approximate the nonlinear function $\boldsymbol{\eta} = \boldsymbol{g}(\boldsymbol{x})$ around $\widehat{\boldsymbol{x}}$ statistically with $\boldsymbol{\eta} = \boldsymbol{G}\boldsymbol{x} + \boldsymbol{b} + \boldsymbol{\zeta}$, where $\boldsymbol{G} = (\boldsymbol{P}^{x\eta})^T(\boldsymbol{P}^{xx})^{-1}$; $\boldsymbol{P}^{x\eta}$ and $\boldsymbol{P}^{xx}$ are the cross-covariance matrix between $\boldsymbol{x}$ and $\boldsymbol{\eta}$, and self-covariance matrix of $\boldsymbol{x}$, respectively; $\boldsymbol{b} = \widehat{\boldsymbol{\eta}} - \boldsymbol{G}\widehat{\boldsymbol{x}}$ and $\widehat{\boldsymbol{\eta}} = \boldsymbol{g}(\widehat{\boldsymbol{x}})$; $\boldsymbol{\zeta}$ is the statistical linerization error. Following these procedures, the statistical linerization of the nonlinear measurement function $\boldsymbol{h}(\cdot)$ around $\widehat{\boldsymbol{x}}_{k|k-1}$ yields

$$\boldsymbol{z}_k = \boldsymbol{H}_k \left(\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k-1}\right) + \widehat{\boldsymbol{z}}_{k|k-1} + \boldsymbol{v}_k + \boldsymbol{\varepsilon}_k, \tag{4.27}$$

where $\boldsymbol{H}_k = (\boldsymbol{\Sigma}^{xz}_{k|k-1})^T(\boldsymbol{\Sigma}^{xx}_{k|k-1})^{-1}$; $\boldsymbol{\varepsilon}_k$ is the statistical linearization error term with zero mean and covariance matrix $\boldsymbol{L}_k = \boldsymbol{\Sigma}^{zz}_{k|k-1} - (\boldsymbol{\Sigma}^{xz}_{k|k-1})^T(\boldsymbol{\Sigma}^{xx}_{k|k-1})^{-1}\boldsymbol{\Sigma}^{xz}_{k|k-1}$, where

$$\boldsymbol{\Sigma}^{xx}_{k|k-1} = \sum_{j=1}^{2n} w_j (\boldsymbol{\chi}^j_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})(\boldsymbol{\chi}^j_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})^T + \boldsymbol{Q}_k. \tag{4.28}$$

$$\boldsymbol{\Sigma}^{zz}_{k|k-1} = \sum_{j=1}^{2n} w_j (\boldsymbol{z}^j_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1})(\boldsymbol{z}^j_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1})^T + \boldsymbol{R}_k, \tag{4.29}$$

$$\boldsymbol{\Sigma}^{xz}_{k|k-1} = \sum_{i=1}^{2n} w_j (\boldsymbol{\chi}^j_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})(\boldsymbol{z}^j_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1})^T, \tag{4.30}$$

and $\widehat{\boldsymbol{z}}_{k|k-1} = \sum_{j=1}^{2n} w_j \boldsymbol{z}^j_{k|k-1}$ is the predicted measurement vector; $\boldsymbol{z}^j_{k|k-1} = \boldsymbol{h}(\boldsymbol{\chi}^j_{k|k-1})$.

Define $\widehat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{x}_k + \boldsymbol{\Delta}_k$, where $\boldsymbol{x}_k$ is the true state vector; $\boldsymbol{\Delta}_k$ is the prediction error and $\mathbb{E}\left[\boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^T\right] = \boldsymbol{\Sigma}^{xx}_{k|k-1}$. Then, the equations associated with the state predictions and measurement functions can be reorganized into the following batch-mode regression form:

$$\begin{bmatrix} \boldsymbol{z}_k + \boldsymbol{H}_k\widehat{\boldsymbol{x}}_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1} \\ \widehat{\boldsymbol{x}}_{k|k-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_k \\ \boldsymbol{I} \end{bmatrix} \boldsymbol{x}_k + \begin{bmatrix} \boldsymbol{v}_k + \boldsymbol{\varepsilon}_k \\ \boldsymbol{\Delta}_k \end{bmatrix} \tag{4.31}$$

where $\boldsymbol{I}$ is an identity matrix. It can be further rewritten in a compact form as

$$\widetilde{\boldsymbol{z}}_k = \widetilde{\boldsymbol{H}}_k \boldsymbol{x}_k + \widetilde{\boldsymbol{e}}_k, \tag{4.32}$$

and the error covariance matrix of $\widetilde{\boldsymbol{e}}_k$ is given by

$$\boldsymbol{W}_k = \begin{bmatrix} \boldsymbol{L}_k + \boldsymbol{R}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma}^{xx}_{k|k-1} \end{bmatrix} = \boldsymbol{S}_k \boldsymbol{S}_k^T, \tag{4.33}$$

where $\boldsymbol{S}_k$ is calculated by the Cholesky decomposition technique.

Note that the state prediction errors of the batch-mode regression form should be uncorrelated before the state estimation. This can be done by pre-multiplying $\boldsymbol{S}_k^{-1}$ on both sides of (7.32), yielding

$$\boldsymbol{S}_k^{-1}\widetilde{\boldsymbol{z}}_k = \boldsymbol{S}_k^{-1}\widetilde{\boldsymbol{H}}_k \boldsymbol{x}_k + \boldsymbol{S}_k^{-1}\widetilde{\boldsymbol{e}}_k, \tag{4.34}$$

which is further organized to the compact form given by

$$\boldsymbol{y}_k = \boldsymbol{A}_k \boldsymbol{x}_k + \boldsymbol{\xi}_k, \tag{4.35}$$

where $\mathbb{E}[\boldsymbol{\xi}_k {\boldsymbol{\xi}_k}^T] = \boldsymbol{I}$.

Outlier Detection and Processing using Multiple Hypothesis Testing

As mentioned before, there are three types of outliers, i.e., observation, innovation and structural outliers. Specifically, observation outliers affect the measurements in $\boldsymbol{z}_k$; innovation outliers are usually caused by unknown system disturbances reflected in the process noise $\boldsymbol{w}_k$ or the incorrect inputs $\boldsymbol{u}_k$; structural outliers are induced by parameter errors of the generator, exciter, governor, etc. It should be noted that structural outliers will affect all the predicted states in this decentralized scheme-based state estimation due to the strong dependence among each predicted state. On the other hand, if the voltage phasor is taken as input vector and has gross errors, it will trigger the following cascading chain: $V\angle\theta \rightarrow V_R \rightarrow V_F \rightarrow E_{fd} \rightarrow E'_q \rightarrow E'_d$, which makes all the the predicted states incorrect, yielding innovation outliers. If the current phasor is taken as input vector and has gross errors, the cascading chain is as follows: $I\angle\phi \rightarrow I_d \rightarrow I_q \rightarrow E'_q \rightarrow E'_d \rightarrow E_{fd} \rightarrow V_F \rightarrow V_R \rightarrow P_e \rightarrow \omega \rightarrow T_M$, which makes all the other predicted states incorrect, yielding innovation outliers. Based on the above analysis, it is clear that there are three hypotheses corresponding to each outlier scenario plus one additional hypothesis to determine whether there are outliers or not. Therefore, we need to validate each hypothesis and determine which type of outliers occurs.

First of all, we need to check the following hypothesis:

$$\text{Hypothesis } 1 : \begin{cases} \mathcal{H}_0 : \text{no outliers} \\ \mathcal{H}_1 : \text{occurrence of outliers} \end{cases}. \tag{4.36}$$

To this end, we propose to apply the projection statistics (PS) at each measurement sample to a 2-dimensional matrix $\boldsymbol{Z}_1$ that contains time-correlated samples of the innovations and of the predicted state variables. Specifically, we have

$$\boldsymbol{Z}_1 = \begin{bmatrix} \boldsymbol{z}_{k-1} - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k-1|k-2}) & \boldsymbol{z}_k - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k|k-1}) \\ \widehat{\boldsymbol{x}}_{k-1|k-2} & \widehat{\boldsymbol{x}}_{k|k-1} \end{bmatrix}, \tag{4.37}$$

where $\boldsymbol{z}_{k-1} - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k-1|k-2})$ and $\boldsymbol{z}_k - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k|k-1})$ are the innovation vectors while $\widehat{\boldsymbol{x}}_{k-1|k-2}$ and $\widehat{\boldsymbol{x}}_{k|k-1}$ are the predicted state vectors at time instants k-1 and k, respectively. The choice of $\boldsymbol{Z}_1$ is motivated by the fact that the innovation vector and the predicted state vector are time series samples of power system responses and have strong temporal correlations. If outliers occur, this relationship is violated. Thus, by checking this statistical property of the matrix $\boldsymbol{Z}_1$, we are able to detect outliers. The PS is defined as [37, 92]:

$$PS_j = \max_{\|\boldsymbol{\ell}\|=1} \frac{\left|\boldsymbol{l}_j^T\boldsymbol{\ell} - med_i\left(\boldsymbol{l}_i^T\boldsymbol{\ell}\right)\right|}{1.4826\ med_\kappa\left|\boldsymbol{l}_\kappa^T\boldsymbol{\ell} - med_i\left(\boldsymbol{l}_i^T\boldsymbol{\ell}\right)\right|}, \tag{4.38}$$

for $i, j, \kappa = 1, 2, ..., m+n$, where n and m are the number of state variables and measurements taken as outputs, respectively; $\boldsymbol{l}_j^T$, $\boldsymbol{l}_i^T$ and $\boldsymbol{l}_\kappa^T$ are the jth, ith and κth row vector of $\boldsymbol{Z}_1$, respectively. The PS values of the predictions and of the innovations are separately calculated. Extensive Monte Carlo simulations revealed that the PS values follow a chi-square distribution with degree of freedom 2. Thus, if we define the set $\Pi_1 = \{\mathrm{PS}_i > \tau = \chi^2_{2,0.975}, i = 1, ..., m+n\}$, then Hypothesis 1 can be rewritten as

$$\text{Hypothesis } 1 : \begin{cases} \mathcal{H}_0 : |\Pi_1| = 0 \\ \mathcal{H}_1 : 1 \leq |\Pi_1| \leq m+n \end{cases}, \tag{4.39}$$

where $|\Pi_1|$ represents the cardinality of the set Π_1; $\mathcal{H}_1$ and $\mathcal{H}_0$ correspond to occurrence of outliers and no outliers, respectively; PS_i denotes the ith PS value; τ is determined according to the chi-square distribution at a significance level 97.5%. If $\mathcal{H}_0$ is accepted, the robust regression shown in the next subsection is performed; otherwise, the outlier detection rule is applied and additional procedures are executed to distinguish between them. It is worth pointing out that we can apply the PS to higher dimensional samples, but we found that applying them to 2-dimensional time series data is sufficient to check if the statistical correlations of the innovation vector and the predicted state vector are broken up by outliers. Note that an unnecessary higher dimensional samples will induce more computing burden of the PS algorithm.

For any generator with n state variables and m outputs, it is noticed that if the number of bad measurements flagged by the PS is less than m, only observation outliers occur. This is because both innovation and structural outliers will affect all n predicted states, and as a result, all m measurements will be flagged as outliers by the PS. The latter is due to the fact that the predicted states dominate the data of the matrix $\boldsymbol{Z}_1 \in \mathbb{R}^{(m+n)\times 2}$ and if they are wrong, the PS method breaks down, that is, it identifies all the good measurements that are associated with the innovation vectors as outliers. However in the case where all the m measurements are outliers, then all bad measurements will be flagged as outliers as well. Therefore, extra test should be performed to differentiate observation outliers that are in extreme scenario with innovation and structural outliers. To this end, we propose to apply the PS to the matrix $\boldsymbol{Z}_2 \in \mathbb{R}^{m\times 3}$ that includes 3 dimensional consecutive samples, that is,

$$\boldsymbol{Z}_2 = \begin{bmatrix} \boldsymbol{z}_{k-2} & \boldsymbol{z}_{k-1} & \boldsymbol{z}_k \end{bmatrix}, \tag{4.40}$$

where the received measurements $\boldsymbol{z}_{k-2}$ and $\boldsymbol{z}_{k-1}$ must be replaced by their corresponding estimated values if they have outliers. Our choice of matrix $\boldsymbol{Z}_2$ is motivated by two facts: i) when all m measurements are outliers at time instant k, two-dimensional consecutive samples are not sufficient to detect outliers due to the lack of redundancy; to address this issue, we propose to include one more dimensional sample, which increases the global redundancy, and thus, the detectability of the outliers; ii) time-correlated measurements are independent from the predicted states and consequently innovation/structural outliers do not affect the detection of observation outliers. Extensive Monte Carlo simulations showed that the PS values calculated from $\boldsymbol{Z}_2$ follow a chi-square distribution with degree of freedom 3. Let us now define the set $\Pi_2 = \{\mathrm{PS}_i > \chi^2_{3,0.975}, i = 1, ..., m\}$. The hypothesis 2 used to distinguish observation outliers with innovation and structural outliers is thus expressed as

$$\text{Hypothesis 2}: \begin{cases} \mathcal{H}_0 : |\Pi_1| - |\Pi_2| \geq 0 \\ \mathcal{H}_1 : |\Pi_2| = m \end{cases}, \tag{4.41}$$

where hypothesis $\mathcal{H}_1$ corresponds to the scenario that all m measurements are outliers while $\mathcal{H}_0$ corresponds to the occurrence of innovation and/or structural outliers. If the hypothesis $\mathcal{H}_1$ is confirmed, we propose to replace all m measurements with their previous estimated values at time instant $k-1$. Then, Hypothesis 1 in (4.39) is tested again and if all newly replaced m measurements are flagged as outliers, the occurrence of innovation and/or structural outliers is double confirmed (note that this indicates the possibility of the occurrence of three types of outliers simultaneously); otherwise, only the extreme outlier scenario associated with the observations presents. To further distinguish innovation and structural outliers, we propose to replace the system inputs $\boldsymbol{u}$ by its estimated values at the previous time instant and obtain new predicted state vector $\widehat{\boldsymbol{x}}^{new}_{k|k-1}$, yielding a new matrix $\boldsymbol{Z}_3$ for outlier detection given by

$$\boldsymbol{Z}_3 = \begin{bmatrix} \boldsymbol{z}_{k-1} - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k-1|k-2}) & \boldsymbol{z}_k - \boldsymbol{h}(\widehat{\boldsymbol{x}}^{new}_{k|k-1}) \\ \widehat{\boldsymbol{x}}_{k-1|k-2} & \widehat{\boldsymbol{x}}^{new}_{k|k-1} \end{bmatrix}, \tag{4.42}$$

where $\boldsymbol{z}_{k-1}$ and $\boldsymbol{z}_k$ must be replaced by their estimated values if they were detected as outliers in the former steps. Next, define set $\Pi_3 = \{\mathrm{PS}_i > \chi^2_{2,0.975}, i = 1, ..., m+n\}$, the hypothesis 3 used to differentiate innovation and structural outliers is expressed as

$$\text{Hypothesis 3}: \begin{cases} \mathcal{H}_0 : |\Pi_3| = 0 \\ \mathcal{H}_1 : |\Pi_3| = m \end{cases}. \tag{4.43}$$

If the hypothesis $\mathcal{H}_1$ is validated, structural outliers exist; otherwise innovation outliers has occurred.

After the detection and differentiation of outliers, we propose to process them using the following rules:

- *Rule 1*: if only observation outliers occur, they are downweighted via the weights given by

$$\varpi_i = \min\left(1, d^2/PS_i^2\right), \tag{4.44}$$

where the parameter d is set as 1.5 to yield good statistical efficiency at Gaussian distribution and other thick-tailed non-Gaussian distributions [38, 92];

- *Rule 2*: if only innovation outliers occur, the inputs are replaced by their estimated values at the last time instant and new predicted states are obtained;
- *Rule 3*: if only structural outliers occur, predicted states are replaced by filtered state variables at the last time instant;
- *Rule 4*: if both observation and innovation outliers occur, the inputs are replaced by their estimated values at the last time instant and new predicted states are obtained, while observations are processed according to rule 1; if both observation and structural outliers occur, rules 1 and 3 are applied;
- *Rule 5*: If three types of outliers occur simultaneously, predicted states are replaced by filtered state variables at the last time instant; bad observations are replaced by their estimated values at the last time instant.

Remark 1: By using our multiple hypothesis testing approach, we are able to distinguish between three types of outliers. If structural outliers occur, additional actions are needed to locate incorrect parameters of the generator, requiring that a model calibration need to be executed. This will be proposed as our future work.

Remark 2: The more types of outliers are present, the more time is required for executing the multiple hypothesis testing rules. However, the computing time is not prohibitively large for our robust DSE as the decentralized scheme enables the outlier detection and processing to be quickly executed at each local generator. This will be shown in the simulation results section.

Robust Regression

After the outlier detection and processing, corresponding weights $\varpi_i, i = 1, ..., m+n$ are assigned to the measurements and the predicted states via (5.20); then the robust regression is performed by using a robust GM-estimator that minimizes the following objective function:

$$J(\boldsymbol{x}_k) = \sum_{i=1}^{m+n} \varpi_i^2 \rho(r_{S_i}), \tag{4.45}$$

where $r_{S_i} = r_i / s\varpi_i$ is the standardized residual; $r_i = y_i - \boldsymbol{a}_i^T \widehat{\boldsymbol{x}}$ is the residual, where $\boldsymbol{a}_i^T$ is the ith row vector of the matrix $\boldsymbol{A}_k$; $s = 1.4826 \cdot b_m \cdot \text{median}_i |r_i|$ is the robust scale estimate; b_m is a correction factor; $\rho(\cdot)$ is the convex Huber-ρ function [38].

To minimize (5.21), the following necessary condition must be satisfied

$$\frac{\partial J\left(\boldsymbol{x}_k\right)}{\partial \boldsymbol{x}_k}=\sum_{i=1}^{l}-\frac{\varpi_i \boldsymbol{a}_i}{s}\psi\left(r_{S_i}\right)=\mathbf{0}, \tag{4.46}$$

where $\psi\left(r_{S_i}\right)=\partial\rho\left(r_{S_i}\right)/\partial r_{S_i}$. By dividing and multiplying the standardized residual r_{S_i} to both sides of (5.23) and putting it in a matrix form, we get

$$\boldsymbol{A}_k^T\boldsymbol{\Lambda}\left(\boldsymbol{y}_k-\boldsymbol{A}_k\boldsymbol{x}_k\right)=\mathbf{0}, \tag{4.47}$$

where $\boldsymbol{\Lambda}$ =diag$(q\left(r_{S_i}\right))$ and $q\left(r_{S_i}\right)=\psi\left(r_{S_i}\right)/r_{S_i}$. By using the IRLS algorithm [37], the state estimates at the j iteration can be calculated

$$\Delta\widehat{\boldsymbol{x}}_{k|k}^{(j+1)}=\left(\boldsymbol{A}_k^T\boldsymbol{\Lambda}^{(j)}\boldsymbol{A}_k\right)^{-1}\boldsymbol{A}_k^T\boldsymbol{\Lambda}^{(j)}\boldsymbol{y}_k, \tag{4.48}$$

where $\Delta\widehat{\boldsymbol{x}}_{k|k}^{(j+1)}=\widehat{\boldsymbol{x}}_{k|k}^{(j+1)}-\widehat{\boldsymbol{x}}_{k|k}^{(j)}$. The algorithm converges when $\left\|\Delta\widehat{\boldsymbol{x}}_{k|k}^{(j+1)}\right\|_\infty\leq 10^{-2}$.

Update the Error Covariance Matrix

Following our previous work, the estimation error covariance matrix $\boldsymbol{\Sigma}_{k|k}^{xx}$ is updated using the total influence function-based approach [38, 79]. Formally, we have

$$\boldsymbol{\Sigma}_{k|k}^{xx}=\frac{\mathbb{E}_F\left[\psi^2\left(r_{S_i}\right)\right]}{\left\{\mathbb{E}_F\left[\psi'\left(r_{S_i}\right)\right]\right\}^2}\left(\boldsymbol{A}_k^T\boldsymbol{A}_k\right)^{-1}\left(\boldsymbol{A}_k^T\boldsymbol{Q}_\varpi\boldsymbol{A}_k\right)\left(\boldsymbol{A}_k^T\boldsymbol{A}_k\right)^{-1} \tag{4.49}$$

where $\boldsymbol{Q}_\varpi=diag\left(\varpi_i^2\right)$.

4.4 Numerical Results

Extensive simulation results of several scenarios are carried out on the IEEE 39-bus system to test the effectiveness and robustness of the proposed robust DSE. In the case studies, Line 15-16 is tripped at $t=0.5$s to simulate system disturbance. The time-domain simulation results are taken as the true values. The collections of simulated voltage phasor, current phasor and frequency at each generator's terminal bus are treated as PMU measurements. A random Gaussian variable with zero mean and covariance matrix equal to $10^{-6}\boldsymbol{I}$ is assumed for system process noise. The generator model assumed for transient simulation is the detailed two-axis generator model, whose parameter values are taken from [70]. The maximum iteration of the IRLS algorithm is 20. The break point of the Huber cost function is set to 2. The root-mean-squared error (RMSE) of all estimated generator state variables is used as the overall performance index while the estimated rotor angle, rotor speed, d- and q-axis voltages of Generator 5 are taken for illustration due to the limitation of space. When the metered

generator terminal voltage phasor is used as model input while the current phasor and frequency are treated as outputs/measurements, the decentralized DSE approach will be called the DUKFV. By contrast, when the metered generator terminal current phasor is used as model input while the voltage phasor and frequency are treated as outputs/measurements, the decentralized DSE approach will be called the DUKFI. Its robust version, which is our proposed robust DSE, will be termed the RDUKFI. The choice of developing a robust version of the DUKFI instead of the DUKFV is motivated by the fact that the former achieves much higher statistical efficiency than the latter in the absence of outliers regardless of the level of measurement noise. This is confirmed by the results shown in Figs. 6.1-6.2. In other words, if the DUKFI is made robust against outliers, it will outperform the robust DUKFV in presence of both measurement noise and outliers.

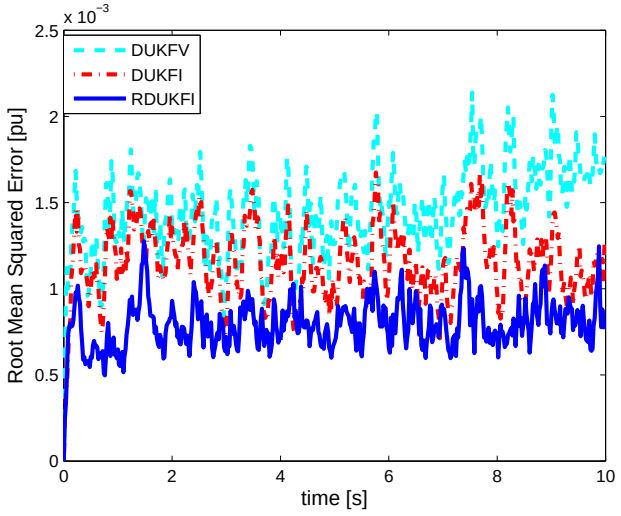


Figure 4.2: Root-mean-squared error of the DUKFV, the DUKFI and the RDUKFI in the presence of normal measurement noise.

4.4.1 Case 1: Sensitivity to Measurement Noise

Measurement noise with different variances is implemented to test the sensitivity of each method. Specifically, normal and large noises are considered; they are assumed to be distributed with zero means and covariance matrices $10^{-6}\boldsymbol{I}$ and $10^{-4}\boldsymbol{I}$, respectively. The test results are displayed in Figs. 6.1-6.2. It can be concluded from these two figures that our RDUKFI outperforms the DUKFV and the DUKFI in presence of both normal and large measurement noises; the DUKFI that uses our model decoupling approach achieves much higher statistical efficiency than the DUKFV. In addition, it is observed that with the increased noise level, all three methods have increased estimation error, which is expected. This is because given certain level of measurement redundancy, any estimator will have increased estimation error with the increase of measurement noise; the higher measurement redun-

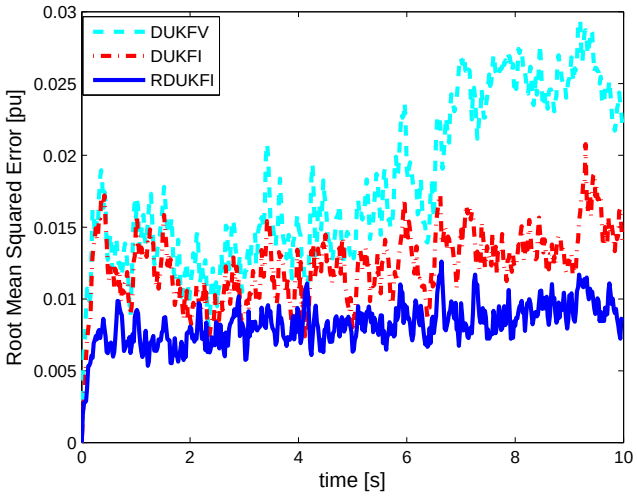


Figure 4.3: Root-mean-squared error of the DUKFV, the DUKFI and the RDUKFI in the presence of large measurement noise.

dancy, the better capability of an estimator to filter out noise. Nevertheless, our RDUKFI is least sensitive to the increase of noise level among three methods.

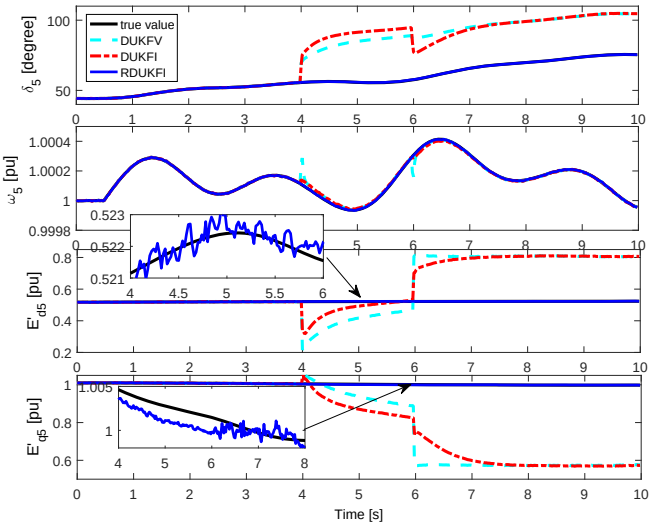


Figure 4.4: Estimated rotor angle, rotor speed, d- and q-axis voltages of Generator 5 by the DUKFV, the DUKFI and the RDUKFI in the presence of observation outlier, where the measured voltage magnitude and angle of Generator 5 is contaminated with 20% error from t=4s to t=6s.

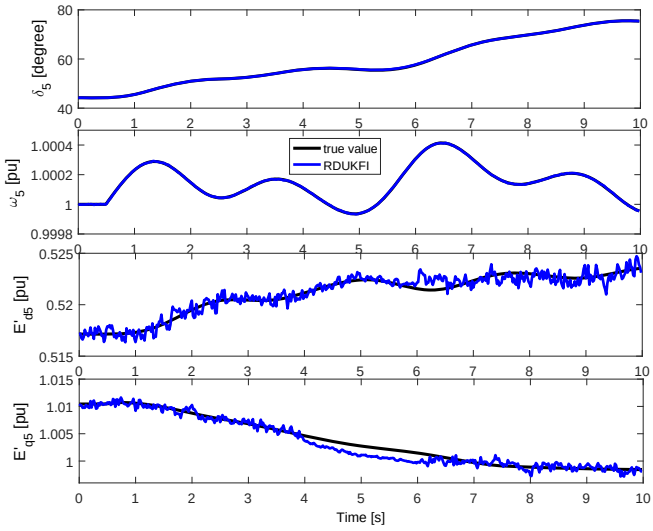


Figure 4.5: Zooming in the Fig. 4.4 for better illustration of the proposed RDUKFI from t=0 to t=10 s.

4.4.2 Case 2: Robustness to Observation Outliers

Due to imperfect phasor synchronization, the saturation of metering current transformers or by metering coupling capacitor voltage transformers (CCVTs), to name a few, gross errors can occur in the PMU measurements [27, 38]. To test the robustness of three methods to observation outliers, the measured voltage magnitude and angle of Generator 5 is contaminated with 20% error from t=4s to t=6s. The results are shown in Fig. 4.4 and its zoomed in version is displayed in Fig. 4.5.

We can find that the estimated generator state variables provided by the DUKFV and the DUKFI are significantly biased in the presence of observation outliers due to their lack of robustness. However, our RDUKFI that uses multiple hypothesis testing is able to detect these two observation outliers and bound their influence, yielding excellent tracking performance. Note that during the multiple hypothesis testing process, the test results on hypothesis 1 indicate that 2 observations are flagged as outliers, and consequently the occurrence of only observations outliers is declared. Then, the observation outliers are downweighted and suppressed by the GM-estimator.

4.4.3 Case 3: Robustness to Innovation Outliers

The system inputs $\boldsymbol{u}$ can be incorrect due to unknown disturbances or erroneous metered values by PMUs, yielding innovation outliers. To test the robustness of the proposed method to this type of outliers, the system inputs are contaminated with 20% error from t=2s to t=4s. Comparison results of the DUKFV, the DUKFI and the RDUKFI are displayed in Fig. 4.6. This figure is further zoomed in and is shown in Fig. 4.7 for better illustration.

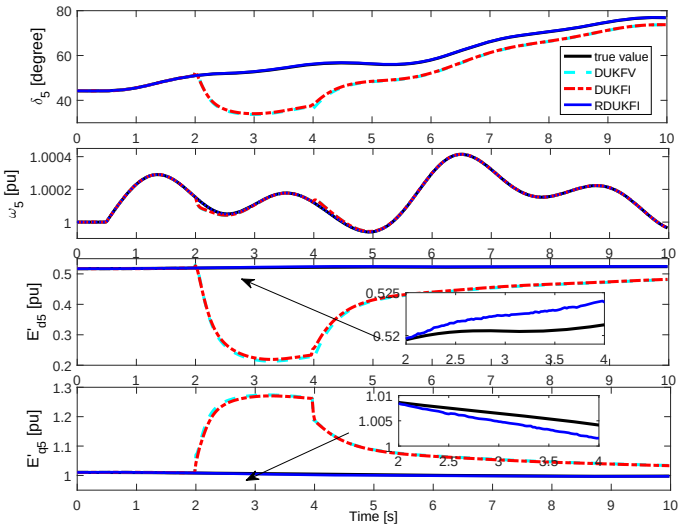


Figure 4.6: Estimated rotor angle, rotor speed, d- and q-axis voltages of Generator 5 by the DUKFV, the DUKFI and the RDUKFI in the presence of innovation outlier, where the model inputs are contaminated with 20% error from t=2s to t=4s.

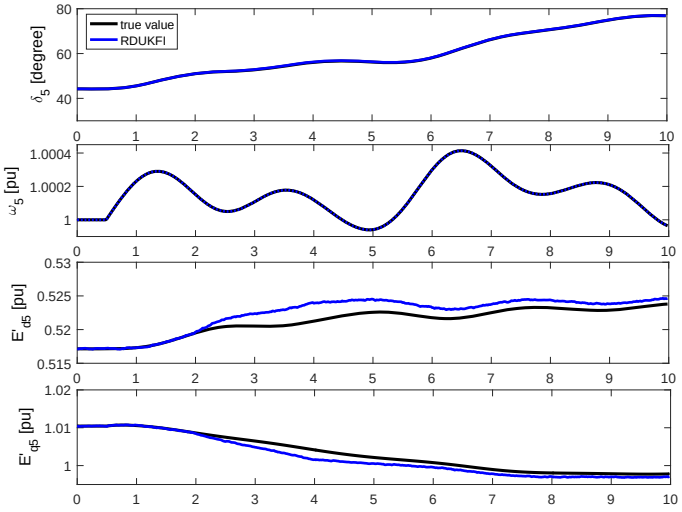


Figure 4.7: Zooming in the Fig. 4.6 for better illustration of the proposed RDUKFI from t=0 to t=10 s.

It can be concluded that the DUKFV and the DUKFI are not robust to innovation outliers as they have very large estimation errors. By contrast, in the multiple hypothesis testing process, Hypothesis 1 indicates that all outputs/measurements are flagged by PS as outliers; then Hypothesis 2 is tested and its result indicates the occurrence of innovation and/or structural outliers. After replacing the model inputs by its estimated value at the previous time instant, new predicted state variables are obtained and Hypothesis 3 is then tested.

The result confirms the occurrence of innovation outliers. As a result, the newly predicted state variables associated with the measurements are used for robust estimation by the GM-estimator, yielding negligible biased state estimates.

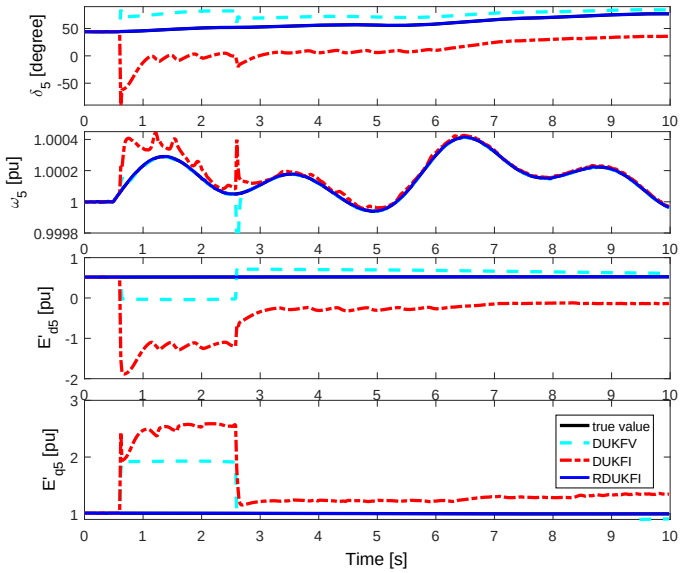


Figure 4.8: Estimated rotor angle, rotor speed, d- and q-axis voltages of Generator 5 by the DUKFV, the DUKFI and the RDUKFI in the presence of structural outlier, where the d- and q-axis transient reactance of Generator 5 are contaminated with 20% Gaussian random errors due to the effect of saturation on generator inductances from $t = 0.6$s to $t = 2.6$s.

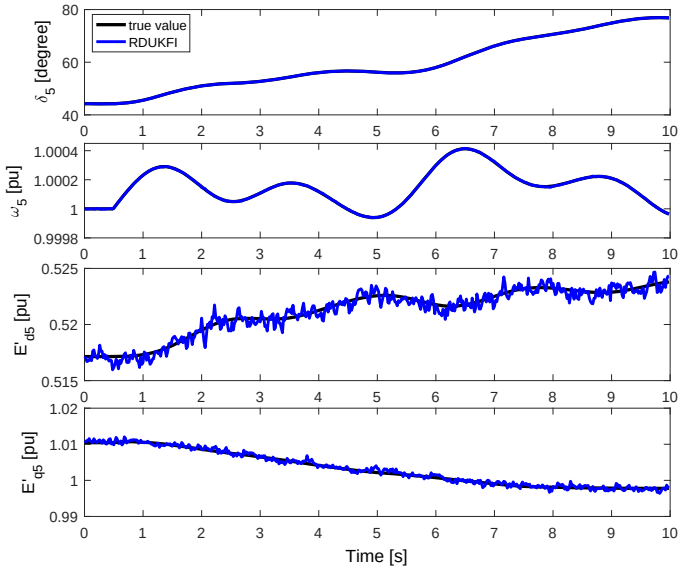


Figure 4.9: Zooming in the Fig. 4.8 for better illustration of the proposed RDUKFI from t=0 to t=10 s.

4.4.4 Case 4: Robustness to Structural Outliers

Due to aging processes, variations of the machine temperature during its operation, the effect of saturation on generator inductances, etc., the parameters of synchronous generator can change significantly while remaining unknown to the operator of the power plant. In this chapter, we consider the case where following a system disturbance, there exist saturations on the inductances of Generator 5 from t= 0.6s to t = 2.6s, yielding structural outliers. Note that the saturation level of a synchronous machine does not change instantaneously; in fact, it has an oscillatory behavior for post-disturbances. Thus, to simulate the oscillatory behavior during the saturation period, 20% random errors simulated by Gaussian distribution are added to the d- and q-axis transient reactance. Estimated rotor angle, rotor speed, d- and q-axis voltages of Generator 5 by the DUKFV, the DUKFI and the RDUKFI are displayed in Fig. 4.8 associated with its zoomed in version shown in Fig. 4.9. We observe from this figure that both the DUKFV and the DUKFI yield large biases on the state estimates in the presence of structural outliers. The DUKFV is less sensitive to this type of outliers than the DUKFI and spends less time re-tracking the system true states. By contrast, thanks to the multiple hypothesis testing and the robust GM-estimator, our RDUKFI can bound the influence of the structural outliers, yielding very small biases on the state estimates when outliers occur.

4.4.5 Computational Efficiency

The average computing times of the DUKFV, the DUKFI and the RDUKFI for Cases 1-4 in each PMU scan are tabulated in Table. 4.1. All the tests are performed on a PC with Intel Core i5, 2.50 GHz, 8GB of RAM. It is observed from Table. 4.1 that all three methods have comparative computational efficiency and their computing times are much lower than the PMU scan rate, which is 16.7ms for 60 sample/s. The DUKFI is the most computational efficient approach, followed by the DUKFV. The proposed RDUKFI is the most time consuming approach among them, which is expected as it needs additional multiple hypothesis testing and robust regression procedures. However, these two procedures just impose slight computing burden to the proposed method since the RDUKFI is implemented at each local generator and very fast to execute. A closer look at the computing time of the proposed method under different conditions reveals that it requires more time in Cases 2-4 than in Case 1. This is because more time is required to test multiple hypotheses when outliers occur. In addition, it should be noted that we need to check all the three hypotheses so that innovation and structural outliers can be distinguished while only two hypothesis are required for the detection of observation outliers. Therefore, the computing times of Cases 3 and 4 are larger than that of Case 2. Finally, the proposed RDUKFI is implemented locally and independent of other generators and other system wide-area measurements, it is thus suitable for large-scale power system online applications.

Table 4.1: Average Computing Times of the Three Decentralized Kalman Filters at Each PMU Scan

Cases	DUKFV	DUKFI	RDUKFI
Case 1	0.282ms	0.296ms	0.85ms
Case 2	0.396ms	0.286ms	1.07ms
Case 3	0.473ms	0.295ms	1.31ms
Case 4	0.275ms	0.264ms	1.35ms

4.4.6 Discussion of Practical Industry Application

The proposed robust DSE is implemented at each generator substation using generator dynamic models as well as local PMU measurements. Since generator model and parameters, measurement devices and local computer have been there already, no hardware modification has to be carried out. The only requirement is to add an additional software block at the local computers that performs the dynamic simulation using measured current phasors as inputs and combine their simulation outputs with the measured voltage phasors and frequency for dynamic state estimation. It should be noted that the results of the robust DSE can be used for power system online monitoring, control and protection [91,93,94]. For example, a recent patent by the vendor ABB demonstrated the value of the DSE for online monitoring [93]. The effectiveness of the DSE assisted adaptive protection scheme has been tested in real system [94]. Thus, although some costs are needed to change the existing code for the DSE implementation, its benefits for system online monitoring, control, and protection are promising.

4.5 Conclusion and Future Work

In this chapter, a fast RDUKFI-based decentralized DSE is proposed. It is implemented at each local generator and independent of the other generators, system parameters, and thus suitable for large-scale power system online applications. Furthermore, our RDUKFI is able to detect, distinguish and suppress the observation, innovation and structural outliers while achieving high statistical efficiency in presence of both normal and large measurement noise. We propose projection statistics-based multiple hypothesis testing approach for outlier detection and processing. The identified outliers are suppressed by a generalized maximum-likelihood type (GM)-estimator. Numerical results carried out on the IEEE 39-bus system demonstrate the effectiveness and robustness of the proposed method.

It should be noted that the proposed robust DSE is general in that it can be easily extended to estimate the dynamic states of other power system components, such as dynamic loads, wind generators and various types of higher order exciters. In addition, the 9th-order two-axis model with DC1A exciter is one of the recommended models by IEEE for transient stability analysis [95]. Thus, we believe that our results have practical meanings.

Chapter 5

Robust Frequency Divider for Power System Online Monitoring and Control

5.1 Introduction

With the increasing penetration of renewable energy-based generations, the total inertia of the synchronous power system is reduced significantly. As a result, the traditional capacity-based requirements for primary reserve definitions may not be able to satisfy the frequency RoCoF and nadir limits [96, 97]. To tackle this potential frequency instability issue, it is expected by the transmission system operators (TSOs) that the wind farms [98], flexible loads [100], and energy storage devices [101] would provide frequency regulations. However, to enable effective frequency regulations, reliable and accurate knowledge of the local bus frequency is a prerequisite.

In the literature, the numerical derivative of the voltage phase angle provided by Phasor Measurement Unit (PMU) through a washout filter is used to define the local bus frequency [102, 103]. However, as shown by Radman *et al.* [104], it may produce physically implausible spikes in frequency due to the numerical derivatives and consequently may exhibit instabilities if controls are taken based on them. An alternative way to obtain local bus frequency is through the phase-locked loop (PLL) technique [105]. But it may be unreliable in presence of large step-input speed changes, and has problems in presence of harmonics, unbalance, etc. Furthermore, only a few load buses/substations have PMUs or PLLs installed, which may prevent the system from taking full advantages of local frequency controls. Finally, measurements provided by PMUs and PLLs are always subject to noise or even gross errors, communication losses, etc [99]. For example, it is shown in [100] that the measurement noise affects the performance of the frequency regulator significantly, not

to mention the gross errors, measurement losses, etc. To address these issues, a model dependent analytical expression of bus frequency is proposed in [106] assuming comprehensive and accurate models of the system. Milano and Ortega [107] improved that approach and proposed a transient stability simulation-based frequency divider. The main idea underlying this method is to solve a steady-state boundary value problem, where the boundary conditions are given by synchronous generator rotor speeds. However, both methods assume accurate power system dynamic models for time-domain simulations, which is difficult to achieve in practice. In addition, the time-domain simulations of large-scale power systems are computational demanding, which may prevent them for the online control applications.

This chapter proposes a robust frequency divider (RFD) for online bus frequency estimation [108]. Our RFD is not dependent on load models, and the knowledge of swing equation parameters, transmission system line parameters and local PMU measurements is sufficient. In addition, it is able to filter out measurement noise, suppress gross measurement errors, handle cyber attacks as well as measurement losses. To develop the proposed RFD, it is observed that to obtain an accurate estimation of bus frequency, accurate generator rotor speeds are required. In the meantime, complicated generator and load models should be avoided. To this end, we propose to divide RFD into two subproblems, namely the decentralized estimation of local generator rotor speed using local measurements and the centralized bus frequency estimation. To address the first problem, we decouple each generator from the rest of the power system by treating metered real power injection as model inputs and the frequency measurements provided by PMU, local meter devices or Frequency monitoring Network devices [109] as outputs. As a result, only the swing equation is required for rotor speed estimation and no detailed generator model is assumed. Since the local measurements can be subject to data quality issue when implementing an estimator, a robust unscented Kalman filter-based dynamic state estimator is proposed. Next, the local estimates are transmitted to the control center for bus frequency estimations, yielding two benefits: 1) the requirement of communication bandwidth is decreased notably as only estimated rotor speeds are communicated instead of voltage and current phasors; 2) the wide-area generator rotor angles are available for operator to achieve better situational awareness and carry out other applications, such as oscillatory modes monitoring, rotor angle stability analysis, etc. Last but not the least, thanks to the decentralized and centralized estimation scheme, the proposed method is suitable for designing controllers of very large-scale power systems.

The remainder of the paper is organized as follows. Section 5.2 presents the problem formulation. Section 5.3 describes the proposed RFD in detail and Section 5.4 shows and analyzes the simulation results. Finally, Section 5.5 concludes the paper.

5.2 Problem Formulation

In this section, the analytical relationship between bus frequency and generator rotor speeds will be presented first; then the limitations of this approach will be discussed thoroughly,

and finally the problem statement will be declared.

5.2.1 Analytical Relationship between Bus Frequency and Generator Rotor Speeds

When a disturbance occurs, such as transmission line faults, load shedding or generator tripping, power mismatch appears between the mechanical torque and electrical power at the generator terminal buses. As a consequence, the generator rotor speeds will deviate from their nominal values. To re-synchronize generator with the rest of the power system, an increase or a decrease in the rotor speed is actuated, which causes rotor angle oscillations as well. Due to such oscillations, the voltage phase angles of the buses that are adjacent to generators will encounter changes, which in turn causes a power mismatch. In this way, the electro-mechanical oscillations will be propagated throughout the entire power system with limited speed. Since we are interested in electro-mechanical oscillations and the propagation speed of such oscillations is much lower than that of the wave, the transient effects of wave propagation are therefore neglected. Based on the analysis above, it is clear that the spatial variations of the system frequency are characterized by synchronous generator rotor speeds. Those frequency variations at each bus of the system are of vital importance for designing local controllers to enhance the frequency regulation capability of a power system with high penetration renewable energy integrations.

Note that electro-mechanical oscillations can be characterized by the magnitude and phase angle modulations of voltages and currents as well since they are corresponding to the movement of rotors of electric machines around the synchronous speed [110, 111]. Thus, to estimate bus frequency of a transmission system, we first need to analyze the relationship of the voltage or current phasors between generators and system buses. This relationship can be expressed by the balanced current injection formula shown as follows:

$$\begin{bmatrix} \boldsymbol{I}_G \\ \boldsymbol{I}_B \end{bmatrix} = \begin{bmatrix} \boldsymbol{Y}_{GG} & \boldsymbol{Y}_{GB} \\ \boldsymbol{Y}_{BG} & \boldsymbol{Y}_{BB} + \boldsymbol{Y}_{B0} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_G \\ \boldsymbol{V}_B \end{bmatrix}, \tag{5.1}$$

where $\boldsymbol{I}_G$ are generator current injections; $\boldsymbol{V}_G$ are generator internal electromotive forces (emfs); $\boldsymbol{I}_B$ and $\boldsymbol{V}_B$ are current and voltage injections of the network buses, respectively; $\boldsymbol{Y}_{BB}$ is the power network admittance matrix; $\boldsymbol{Y}_{GG}$, $\boldsymbol{Y}_{GB}$ and $\boldsymbol{Y}_{BG}$ are admittance matrices calculated by including the internal impedances of the synchronous generators; $\boldsymbol{Y}_{B0}$ is a diagonal matrix, which takes into account the internal impedances of synchronous generators at the generator buses. Since the load current injections are negligible compared with that of the synchronous generators [84, 107], (5.1) can be rewritten as:

$$\begin{bmatrix} \boldsymbol{I}_G \\ \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} \boldsymbol{Y}_{GG} & \boldsymbol{Y}_{GB} \\ \boldsymbol{Y}_{BG} & \boldsymbol{Y}_{BB} + \boldsymbol{Y}_{B0} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_G \\ \boldsymbol{V}_B \end{bmatrix}. \tag{5.2}$$

By taking simple algebraic operations on the second row of (5.2), we can derive the relation-

ship between bus voltage vector $\boldsymbol{V}_B$ and the emfs of generators as follows:

$$\boldsymbol{V}_B = -(\boldsymbol{Y}_{BB} + \boldsymbol{Y}_{B0})^{-1}\boldsymbol{Y}_{BG}\boldsymbol{V}_G = \boldsymbol{D}\boldsymbol{V}_G. \tag{5.3}$$

Taking time derivatives on both sides of (5.3) in rotating reference frame, i.e., dq frame, we get

$$\frac{d\boldsymbol{V}_B}{dt} + j\omega_0\boldsymbol{V}_B = \boldsymbol{D}\cdot\frac{d\boldsymbol{V}_G}{dt} + j\omega_0\boldsymbol{D}\boldsymbol{V}_G, \tag{5.4}$$

where ω_0 is the nominal rotor speed. For more details of deriving (5.4), please see [107]. Define $\Delta\boldsymbol{\omega}_B = \boldsymbol{\omega}_B - \boldsymbol{\omega}_0$, $\Delta\boldsymbol{\omega}_G = \boldsymbol{\omega}_G - \boldsymbol{\omega}_0$, the analytical relationship between bus frequency and generator rotor speeds can be derived from (5.4), which is expressed as follows [107]:

$$\boldsymbol{\omega}_B = \boldsymbol{\omega}_0 + \boldsymbol{D}\left(\boldsymbol{\omega}_G - \boldsymbol{\omega}_0\right). \tag{5.5}$$

It can be observed from (5.5) that the bus frequency is correlated with each synchronous generator in the system, but the degree of participation of each generator on bus frequency is determined by the transmission parameters. To speed up the calculation of bus frequency from (5.5) without a relevant loss of accuracy, the conductances of transmission lines utilized to calculate $\boldsymbol{D}$ can be neglected [107].

Remark: Although the matrix $\boldsymbol{D}$ depends on the frequency, the consideration of this factor has negligible impact on the frequency divider. Therefore, transmission lines are lumped and assumed to have constant impedances. This assumption is consistent with the standard model for transient stability analysis.

5.2.2 Limitations and Problem Statement

When implementing (5.5) to estimate bus frequency, there are two possible ways: i) the dynamic simulation based-approach and ii) the measurement-based approach. In the former approach, the transient stability program is used to obtain the rotor speed of each generator, followed by the calculations of the bus frequencies through (5.5). However, to obtain good time-domain simulation results, accurate and detailed generator and load models are required, which may be difficult to achieve in practice. In addition, the time-domain simulations of large-scale power systems are computational demanding, which may not be suitable for the online control applications. By contrast, the measurement-based approach does not have such issues, but the metered generator rotor speeds and frequencies by PMUs or Frequency monitoring Network devices are assumed to be of high quality. However, this assumption may not hold true for practical power systems as the PMU measurements are usually subject to noise or even impulsive noise, gross errors, cyber attacks, communication losses, etc. Under those conditions, this approach will produce significantly biased results and subsequently the control actions based on them may make the system even worse. Furthermore, it should be noted that the frequency of each bus is correlated with most generators,

thus if just one rotor speed measurement is corrupted, its error may propagate to many other bus frequency estimations. It is thus indispensable to make sure that every generator rotor speed is of good accuracy.

Problem statement: given a limited number of PMUs installed at the terminal bus of each generator, a robust frequency divider is developed to address the data quality issues of PMU measurements; in the meantime, it should be model independent so as to mitigate the strong assumptions on the generator and load models, and finally, it should be fast to calculate and suitable for large-scale system online control applications. Note that in this chapter, the rotor speed and frequency of each generator at its terminal bus are assumed to be monitored by PMUs. This is a reasonable assumption due to several reasons: 1) online monitoring of generators plays a major role in power system operation and control, thus it is given a high priority for PMU placement according to the NERC PMU placement standard [90]; 2) it is required by NERC standard [112] to have PMUs installed at the point of interconnection for power plant model validation and verification.

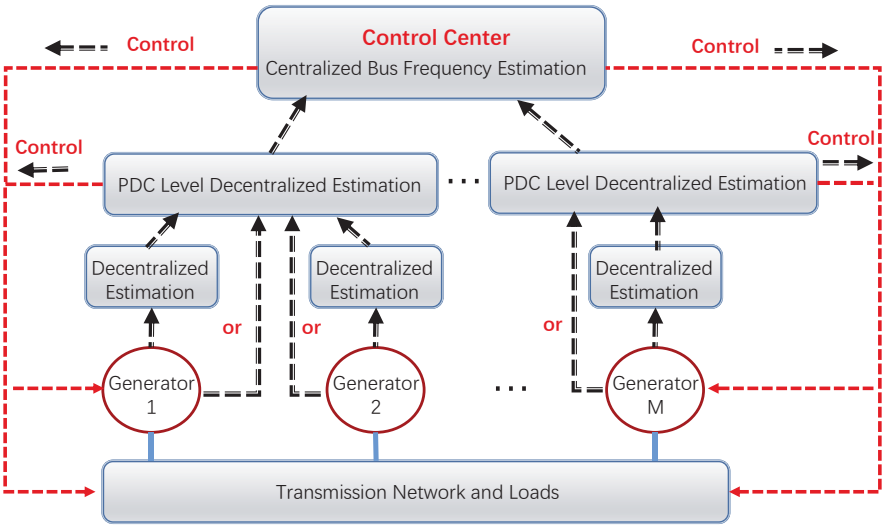


Figure 5.1: Proposed decentralized-centralized bus frequency estimation framework.

5.3 Proposed Robust Frequency Divider

By looking at (5.5), one may easily come up with the idea that this equation can be formulated as a regression problem, where the generator rotor speeds are measurements provided by PMUs while the frequency of each bus is the unknown state vector to be estimated. However, since the number of rows of the matrix $\boldsymbol{D}$ is larger than that of the columns, (5.5) cannot be treated as an estimation problem. Therefore, alternative approaches should be developed.

Note that to obtain an accurate estimation of bus frequency, accurate generator rotor speeds

are required. Thus, if measurement quality issues can be addressed locally at the generator bus through a local robust estimator, we are able to obtain accurate bus frequency estimates. To this end, we propose a decentralized-centralized bus frequency estimation framework shown in Fig. 5.1. Particularly, we first perform the robust unscented Kalman filter-based dynamic state estimator at the local generator substation or phasor data concentrator (PDC) level by using the proposed model decoupling approach; then these local estimates (generator rotor speeds and angles) are transmitted to control center for bus frequency estimation through (5.5). Note that if the decentralized dynamic state estimation at each generator substation is costly, the data of those generators associated with PMU measurements will be transmitted to the PDC or regional system operating center for decentralized estimation. After that, local controls can be initiated if required, otherwise, they will be further communicated to the control center for bus frequency estimation and coordinated control. In the following subsections, we will elaborate on each of the block.

5.3.1 Generator Model Decoupling Approach

Due to power unbalance and other control actions, the rotor of the generator will accelerate or deaccelerate. Those electro-mechanical dynamics can be captured by the swing equations shown as follows [84]:

$$\frac{d\delta}{dt} = \omega - \omega_0, \tag{5.6}$$

$$\frac{2H}{\omega_0}\frac{d\omega}{dt} = T_M - P_e - D_p(\omega - \omega_0), \tag{5.7}$$

where δ is the rotor angle; H is the generator inertia constant; T_M and P_e are the generator mechanical power and electrical power outputs, respectively; it is assumed that T_M remains the same value as that in steady-state condition during transient process, which is reasonable as the time constant of the governor is very large; D_p is the generator damping constant. Note that the generator first swing dynamics are affected by P_e significantly. In other words, the generator swing equation is coupled with its dq windings and the rest of the system through P_e. For example, in the two axis generator model, $P_e = E'_d I_d + E'_q I_q + (X'_q - X'_d) I_d I_q$, where E'_d and E'_q are the d-axis and q-axis transient voltages, respectively; X'_d and X'_q are generator d-axis and q-axis transient reactance, respectively; I_d and I_q are the d and q axis currents, respectively. For more detailed model, the expression is different.

Motivated by the aforementioned analysis, if P_e is measured and taken as the model input while the measured frequency by PMUs is treated as outputs, the swing equation can be decoupled from the rest of the system. Furthermore, no information of the d and q damper windings is required. By doing that, no complex generator model is assumed. The physical mean of this decoupling approach can be explained as follows: when there is a disturbance at one point of the power system, synchronous generators will response to it through actions on the rotors; these responses reveal themselves in their real power injections and frequency. In other words, the generator swing dynamics are coupled with the rest of the system at

the point of connection, and its interactions with the rest of the system are through the real power injections and frequency. If real power injections and frequency are measured by a PMU, its swing responses to the disturbance are captured completely and no other system information is required.

5.3.2 Proposed Robust Decentralized Dynamic State Estimator

The model decoupling approach enables a generator swing equation to be decoupled from the rest of the system model, which in turn allows us to rely only on local measurements to estimate the rotor speed and angle of a generator. The discrete-time state representation of the ith synchronous generator is

$$\boldsymbol{x}_k^i = \boldsymbol{f}_i\left(\boldsymbol{x}_{k-1}^i, \boldsymbol{u}_k^i\right) + \boldsymbol{w}_k^i, \tag{5.8}$$

$$\boldsymbol{z}_k^i = \boldsymbol{h}_i\left(\boldsymbol{x}_k^i, \boldsymbol{u}_k^i\right) + \boldsymbol{v}_k^i, \tag{5.9}$$

where $\boldsymbol{x}_k^i$ is the state vector, including the generator rotor speed ω_i and rotor angle δ_i; $\boldsymbol{z}_k^i$ is the measurement vector that contains rotor speed z_{k1} provided by PMUs and frequency z_{k2} by local metering devices or frequency monitoring network devices [109]; $\boldsymbol{f}_i(\cdot)$ represents the discrete-time form of (7.5) and (7.6) while $\boldsymbol{z}_k^i = [z_{k1}\ z_{k2}]^T$ and $z_{k1} = \omega_i + v_{k1}$, $z_{k2} = (1 + \Delta\omega_i)f_0 + v_{k2}$; f_0 is the nominal system frequency; $\boldsymbol{h}_i(\cdot) = \boldsymbol{H}_k^i\boldsymbol{x}_k^i$ and $\boldsymbol{H}_k^i$ is a constant matrix that can be derived directly from the measurement equations of z_{k1} and z_{k2}; $\boldsymbol{w}_k^i$ and $\boldsymbol{v}_k^i = [v_{k1}\ v_{k2}]^T$ are the process and observation noise, respectively; they are assumed to be Gaussian with zero mean and covariance matrices $\boldsymbol{Q}_k^i$ and $\boldsymbol{R}_k^i$, respectively; $\boldsymbol{u}_k^i$ is the input that contains the real power injection of the ith generator.

Based on the derived discrete-time state space equations (5.8) and (5.9), the dynamic state estimator (DSE) can be used to estimate the generator rotor speed and angle using local measurements. In this chapter, the UKF is chosen as the basic DSE as it achieves a more balanced performance between computational efficiency and ability to cope with strong system nonlinearities than the extended Kalman filter, or the particle filter [113]. However, UKF has been proved to be sensitive to gross errors, cyber attacks and loss of measurements, etc [79, 80]. To handle these issues, a robust Generalized Maximum-likelihood-type UKF (GM-UKF) is proposed. It consists of four major steps, namely a batch-mode regression form step, a robust pre-whitening step, a robust regression state estimation step, and a robust error covariance matrix updating step. In the following subsections, we will discuss them in detail. Note that the index i is neglected for simplicity but without the loss of generality.

Derive Batch-Mode Regression Model

Given a state estimate at time step k-1, $\widehat{\boldsymbol{x}}_{k-1|k-1} \in \mathbb{R}^{n\times 1}$, having a covariance matrix given by $\boldsymbol{P}^{xx}_{k-1|k-1}$, its statistics are captured by $2n$ weighted sigma points defined as [113]

$$\boldsymbol{\chi}^{i}_{k-1|k-1} = \widehat{\boldsymbol{x}}_{k-1|k-1} \pm \left(\sqrt{n\boldsymbol{P}^{xx}_{k-1|k-1}}\right)_i, \tag{5.10}$$

with weights $w_i = \frac{1}{2n}, i = 1, ..., 2n$, where n is the number state variables for each generator. Then, each sigma point is propagated through the nonlinear system process model (5.8), yielding a set of transformed samples expressed as

$$\boldsymbol{\chi}^{i}_{k|k-1} = \boldsymbol{f}\left(\boldsymbol{\chi}^{i}_{k-1|k-1}\right). \tag{5.11}$$

Next, the predicted sample mean and sample covariance matrix of the state vector are calculated by

$$\widehat{\boldsymbol{x}}_{k|k-1} = \sum_{i=1}^{2n} w_i \boldsymbol{\chi}^{i}_{k|k-1}, \tag{5.12}$$

$$\boldsymbol{P}^{xx}_{k|k-1} = \sum_{i=1}^{2n} w_i (\boldsymbol{\chi}^{i}_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})(\boldsymbol{\chi}^{i}_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})^T + \boldsymbol{Q}_k. \tag{5.13}$$

We define $\widehat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{x}_k - \boldsymbol{\Delta}_k$, where $\boldsymbol{x}_k$ is the true state vector; $\boldsymbol{\Delta}_k$ is the prediction error and $\mathbb{E}\left[\boldsymbol{\Delta}_k \boldsymbol{\Delta}_k^T\right] = \boldsymbol{P}^{xx}_{k|k-1}$. By processing the predictions and observations simultaneously, we have the following batch-mode regression form:

$$\begin{bmatrix} \boldsymbol{z}_k \\ \widehat{\boldsymbol{x}}_{k|k-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_k \\ \boldsymbol{I} \end{bmatrix} \boldsymbol{x}_k + \begin{bmatrix} \boldsymbol{v}_k \\ -\boldsymbol{\Delta}_k \end{bmatrix} \tag{5.14}$$

which can be rewritten in a compact form

$$\widetilde{\boldsymbol{z}}_k = \widetilde{\boldsymbol{H}}_k \boldsymbol{x}_k + \widetilde{\boldsymbol{e}}_k, \tag{5.15}$$

and the error covariance matrix is

$$\boldsymbol{W}_k = \mathbb{E}\left[\widetilde{\boldsymbol{e}}_k \widetilde{\boldsymbol{e}}_k^T\right] = \begin{bmatrix} \boldsymbol{R}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}^{xx}_{k|k-1} \end{bmatrix} = \boldsymbol{S}_k \boldsymbol{S}_k^T, \tag{5.16}$$

where $\boldsymbol{I}$ is an identity matrix; $\boldsymbol{S}_k$ is calculated by the Cholesky decomposition technique.

Perform Robust Pre-whitening

Before carrying out a robust regression, the state prediction errors of the batch-mode regression form need to be uncorrelated. This can be done by pre-multiplying $\boldsymbol{S}_k^{-1}$ on both sides of (7.32), yielding

$$\boldsymbol{S}_k^{-1} \widetilde{\boldsymbol{z}}_k = \boldsymbol{S}_k^{-1} \widetilde{\boldsymbol{H}}_k \boldsymbol{x}_k + \boldsymbol{S}_k^{-1} \widetilde{\boldsymbol{e}}_k, \tag{5.17}$$

which can be further organized to the compact form

$$\boldsymbol{y}_k = \boldsymbol{A}_k \boldsymbol{x}_k + \boldsymbol{\xi}_k, \tag{5.18}$$

where $\mathbb{E}[\boldsymbol{\xi}_k \boldsymbol{\xi}_k{}^T] = \boldsymbol{I}$. However, if outliers occur, the use of $\boldsymbol{S}_k^{-1}$ for prewhitening will cause negative smearing effect. To handle this issue, we first detect and downweight the outliers by means of weights calculated using the projection statistics (PS) [37] and a statical test applied to them. Those weights contribute to the robust prewhitening and their functionals will be shown later in the objective function. Specifically, we apply the PS to a 2-dimensional matrix $\boldsymbol{Z}$ that contains serially correlated samples of the innovations and of the predicted state variables. Formally, we have

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{z}_{k-1} - \boldsymbol{H}_k \widehat{\boldsymbol{x}}_{k-1|k-2} & \boldsymbol{z}_k - \boldsymbol{H}_k \widehat{\boldsymbol{x}}_{k|k-1} \\ \widehat{\boldsymbol{x}}_{k-1|k-2} & \widehat{\boldsymbol{x}}_{k|k-1} \end{bmatrix}, \tag{5.19}$$

where $\boldsymbol{z}_{k-1} - \boldsymbol{H}_k \widehat{\boldsymbol{x}}_{k-1|k-2}$ and $\boldsymbol{z}_k - \boldsymbol{H}_k \widehat{\boldsymbol{x}}_{k|k-1}$ are the innovation vectors while $\widehat{\boldsymbol{x}}_{k-1|k-2}$ and $\widehat{\boldsymbol{x}}_{k|k-1}$ are the predicted state vectors at time instants k-1 and k, respectively. The PS values of the predictions and of the innovations are separately calculated because the values taken by the former and the latter are centered around different points. The implementation of PS can be found in [37].

Once the PS values are calculated, they are compared to a threshold to identify outliers. According to our previous work [79, 80], $\boldsymbol{Z}$ can be shown to follow a bivariate Gaussian probability distribution and the calculated PS values using $\boldsymbol{Z}$ follow a chi-square distribution with degree of freedom 2. As a result, the outliers can be flagged if their PS values satisfy $\mathrm{PS}_i > \chi^2_{2,0.975}$ at a significance level 97.5% in the statistical test, and are further downweighted via [79]

$$\varpi_i = \min\left(1, d^2 / PS_i^2\right), \tag{5.20}$$

where the parameter d is set as 1.5 to yield good statistical efficiency at Gaussian distribution.

Remark: Except for the occurrence of outliers in rotor speed measurements, outliers may occur in the measured real power injections as well, and consequently, yielding incorrect predicted states, called innovation outliers [79, 80]. In such condition, the predicted state corresponding to incorrect real power injection will be flagged as outliers. Since the local measurement redundancy is not high, we will not downweight it directly, instead we propose to replace the current real power injection by its previous value and obtain the new state predictions. By doing that, we can achieve a better statistical efficiency.

Carry out Robust Regression

To address the data quality issues, we develop a robust GM-estimator that minimizes the following objective function:

$$J(\boldsymbol{x}_k) = \sum_{i=1}^{l} \varpi_i^2 \rho(r_{S_i}), \tag{5.21}$$

where $l = m + n$ and m is the number of measurements; ϖ_i is calculated by (5.20); $r_{S_i} = r_i/s\varpi_i$ is the standardized residual; $r_i = y_i - \boldsymbol{a}_i^T\widehat{\boldsymbol{x}}$ is the residual, where $\boldsymbol{a}_i^T$ is the ith row vector of the matrix $\boldsymbol{A}_k$; $s = 1.4826 \cdot b_m \cdot \text{median}_i\,|r_i|$ is the robust scale estimate; b_m is a correction factor; $\rho(\cdot)$ is the convex Huber-ρ function, that is

$$\rho\left(r_{S_i}\right) = \begin{cases} \frac{1}{2}r_{S_i}^2, & \text{for } |r_{S_i}| < \lambda \\ \lambda\,|r_{S_i}| - \lambda^2/2, & \text{elsewhere} \end{cases}, \tag{5.22}$$

where the parameter λ is typically chosen to be between 1.5 and 3 to achieve high statistical efficiency in the literature [82].

To minimize (5.21), the following necessary condition must be satisfied

$$\frac{\partial J\left(\boldsymbol{x}_k\right)}{\partial \boldsymbol{x}_k} = \sum_{i=1}^{l} -\frac{\varpi_i \boldsymbol{a}_i}{s}\psi\left(r_{S_i}\right) = \boldsymbol{0}, \tag{5.23}$$

where $\psi\left(r_{S_i}\right) = \partial\rho\left(r_{S_i}\right)/\partial r_{S_i}$is the so-called ψ-function. By dividing and multiplying the standardized residual r_{S_i} to both sides of (5.23) and putting it in a matrix form, we get

$$\boldsymbol{A}_k^T\boldsymbol{\Lambda}\left(\boldsymbol{y}_k - \boldsymbol{A}_k\boldsymbol{x}_k\right) = \boldsymbol{0}, \tag{5.24}$$

where $\boldsymbol{\Lambda}$ =diag$(q\left(r_{S_i}\right))$ and $q\left(r_{S_i}\right) = \psi\left(r_{S_i}\right)/r_{S_i}$. By using the IRLS algorithm [37], the state estimates at the j iteration can be updated using

$$\Delta\widehat{\boldsymbol{x}}_{k|k}^{(j+1)} = \left(\boldsymbol{A}_k^T\boldsymbol{\Lambda}^{(j)}\boldsymbol{A}_k\right)^{-1}\boldsymbol{A}_k^T\boldsymbol{\Lambda}^{(j)}\boldsymbol{y}_k, \tag{5.25}$$

where $\Delta\widehat{\boldsymbol{x}}_{k|k}^{(j+1)} = \widehat{\boldsymbol{x}}_{k|k}^{(j+1)} - \widehat{\boldsymbol{x}}_{k|k}^{(j)}$. The algorithm converges when $\left\|\Delta\widehat{\boldsymbol{x}}_{k|k}^{(j+1)}\right\|_\infty \leq 10^{-2}$.

Update Error Covariance Matrix

After the convergence of the algorithm, the estimation error covariance matrix $\boldsymbol{P}_{k|k}^{xx}$ of the GM-UKF needs to be updated so that the state prediction at the next time sample can be performed. Following the work from [79], we derive the estimation error covariance matrix of our GM-UKF as

$$\boldsymbol{P}_{k|k}^{xx} = \frac{\mathbb{E}_F\left[\psi^2\left(r_{S_i}\right)\right]}{\left\{\mathbb{E}_F\left[\psi'\left(r_{S_i}\right)\right]\right\}^2}\left(\boldsymbol{A}_k^T\boldsymbol{A}_k\right)^{-1}\left(\boldsymbol{A}_k^T\boldsymbol{Q}_\varpi\boldsymbol{A}_k\right)\left(\boldsymbol{A}_k^T\boldsymbol{A}_k\right)^{-1} \tag{5.26}$$

where $\boldsymbol{Q}_\varpi = diag\left(\varpi_i^2\right)$.

Remark: the proposed robust DSE is supposed to be performed for each generator substation. It can be implemented at the control center as well if all the generator data and PMU measurements are transmitted from local substations to it. However, there exist several concerns by doing so, such as increased communication burden that is discussed in the next

subsection and the delayed local control, etc. Indeed, if all the calculations are performed at the control center while some local controls are required at this period, the estimated bus frequency for those local controls can be delayed; by contrast, our robust DSE is first conducted locally using local PMU measurements and its estimated rotor speeds and angles can be used for local controls; in the meantime, they can be transmitted to control center for bus frequency estimation and coordinated control. In practice, if it is costly to implement the decentralized DSE for each generator substation, we can do it at the local phasor data concentrator (PDC) level or regional system level. This can still save a lot of communication burden and enable the timely local control actions compared with the fully centralized strategy.

5.3.3 Bus Frequency Estimation

When the rotor speed and rotor angle of each generator are obtained, they need to be communicated to the control center for bus frequency estimation. Depending on the applications, there are two ways to communicate the estimation results.

- If the control center is only interested in monitoring and regulating the system frequency, the rotor speed of each generator is transmitted and the bus frequency is estimated using (5.5). Compared with the conventional strategy, that is, all the measured voltage magnitudes and angles, current magnitudes and angles, and frequency by PMUs are communicated, the communication burden of our proposed approach is only 20% of it;
- If the control center are interested in both frequency and rotor angle stability monitoring and control, rotor speed and angle estimates are communicated. In this case, it only requires 40% communication burden of the conventional strategy.

Note that the total computing time of the proposed approach consists of two parts: the decentralized DSE and the projection of the rotor speed to bus frequency through (5.5). Since each robust DSE is performed locally and its estimates are communicated to the control center through its communication link, the proposed DSE is independent of the size of the power system. The only concern of the proposed RFD for very large-scale power system online applications is that $\boldsymbol{D}$ becomes rather dense, which requires a lot of computer memory. However, this is not a problem if we use the sparse matrices $\boldsymbol{B}_{BB}$, $\boldsymbol{B}_{G0}$ and $\boldsymbol{B}_{BG}$ instead of $\boldsymbol{D}$ for the projection of rotor speeds to bus frequencies (see equation (20) in [107]). As a result, the computational burden of this step is negligible.

Remark: The reason that the complicated generator model is not required has been discussed in the model decoupling section. We discuss here how the proposed RFD is not dependent on load model. It is well-known that the power system dynamics are different if different load models are assumed. However, although system dynamic behaviors are different, they are reflected on the variations of the rotor speeds of synchronous generators. Since the proposed

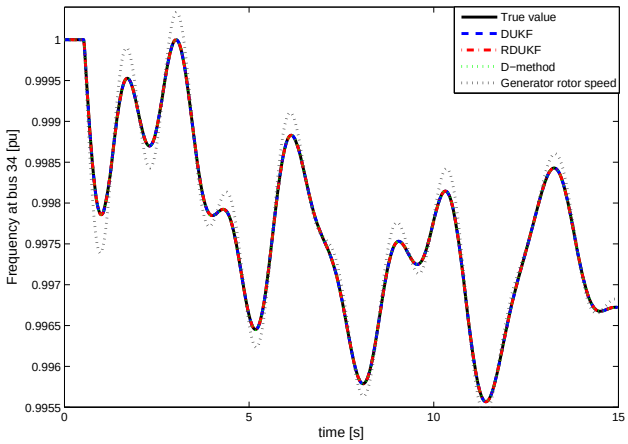


Figure 5.2: Comparing the estimated frequency at bus 34 by DUKF, RDUKF and D-method with normal measurement noise in the IEEE 39-bus system.

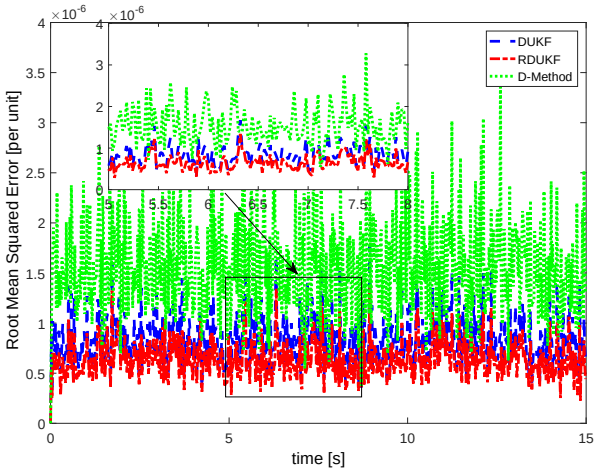


Figure 5.3: RMSE of DUKF, RDUKF and D-method with normal measurement noise in the IEEE 39-bus system.

RFD is based on such variations, load models have been implicitly taken into account. This conclusion has also been verified by extensive simulation results in [107].

5.4 Numerical Results

In this section, extensive simulations on the IEEE 39-bus test system will be carried out to demonstrate the effectiveness and robustness of the proposed RFD. Specifically, at t=0.5s,

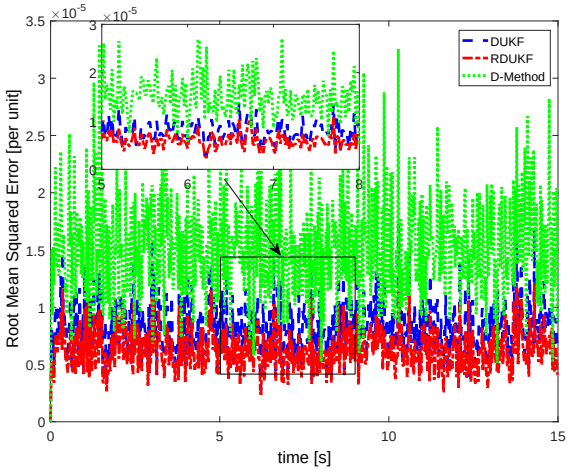


Figure 5.4: RMSE of DUKF, RDUKF and D-method with large measurement noise in the IEEE 39-bus system, where the noise covariance matrix is changed from $10^{-6}\boldsymbol{I}$ to $10^{-4}\boldsymbol{I}$.

the Generator 4 connected to bus 33 is tripped to simulate system disturbance. The transient stability simulations are performed to generate measurements and true state variables using the Matlab-based software PST with some revisions [114]. The fourth order Ruger-Kutta approach is adopted with integration step t=1/120s to solve differential and algebraic equations. The measured real power injection of each generator is taken as model input, while the measured generator rotor speed and frequency by PMU are treated as outputs/measurements. A random Gaussian variable with zero mean and variance equal to 10^{-6} is assumed for system process noise. The generator model assumed for transient simulation is the detailed two-axis generator model, whose parameter values are taken from [70]. The root-mean-squared error (RMSE) of all bus frequencies is used as the performance index while the estimated frequency at bus 34 is taken for illustration. Note that, Generator 5 is connected to bus 34. The proposed non-robust UKF based method will be called DUKF, and the proposed robust UKF based method is called RDUKF while the original proposal [107] that works on $\boldsymbol{D}$ matrix directly will be called the D-method.

5.4.1 Estimation Results with Noisy Measurements

All the methods are tested with noisy measurements. Normal and large noise are considered; their noise covariance matrices are assumed to be $10^{-6}\boldsymbol{I}$ and $10^{-4}\boldsymbol{I}$ with appropriate dimensions, respectively. The results are shown in Figs. 5.2- 5.4, where in Fig. 5.2 the rotor speed of Generator 5 is shown. From Fig. 5.2, we observe that the rotor speed of Generator 5 is different from its terminal bus frequency. This difference is caused by two factors: (i) generator internal impedance and (ii) severity of the transient (e.g., how much rotor speeds differ from each other). On the other hand, by observing both Figs. 5.2 and 5.3, it is found that the D-method is one of the most sensitive method to measurement noise while our

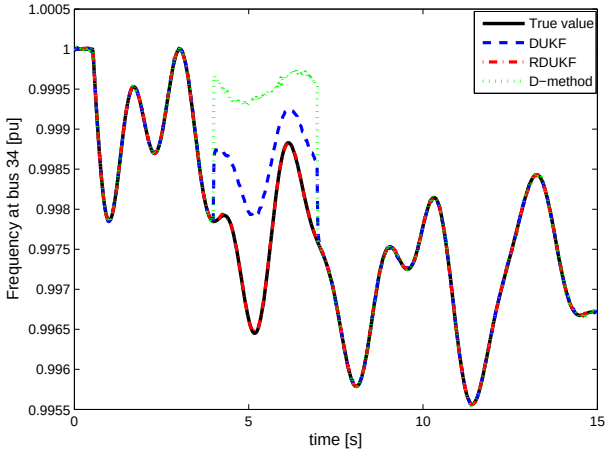


Figure 5.5: Estimated frequency at bus 34 by DUKF, RDUKF and D-method with observation outliers in the IEEE 39-bus system, where the measured rotor speed of Generator 5 is contaminated with 20% error from t=4s to t=6s.

DUKF and RDUKF approaches are able to filter them out. This is because in D-method, the noise is directly propagated through the rotor speed measurements to the bus frequency. By contrast, our methods adopt the UKF to filter out the noise, yielding better performance. When we increase the measurement noise level, i.e., the covariance matrix is changed from $10^{-6}\boldsymbol{I}$ to $10^{-4}\boldsymbol{I}$, the results of D-method become even worse while our methods can achieve comparable performance as those in the former case (see Fig. 5.4).

5.4.2 Impact of Observation and Innovation Outliers

Due to cyber attacks, imperfect phasor synchronization, the saturation of metering current transformers or by metering Couple Capacitor Voltage Transformers (CCVTs), to name a few, gross errors can occur in the PMU measurements [79, 99]. As for our decentralized DSE-based bus frequency estimation problem, there are two ways to induce outliers: i) the measured rotor speed by PMUs is contaminated with gross error, which is called observation outlier; ii) since the real power injection measured by PMUs can be contaminated with gross error, treating it as model input can yield incorrect rotor speed predictions, which is called innovation outlier. Note that, as D-method is working directly with rotor speed measurements, it is affected by observation outliers while being independent of the innovation outlier caused by incorrect real power injection measurements. To this end, two cases are considered:

Case 1: the measured rotor speed of Generator 5 is contaminated with 20% error from t=4s to t=6s to simulate observation outlier.

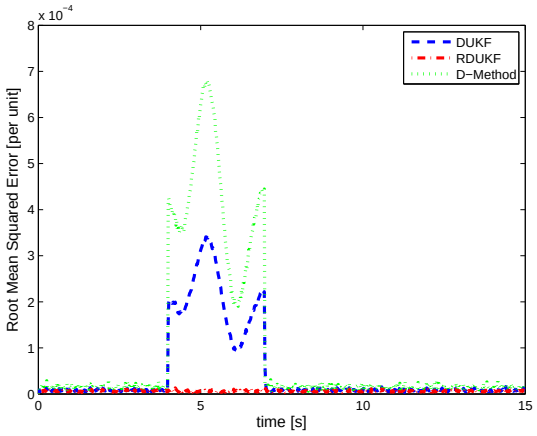


Figure 5.6: RMSE of DUKF, RDUKF and D-method with observation outliers in the IEEE 39-bus system, where the measured rotor speed of Generator 5 is contaminated with 20% error from t=4s to t=6s.

Case 2: the measured real power injection of Generator 5 is contaminated with 30% error from t=3s to t=6s to simulate innovation outlier.

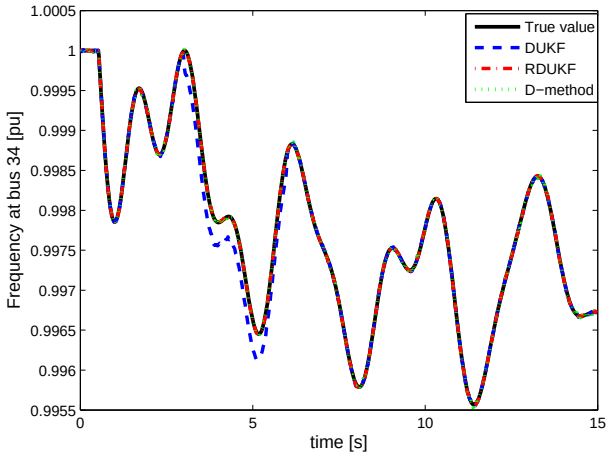


Figure 5.7: Estimated frequency at bus 34 by DUKF, RDUKF and D-method with innovation outliers in the IEEE 39-bus system, where the measured real power injection of Generator 5 is contaminated with 30% error from t=3s to t=6s.

The test results for Case 1 are shown in Figs. 5.5 and 5.6. From these two figures, we find that the estimation results of the DUKF and the D-method are significantly biased in the presence of observation outliers. DUKF is less sensitive to the observation outlier compared with the D-method. By contrast, our RDUKF is able to suppress the observation outliers thanks to the robustness provided by PS and the GM-estimator, yielding negligible bias of

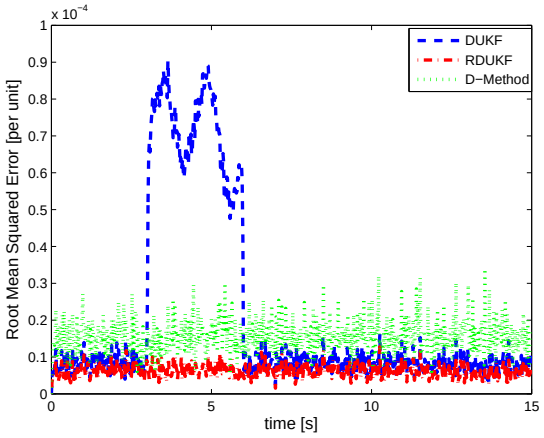


Figure 5.8: RMSE of DUKF, RDUKF and D-method with innovation outliers in the IEEE 39-bus system, where the measured real power injection of Generator 5 is contaminated with 30% error from t=3s to t=6s.

the estimation. It should be noted that due to the smearing effect of applying (5.5) for bus frequency estimation, the estimated frequencies at many buses are affected by the incorrect rotor speed of the generator 5. This however does not happen in our RDUKF.

The test results for Case 2 are shown in Figs. 5.7 and 5.8. As expected, the results of DUKF are biased in the presence of innovation outliers while the D-method is not affected. Due to the robustness of the proposed DSE, this innovation outlier has been suppressed. By comparison, RDUKF still outperforms D-method, yielding the best results.

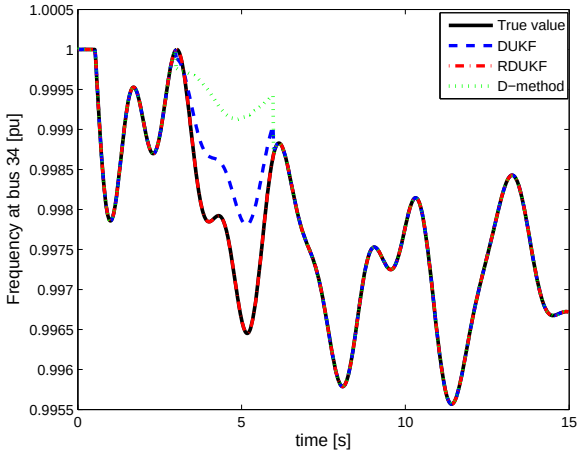


Figure 5.9: Estimated frequency at bus 34 by DUKF, RDUKF and D-method with measurement losses from t=3s to t=6s in the IEEE 39-bus system.

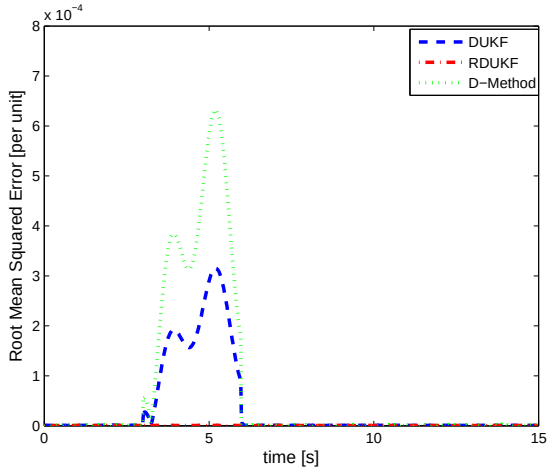


Figure 5.10: RMSE of DUKF, RDUKF and D-method with measurement losses from t=3s to t=6s in the IEEE 39-bus system.

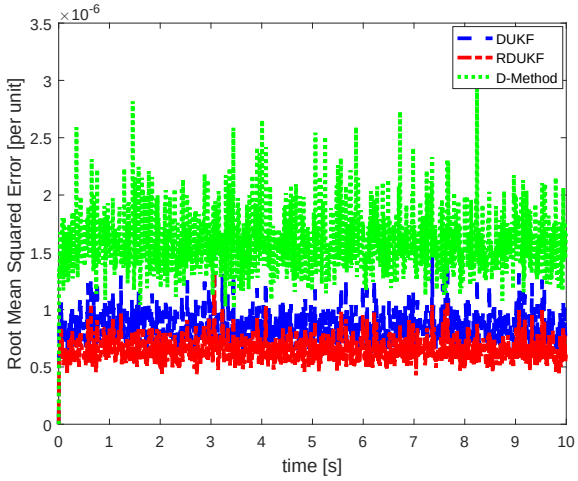


Figure 5.11: RMSE of DUKF, RDUKF and D-method with normal measurement noise in the IEEE 145-bus system.

5.4.3 Loss of PMU Measurements

Due to the failures of communication links between PMU and phasor data concentrator or cyber attacks, the PMU placed at Bus 34, where Generator 5 is connected, is assumed to loss its measurements from t=3s to t=6s. Therefore, the measurement set becomes unavailable during this time interval and their values are set equal to zero for simulation purpose. Please note that in this extreme case, both predicted and measured rotor speeds will be flagged as outliers. To enable the estimation of bus frequency by the proposed method, we advocate to either recover the missing data using [115] or perform short-term forecasting of the PMUs

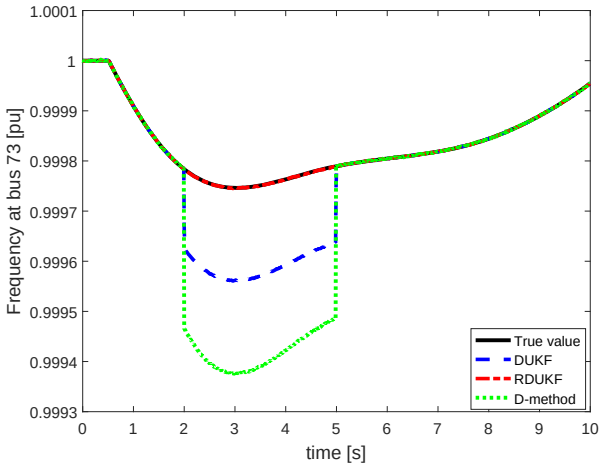


Figure 5.12: Estimated frequency at bus 73 by DUKF, RDUKF and D-method with outliers in the IEEE 145-bus system, where 10% of the measured generator rotor speeds is contaminated with 20% error from t=2s to t=5s.

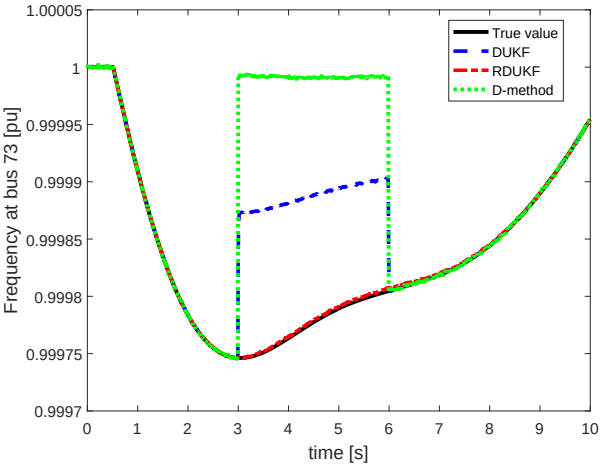


Figure 5.13: Estimated frequency at bus 73 by DUKF, RDUKF and D-method with measurement losses in the IEEE 145-bus system, where 50% of the measured generator rotor speeds are lost from t=3s to t=6s.

using their spatial and temporal correlations [116]. The test results are presented in Figs. 5.9 and 5.10. It can be seen from these two figures that the estimated bus frequencies of both the DUKF and the D-method are biased significantly. But the DUKF approach is less sensitive to the measurement losses than the D-method. As a result, the proposed RDUKF can always track the bus frequency reliably and accurately. It should be noted that the two mitigation approaches [115, 116] can be used in the situation that the measurements are temporally lost (a few seconds). For longer period, they may not be valid. Otherwise, the

operator will be warned and the decentralized DSE stops before further careful investigations are done.

5.4.4 Results on Large-Scale Systems

To demonstrate the applicability of the proposed robust frequency divider for large-scale system, the 50-machine IEEE 145-bus system is used. The dynamic data can be found through [117]. The generator located at bus 106 is tripped at t=0.5s to simulate the system disturbance. The following three scenarios are considered and tested:

- Scenario 1: only normal Gaussian noise is added to the simulated data like the test done in Section IV-A;
- Scenario 2: 10% of the measured generator rotor speeds is contaminated with 20% error from t=2s to t=5s;
- Scenario 3: 50% of the measured generator rotor speeds are lost from t=3s to t=6s.

The test results of all three scenarios are displayed in Figs. 5.11-5.13, where the RMSE of scenario 1 and the estimated frequencies of bus 73 under scenarios 2 and 3 are represented accordingly. Note that bus 73 is close to the tripped generator. Based on these figures, we conclude that our proposed robust frequency estimator still outperforms the other two alternatives in a larger-scale system. These results are consistent with those for the IEEE 39-bus system. It is interesting to find that even without using missing data recovering approach [115, 116], our proposed method is able to rely on good PMU measurements and the predicted dynamic state variables for filtering, yielding good estimation results.

5.4.5 Computational Efficiency

To validate the capability of the proposed method for online estimation, that is, to be compatible with PMU sampling rate, its computational efficiency is analyzed. All cases and scenarios simulated in the previous sections are considered. All the tests are performed on a PC with Intel Core i5, 2.50 GHz, 8GB of RAM. The average computing time of each method for every PMU sample is displayed in the Table 5.1. We observe from this table that all methods have comparative computational efficiency and their computing times are much lower than the PMU sampling period, which are 16.7ms and 8.3ms for 60 samples/s and 120 samples/s, respectively. On the other hand, although RDUKF is the most time consuming method compared with RDUKF and D-method, its computing time is negligible for practical applications considering the PMU sampling speed. Note that the decentralized DSE for each generator can be carried out independently and in a parallel manner, which are very fast to calculate and independent of the scale of the power system. For very large-scale power systems, $\boldsymbol{D}$ becomes rather dense and a numerical stable and computational

Table 5.1: Average Computing Times of The Three Bus Frequency Estimators At Each PMU Sample

Scenarios	D-method	DUKF	RDUKF
Section IV-A	0.02ms	0.15ms	0.37ms
Section IV-B Case 1	0.022ms	0.14ms	0.45ms
Section IV-B Case 2	0.021ms	0.16ms	0.46ms
Section IV-C	0.021ms	0.15ms	0.46ms
Scenario 1	0.08ms	0.66ms	1.6ms
Scenario 2	0.09ms	0.78ms	2.1ms
Scenario 3	0.085ms	0.79ms	2.14ms

efficient approach proposed in [107] is used to project the rotor speeds to bus frequencies. The computational burden of this step has been shown to be negligible [107]. In conclusion, the proposed method is suitable for large-scale power system online applications.

5.5 Conclusion

A robust frequency divider (RFD) is proposed to estimate the frequency of each bus in a power system. Our RFD consists of two steps, namely the estimation of generator rotor speeds through a robust decentralized UKF using local measurements, and the projection of all generator rotor speeds to bus frequency. The proposed RFD is model independent, and the knowledge of local PMU measurements at each generator terminal bus and transmission system line parameters is sufficient. Furthermore, the proposed RFD is able to filter out measurement noise, suppress gross measurement errors, handle cyber attacks as well as measurement losses. Extensive results carried out on the IEEE 39-bus and 145-bus systems demonstrate the effectiveness and the robustness of the proposed method. A possible issue of the proposed RFD is that the decentralized and centralized scheme may produce some delays for the estimated bus frequencies used for controllers. However, thanks to the advancement of control techniques, the time delays can be effectively mitigated [118, 119]. In the future work, we will design this type of robust frequency regulator based on our RFD.

Chapter 6

A Decentralized H-infinity Unscented Kalman Filter for Power System Dynamic State Estimation with Uncertainties

6.1 Introduction

Real-time dynamic state variables of synchronous generators are of vital importance for power system monitoring and implementing effective local and global system controls. With the wide-area deployment of phasor measurement units (PMUs), online application of a dynamic state estimator (DSE) becomes feasible.

To date, various Kalman filter-type DSEs have been proposed in the literature, including the extended Kalman filter (EKF) and the unscented Kalman filter, among others [22–27, 38, 86, 87, 120]. The proposed methods can be grouped into centralized and decentralized ones. The centralized methods are based on a global DSE executed at the control center, which processes information collected from all the system devices and the real-time PMU measurements [22, 25, 27, 38, 120]. Since all the measurements are communicated to the control center, a good global measurement redundancy is achieved, which in turn enhances the capability of the DSE to handle measurement losses and outliers. However, for large-scale interconnected power systems, a global DSE may be incompatible with real-time applications due to prohibitively large executing times. By contrast, decentralized methods are implemented at each local synchronous generator, and therefore are independent of system size. Consequently, they are suitable for large-scale power system online monitoring. In [24, 86], a decentralized DSE based on EKF is proposed to simultaneously estimate the states and the parameters of a classical generator model. In [26], a decentralized DSE based

on UKF is advocated to overcome the weaknesses of the EKF in handling strong nonlinearity of the dynamical model, yielding more accurate state estimates. In [23], a fourth-order generator model is considered and the mechanical torque and field voltage of the exciter are taken as system inputs when implementing the decentralized EKF-based DSE. Anagnostou and Pal [87] improved that approach by proposing a derivative free UKF-based DSE, where the unknown mechanical torque and field voltage of the exciter are estimated along with the synchronous generator state variables. [23, 26, 87] showed that their methods are highly sensitive to generator parameter uncertainties and large measurement noise with extensive simulation results. However, they did not propose related mitigation approaches.

It turns out that all the decentralized DSE methods proposed so far in the literature do not consider the cases of unknown or thick-tailed process and measurement noises, a possibility demonstrated by a PNNL report [35]. Indeed, due to unknown disturbances and changes of system inputs, the assumed dynamical model may not properly predict the dynamical behaviour of a power system, yielding unknown process noise. Other problems include: 1) measurement biases induced by CCVTs that are known to be sensitive to changing temperature and to aging processes; and 2) additional noise induced by the synchronization process of the received PMU measurements, yielding unknown or non-Gaussian measurement noise. As a result, the performances of Kalman filter-type DSEs may significantly degrade, resulting in unreliable state estimates.

To handle a variety of system uncertainties, we resort to robust control theory and develop a decentralized H-infinity UKF for power system dynamic state estimation. Our choice of an H-infinity filter is motivated by the fact that it has a proved performance in a variety of engineering applications such as designing controllers to address measurement delays, unknown inputs, signals that contain unknown noise statistics, among others [37, 121–123]. Unlike the Kalman filter that minimizes the mean-square error, the H-infinity filter minimizes the maximum variance of the state estimates due to measurement and process noise and model uncertainties [121]. To overcome the fact that the H-infinity filter is designed for linear dynamical system and thus may not be effective for power system DSE applications, first-order Taylor series approximations-based H-infinity EKF are proposed [124, 125]. However, a power system may exhibit strong nonlinearity on the dynamical model when it is heavily loaded or subject to large disturbances. As a result, first-order Taylor series approximation will induce large estimation errors, which may cause a divergence of the H-infinity EKF.

To address the above challenges, this chapter proposes a decentralized H-infinity UKF [126] that has the following salient features:

- It stems from the conversion of the traditional nonlinear UKF into an equivalent linear like batch-mode regression form through the use of the statistical linerization approach. Note that the latter does not involve any Jacobian matrix while being equivalent to the unscented transformation; therefore, it is able to handle strong system nonlinearities. By contrast, the first-order Taylor series approximation based H-infinity EKF involves cumbersome linerization through the calculation of Jacobian matrices, which results

in significant biased results in presence of strong nonlinearity;

- It involves a linear like batch-regression form that preserves all the properties of the UKF when casted into the H-infinity framework, resulting in an H-infinity UKF whose parameters can be tuned via an analytical method;
- It leverages two ways for the decentralized implementation using local PMU measurements; their effectiveness are compared as well;
- It has better robustness to unknown process and measurement noise and uncertain generator model and controller parameter values when compared to other approaches, including the H-infinity EKF [125], the traditional decentralized UKF [26] and the robust EKF [38].

The remainder of the paper is organized as follows. Section 6.2 presents the dynamical model of synchronous generator and the problem statement. Section 6.3 describes the proposed decentralized H-infinity UKF and Section 6.4 shows and analyzes the simulation results. Finally, Section 6.5 concludes the paper.

6.2 Problem Formulation

6.2.1 Synchronous Generator Modeling

In this chapter, the two-axis generator model with the IEEE-DC1A exciter and the TGOV1 turbine-governor is considered. Note that this is one of the recommended models by IEEE for transient stability analysis [95]. It consists of the following differential and algebraic equations [84]:

Differential equations:

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

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

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

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

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

$$T_{\mathrm{F}}\frac{dV_F}{dt} = -V_F + \frac{K_F}{T_E}V_R - \frac{K_F}{T_E}\left(K_E + S_E\left(E_{fd}\right)\right) E_{fd}, \tag{6.6}$$

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

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

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

Algebraic equations:

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

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

The discrete-time state space representation of the above equations can be expressed as

$$\boldsymbol{x}_k = \boldsymbol{f}\left(\boldsymbol{x}_{k-1}, \boldsymbol{u}_k\right) + \boldsymbol{w}_k, \tag{6.12}$$

$$\boldsymbol{z}_k = \boldsymbol{h}\left(\boldsymbol{x}_k\right) + \boldsymbol{v}_k, \tag{6.13}$$

where $\boldsymbol{x}_k$ is the state vector that includes generator state variables ω, δ, E'_d and E'_q, exciter state variables E_{fd}, V_F and V_R, and governor state variables T_M and P_{SV}; $\boldsymbol{f}(\cdot)$ is the nonlinear state transition function corresponding to the equations (7.3)-(7.11); note that ω is in radians while the other state variables are in p.u., and after estimation, $\widehat{\boldsymbol{\omega}}_{est} = \mathbf{1} + \widehat{\boldsymbol{\omega}}/\boldsymbol{\omega}_s$; $\boldsymbol{u}_k$ is the system input vector; $\boldsymbol{z}_k$ is the measurement vector, including metered generator terminal voltage or current phasors and frequency; $\boldsymbol{h}(\cdot)$ is the vector-valued measurement function corresponding to the equations (7.12)-(7.13); $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ are system process and measurement noises, respectively, and their covariance matrices are $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$.

6.2.2 Problem Statement

To estimate the state variables of the synchronous generator using (6.12)-(6.13), several decentralized EKF and UKF-based approaches have been proposed in [23,24,26,86,87]. These Kalman filter-type DSEs work well if the following conditions are satisfied [121]: i) the system process and observation noises are assumed to have zero mean and zero correlation at each time instant; ii) the covariance matrices $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$ are known; iii) the discrete-time nonlinear state-space model is accurate. However, these assumptions may be violated for practical systems. For instance, the disturbances and the system inputs may be unknown under some operating conditions of the system; the synchronization process may induce additional noise to PMU measurements, yielding unknown or even non-Gaussian measurement noise; the synchronous generator reactance and transient reactance can vary [127] due to the magnetic saturation of the iron core of the generator and to a changing temperature of the machine windings; due to the aging process, the addition of new digital protective relaying packages,

replacement of the old rotating pilot exciters and voltage regulators, and the controller failures, to name a few, controller parameters can change from time to time [128, 129]. As a result, their performances can be significantly degraded. To bound the influence of these uncertainties on the state estimates while handling strong nonlinear dynamical system model, we develop a decentralized H-infinity UKF as described next.

6.3 The Proposed Decentralized H-infinity UKF

This section presents two model decoupling approaches of a synchronous generator that enable the decentralized implementation of a DSE. Statistical linearization approach is used to convert the traditional nonlinear UKF into an equivalent linear batch-mode regression form. The latter allows us to resort to the linear H-infinity filter framework for the development of an H-infinity UKF.

6.3.1 Discrete-Time State Space Model for Decentralized DSE

Evidently, the dynamics of a synchronous generator can be accurately captured if one pair of the voltage or current phasors provided by PMUs are taken as inputs while the other pair of measurements are considered as outputs [26]. By doing so, a synchronous generator can be modeled as an isolated entity from a DSE prospective, without modeling the remaining part of the system to which it is interconnected. Depending on the way of treating the voltage and current phasors, two model decoupling approaches can be implemented. They are presented below as Approach A1 and Approach A2. The comparisons between A1 and A2 are conducted to demonstrate their advantages and disadvantages through extensive numerical studies. Note that the i-th generator is used as an example and the q-axis is assumed to lead d-axis.

Approach A1: Consider a synchronous generator where the stator current phasor $I_i\angle\phi_i$ is supposed to be a known input, i.e., $\boldsymbol{u}_i = [I_i \ \phi_i]^T$, and terminal bus voltage phasor $V_i\angle\theta_i$ and frequency f_i as outputs. Consequently, we have $I_{di} = \text{real}[I_i e^{j(\phi_i-\delta_i+\frac{\pi}{2})}]$, $I_{qi} = \text{Imag}[I_i e^{j(\phi_i-\delta_i+\frac{\pi}{2})}]$, $V_{di} = E'_{di} + I_{qi}X'_{qi}$, and $V_{qi} = E'_{qi} - I_{di}X'_{di}$. Furthermore, we have $\boldsymbol{z}_i = [z_1 \ z_2 \ z_3]^T$ and their expressions can be shown as follows:

$$z_1 = V_i = \sqrt{V_{di}^2 + V_{qi}^2} + v_{Vi}, \tag{6.14}$$

$$z_2 = \theta_i = \delta_i - arctan(\frac{V_{di}}{V_{qi}}) + v_{\theta i}, \tag{6.15}$$

$$z_3 = f_i = f_0(\Delta\omega_i + 1) + v_{fi}, \tag{6.16}$$

where $\boldsymbol{v}_i = [v_{Vi} \ v_{\theta i} \ v_{fi}]^T$ is the measurement noise with zero mean and covariance matrix $\boldsymbol{R}_k^i$.

Approach A2: Consider a synchronous generator where the terminal bus voltage phasor, $V_i\angle\theta_i$, is supposed to be known input, i.e., $\boldsymbol{u}_i = [V_i\ \theta_i]^T$, and current phasor $I_i\angle\phi_i$ and f_i are considered as outputs. Consequently, we have, $V_{di} = V_i\sin(\delta_i - \theta_i)$, $V_{qi} = V_i\cos(\delta_i - \theta_i)$, $I_{di} = (E'_{qi} - V_{qi})/X'_{di}$, and $I_{qi} = (-E'_{di} + V_{di})/X'_{qi}$. Furthermore, we have $\boldsymbol{z}_i = [z_1\ z_2\ z_3]^T$ and

$$z_1 = I_i = \sqrt{I_{di}^2 + I_{qi}^2} + v_{Ii}, \tag{6.17}$$

$$z_2 = \phi_i = \arg(I_{qi} - jI_{di}) + \delta_i + v_{\phi i}, \tag{6.18}$$

$$z_3 = f_i = f_0(\Delta\omega_i + 1) + v_{fi}, \tag{6.19}$$

where $\boldsymbol{v}_i = [v_{Ii}\ v_{\phi i}\ v_{fi}]^T$ is the measurement noise with zero mean and covariance matrix $\boldsymbol{R}_k^i$. It should be emphasized that in both approaches, the rotor speed deviation $\Delta\omega_i$ can be either obtained from the PMUs or calculated using the approach shown in [23]. The measurement or approximation error is characterized by v_{fi} and treated as one of the uncertainties.

Consequently, the discrete time state-space model of the ith generator at the time sample k is derived as

$$\boldsymbol{x}_k^i = \boldsymbol{f}_i\left(\boldsymbol{x}_{k-1}^i, \boldsymbol{u}_k^i\right) + \boldsymbol{w}_k^i, \tag{6.20}$$

$$\boldsymbol{z}_k^i = \boldsymbol{h}_i\left(\boldsymbol{x}_k^i\right) + \boldsymbol{v}_k^i, \tag{6.21}$$

where $\boldsymbol{w}_k^i$ is the system process noise that includes the noise associated with the input vector $\boldsymbol{u}_i$ and the model approximation error. Under such condition, the covariance matrix $\boldsymbol{Q}_k^i$ becomes uncertain or even unknown since it has a nonlinear relationship with $\boldsymbol{u}_i$.

6.3.2 Derivation of the H-infinity UKF

The model decoupling approaches presented above allow us to estimate the state variables of a generator through (6.20) and (6.21) using only local PMU measurements. As elaborated before, the derived discrete-time state-space model is subject to uncertainties, which will affect significantly the performance of the traditional UKF. To overcome this difficulty, we develop an H-infinity UKF that bounds the influence of the uncertainties on the state estimates, and thereby provides reasonably good state estimates of the synchronous generators. The proposed H-infinity UKF consists of two key steps, namely the derivation of an equivalent linear regression form of the UKF and of the state filtering using the H-infinity criterion. For the sake of simplicity, in the sequel the index i associated with the ith synchronous generator under study is neglected.

Derivation of the Equivalent Linear Regression Form of the Nonlinear UKF

The main idea of the UKF is to rely on the unscented transformation for nonlinear state transition and measurement prediction by strategically selecting sigma points within the

Kalman filter framework. The use of the sigma points for statistical information propagation avoids the linearization of the nonlinear state transition and measurement functions, and thus yields accurate state estimates [113].

Given a state estimate with mean $\widehat{\boldsymbol{x}}_{k-1|k-1} \in \mathbb{R}^{n\times 1}$ and covariance matrix $\boldsymbol{P}^{xx}_{k-1|k-1}$ at time step k-1, $2n$ weighted sigma points are generated through

$$\boldsymbol{\chi}^{j}_{k-1|k-1} = \widehat{\boldsymbol{x}}_{k-1|k-1} \pm \left(\sqrt{n\boldsymbol{P}^{xx}_{k-1|k-1}}\right)_j, \tag{6.22}$$

where the weights $w_j = 1/2n, j = 1, ..., 2n$. These sigma points are propagated through the nonlinear system process model (6.20) to obtain the following transformed sigma points

$$\boldsymbol{\chi}^{j}_{k|k-1} = \boldsymbol{f}\left(\boldsymbol{\chi}^{j}_{k-1|k-1}\right). \tag{6.23}$$

Then, the predicted state vector $\widehat{\boldsymbol{x}}_{k|k-1}$ and its covariance matrix are approximated by the weighted sample mean and sample covariance matrix of the transformed sigma points, respectively, yielding

$$\widehat{\boldsymbol{x}}_{k|k-1} = \sum_{j=1}^{2n} w_j \boldsymbol{\chi}^{j}_{k|k-1}, \tag{6.24}$$

$$\boldsymbol{P}^{xx}_{k|k-1} = \sum_{j=1}^{2n} w_j (\boldsymbol{\chi}^{j}_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})(\boldsymbol{\chi}^{j}_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})^T + \boldsymbol{Q}_k. \tag{6.25}$$

Next, the sigma points are updated to capture the information of system process noise through

$$\boldsymbol{\chi}^{j}_{k|k-1} = \widehat{\boldsymbol{x}}_{k|k-1} \pm \left(\sqrt{n\boldsymbol{P}^{xx}_{k|k-1}}\right)_j. \tag{6.26}$$

Then, the predicted measurement vector is given by $\widehat{\boldsymbol{z}}_{k|k-1} = \sum_{j=1}^{2n} w_j \boldsymbol{h}(\boldsymbol{\chi}^{j}_{k|k-1})$ and its associated error covariance matrix is estimated via

$$\boldsymbol{P}^{zz}_{k|k-1} = \sum_{j=1}^{2n} w_j (\boldsymbol{z}^{j}_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1})(\boldsymbol{z}^{j}_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1})^T + \boldsymbol{R}_k, \tag{6.27}$$

where $\boldsymbol{z}^{j}_{k|k-1} = \boldsymbol{h}(\boldsymbol{\chi}^{j}_{k|k-1})$.

In the work [54], the unscented transformation has been shown to be equivalent to the statistical linerization. As a result, the nonlinear regression model can be converted to its equivalent linear regression form. Motivated by this property, we apply statistical linerization to the nonlinear state transition and measurement functions around $\widehat{\boldsymbol{x}}_{k-1|k-1}$ and $\widehat{\boldsymbol{x}}_{k|k-1}$, respectively, yielding

$$\boldsymbol{x}_k = \boldsymbol{F}_k\left(\boldsymbol{x}_{k-1} - \widehat{\boldsymbol{x}}_{k-1|k-1}\right) + \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{e}_k + \boldsymbol{w}_k, \tag{6.28}$$

$$\boldsymbol{z}_k = \boldsymbol{H}_k\left(\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k-1}\right) + \widehat{\boldsymbol{z}}_{k|k-1} + \boldsymbol{\varepsilon}_k + \boldsymbol{v}_k, \tag{6.29}$$

where $\boldsymbol{F}_k = (\boldsymbol{P}^{x\chi}_{k|k-1})^T(\boldsymbol{P}^{xx}_{k-1|k-1})^{-1}$;

$$\boldsymbol{P}^{x\chi}_{k|k-1} = \sum_{j=1}^{2n} w_j(\boldsymbol{\chi}^j_{k-1|k-1} - \widehat{\boldsymbol{x}}_{k-1|k-1})(\boldsymbol{\chi}^j_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})^T, \tag{6.30}$$

and $\boldsymbol{e}_k$ is the statistical linearization error term on the nonlinear state transition function; it is normally distributed with zero mean and covariance matrix $\boldsymbol{L}_k = \boldsymbol{P}^{xx}_{k|k-1} - (\boldsymbol{P}^{x\chi}_{k|k-1})^T(\boldsymbol{P}^{xx}_{k-1|k-1})^{-1}\boldsymbol{P}^{x\chi}_{k|k-1}$; $\boldsymbol{H}_k = (\boldsymbol{P}^{xz}_{k|k-1})^T(\boldsymbol{P}^{xx}_{k|k-1})^{-1}$;

$$\boldsymbol{P}^{xz}_{k|k-1} = \sum_{i=1}^{2n} w_j(\boldsymbol{\chi}^j_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})(\boldsymbol{z}^j_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1})^T, \tag{6.31}$$

and $\boldsymbol{\varepsilon}_k$ is the statistical linearization error term on the nonlinear measurement function; it is normally distributed with zero mean and covariance matrix:

$$\boldsymbol{\Sigma}_k = \boldsymbol{P}^{zz}_{k|k-1} - (\boldsymbol{P}^{xz}_{k|k-1})^T(\boldsymbol{P}^{xx}_{k|k-1})^{-1}\boldsymbol{P}^{xz}_{k|k-1}. \tag{6.32}$$

As a result, the nonlinear dynamical model expressed by (6.20) and (6.21) can be converted into the following equivalent linear regression form:

$$\boldsymbol{x}_k = \boldsymbol{F}_k\boldsymbol{x}_{k-1} + \widehat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{F}_k\widehat{\boldsymbol{x}}_{k-1|k-1} + \boldsymbol{e}_k + \boldsymbol{w}_k, \tag{6.33}$$

$$\boldsymbol{z}_k = \boldsymbol{H}_k\boldsymbol{x}_k + \widehat{\boldsymbol{z}}_{k|k-1} - \boldsymbol{H}_k\widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{\varepsilon}_k + \boldsymbol{v}_k. \tag{6.34}$$

Note that $\boldsymbol{F}_k$ and $\boldsymbol{H}_k$ are no longer the Jacobian matrices, but $\boldsymbol{e}_k$ and $\boldsymbol{\varepsilon}_k$ are the statistical linearization errors that are used to preserve the nonlinearities of the state transition and the measurement functions, respectively. These terms are implicitly contained in the original unscented transformation-based UKF.

Derivation of the H-infinity UKF

The derived linear regression form allows us to resort to the original linear H-infinity Kalman filter framework for the development of our H-infinity UKF. Note that we aim to design a filter that guarantees a finite upper bound on the estimation error [121]. If the metered sequence values are provided up to time N, the foregoing objective can be achieved by satisfying the following inequality:

$$\frac{\sum_{k=0}^{N} \|\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k}\|^2_{\boldsymbol{P}^{-1}_{k|k}}}{\|\boldsymbol{x}_0 - \widehat{\boldsymbol{x}}_{0|0}\|^2_{\boldsymbol{P}^{-1}_{0|0}} + \sum_{k=0}^{N-1} \|\boldsymbol{w}_k\|^2_{\boldsymbol{Q}^{-1}_k} + \sum_{k=0}^{N} \|\boldsymbol{v}_k\|^2_{\boldsymbol{R}^{-1}_k}} \leq \gamma^2, \tag{6.35}$$

where $\boldsymbol{x}_0$ and $\boldsymbol{P}_{0|0}$ are the initial state vector and its covariance matrix, respectively; γ is a positive scalar parameter that bounds the uncertainties.

By substituting (7.25) and (7.26) into (6.35) and using the dynamic optimization approach advocated by Simon [121], we derive the filtered state vector as

$$\widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k[\boldsymbol{z}_k - \widehat{\boldsymbol{z}}_{k|k-1}], \tag{6.36}$$

$$\boldsymbol{K}_k = \boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k^T(\boldsymbol{H}_k\boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k^T + \boldsymbol{R}_k)^{-1} \tag{6.37}$$

$$= \boldsymbol{P}^{xz}_{k|k-1}(\boldsymbol{P}^{zz}_{k|k-1})^{-1} \tag{6.38}$$

$$\boldsymbol{P}^{xx}_{k|k} = (\boldsymbol{I} - \boldsymbol{P}^{xx}_{k|k-1}[\boldsymbol{H}_k^T\ \boldsymbol{I}]\boldsymbol{R}^{-1}_{e,k}[\boldsymbol{H}_k^T\ \boldsymbol{I}]^T)\boldsymbol{P}^{xx}_{k|k-1} \tag{6.39}$$

$$= \boldsymbol{P}^{xx}_{k|k-1} - [\boldsymbol{P}^{xz}_{k|k-1}\ \boldsymbol{P}^{xx}_{k|k-1}]\boldsymbol{R}^{-1}_{e,k}[\boldsymbol{P}^{xz}_{k|k-1}\ \boldsymbol{P}^{xx}_{k|k-1}]^T, \tag{6.40}$$

$$\boldsymbol{R}_{e,k} = \begin{bmatrix} \boldsymbol{R}_k + \boldsymbol{H}_k\boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k^T & (\boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k^T)^T \\ \boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k^T & -\gamma^2\boldsymbol{I} + \boldsymbol{P}^{xx}_{k|k-1} \end{bmatrix} \tag{6.41}$$

$$= \begin{bmatrix} \boldsymbol{R}_k + \boldsymbol{P}^{zz}_{k|k-1} & [\boldsymbol{P}^{xz}_{k|k-1}]^T \\ \boldsymbol{P}^{xz}_{k|k-1} & -\gamma^2\boldsymbol{I} + \boldsymbol{P}^{xx}_{k|k-1} \end{bmatrix}, \tag{6.42}$$

and where $\boldsymbol{I}$ is an identity matrix.

To show the relationship between the tuning parameter γ and the estimation error covariance matrix, we apply the matrix inversion lemma to (7.59), yielding

$$(\boldsymbol{P}^{xx}_{k|k})^{-1} = (\boldsymbol{P}^{xx}_{k|k-1})^{-1} + [\boldsymbol{H}_k^T\ \boldsymbol{I}]\begin{bmatrix} \boldsymbol{R}_k & \boldsymbol{0} \\ \boldsymbol{0} & -\gamma^2\boldsymbol{I} \end{bmatrix}^{-1}[\boldsymbol{H}_k^T\ \boldsymbol{I}]^T \tag{6.43}$$

$$= (\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k - \gamma^{-2}\boldsymbol{I}. \tag{6.44}$$

Thus, when γ tends to infinity, the H-infinity UKF will reduce to the traditional UKF. Therefore, γ can be seen as a tuning parameter to balance the tradeoff between the H-infinity and the minimum mean-square error performance. Extensive simulations reveal that the H-infinity UKF always achieves a good performance as long as the chosen γ is not very large. The suggested values for power system DSE vary between 10 and 20.

Another way to tune the parameter γ is motivated by the fact that the error covariance matrix $\boldsymbol{P}^{xx}_{k|k}$ must be positive-definite so that the sigma points can be generated [121,123,125]. This means that $(\boldsymbol{P}^{xx}_{k|k})^{-1} + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k - \gamma^{-2}\boldsymbol{I} \succeq \boldsymbol{0}$. In other words, the parameter γ should satisfy

$$\gamma^2 \geq \max\{eig(\boldsymbol{P}^{-1}_{k|k-1} + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k)^{-1}\}, \tag{6.45}$$

where $eig(\boldsymbol{A})^{-1}$ denotes the eigenvalue of matrix $\boldsymbol{A}^{-1}$. Interestingly, we find that both ways of tuning γ yield negligible differences on the estimation results.

Remark: the system model is always with uncertainties and the measurement noise statistics may change from time to time, yielding unknown and time-varying $\boldsymbol{Q}$ and $\boldsymbol{R}$. It is thus difficult to tune them for the traditional UKF in a real-time manner. By contrast, our proposed H-infinity UKF adopts the min-max criteria to bound these uncertainties and does not require the online tuning of them.

6.4 Numerical Results

We carry out extensive simulations on the IEEE 39-bus system to demonstrate the performance of the proposed DSE subject to various types of uncertainties. The two-axis model provided with the IEEE-DC1A exciter and the TGOV1 turbine-governor is assumed for each generator, where the parameter values can be found in [70]. Here, the transient stability simulations are performed to generate measurements and true state variables using the Matlab-based software PST [114] with some modifications. The 4th order Ruger-Kutta approach is adopted with an integration step of t=1/120 s to solve differential and algebraic equations, which complies with the time requirement found in the literature. The simulations consist of the following steps: Line 15-16 is tripped at t=0.5s to simulate a system disturbance; the voltage phasor, current phasor and frequency at each generator's terminal bus are corrupted by additive noise to simulate realistic PMU measurements; the scan rate of the PMU measurements is 20ms; note that if the integration step of the dynamic simulation is different from the PMU sampling rate, interpolation technique [130] is used to coordinate the simulations and real-time measurements. Of course, this would induce some errors. But, they have negligible effects on the DSE as demonstrated in [130]. On the other hand, these errors can be regarded as model uncertainties and are bounded by our proposed H-infinity filter. Furthermore, the computing speed of our approach is smaller than the typical PMU scan rate, i.e., 30-120 samples/s, which will be shown by simulation results. If the PMU sampling rate is 120 samples/s, no interpolation technique is required; the tuning parameter of the H-infinity filter is 10; the root-mean-squared error (RMSE) of all estimated generator state variables is used as the overall performance index while the estimated state variables of Generator 5 are taken for illustration. The RMSE is defined as follows:

$$RMSE = \frac{1}{N_m}\sum_{i=1}^{N_m}\sqrt{\frac{1}{N_s}\sum_{j=1}^{N_s}\left(\widehat{x}_j^i - x_j^i\right)^2}, \tag{6.46}$$

where N_m is the number of Monte Carlo runs, which is selected to be 100 in this chapter; x_j^i is the j-th true state variable at the ith run obtained from time-domain simulation results while $\widehat{x}_j^i$ is its state estimate; N_s is the total number of state variable at each run. Note that the rotor angle is in degree while all other state variables are in p.u. The reasons of choosing RMSE as the overall index are: i) To show the performance of different state variables, all state variables may need to be displayed, which in turn would require a lot of space; the use of RMSE is thus a convenient way to save space without losing its capability of assessing the overall performance of each estimator; ii) most importantly, our proposed H-infinity UKF is theoretically guaranteed to outperform other techniques that are not based on robust control theory. In other words, if it outperforms other approaches from the RMSE prospective, it is highly likely that every state variable obtained by the H-infinity UKF has a better performance. This has been verified by simulation results, such as Figs. 6.5 and 6.8. Three approaches are considered and compared, namely DUKFV, DUKFI and HDUKFI. They are defined as follows:

- DUKFV: decentralized unscented Kalman filter with measured terminal voltage phasor as model inputs while treating the current phasor and frequency as outputs/measurements; this is the existing approach proposed in [26];
- DUKFI: decentralized unscented Kalman filter with measured terminal current phasor as model inputs while treating the voltage phasor and frequency as outputs/measurements; this is one of our proposed approach but not based on H-infinity criteria;
- HDUKFI: proposed decentralized H-infinity unscented Kalman filter with measured terminal current phasor as model inputs while treating the voltage phasor and frequency as outputs/measurements.

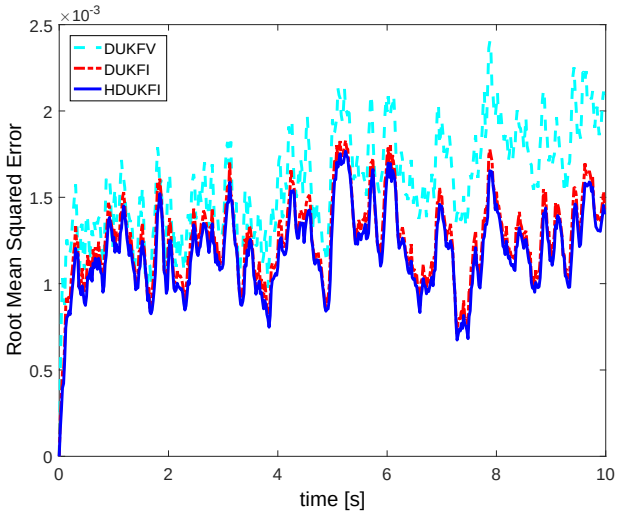


Figure 6.1: Root-mean-squared error of the DUKFV, the DUKFI and the HDUKFI in the presence of normal measurement noise.

6.4.1 Case 1: Known Gaussian Process and Measurement Noises

To test the sensitivity of each method to known Gaussian noises, Gaussian random variables with zero means and different covariance matrices are assumed for process and measurement noises. Specifically, normal and large noises are considered; they are assumed to be distributed with zero means and covariance matrices $\boldsymbol{Q} = 10^{-6}\boldsymbol{I}_{9\times 9}$ and $\boldsymbol{R} = 10^{-6}\boldsymbol{I}_{3\times 3}$, and $\boldsymbol{Q} = 10^{-4}\boldsymbol{I}_{9\times 9}$ and $\boldsymbol{R} = 10^{-4}\boldsymbol{I}_{3\times 3}$, respectively. The results are shown in Figs. 6.1-6.2, where Figs. 6.1 and 6.2 display the root-mean-squared errors of the DUKFV, the DUKFI and the HDUKFI in the presence of normal and large noises, respectively. The conclusions are as follows. Our HDUKFI achieves the highest statistical efficiency among the three approaches in both scenarios, followed by the DUKFI. The DUKFV is the most sensitive method to noises and has significantly biased estimation results in presence of large noises. In addition,

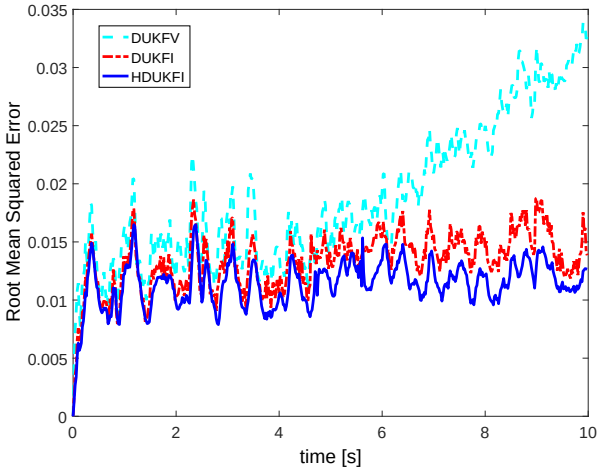


Figure 6.2: Root-mean-squared error of the DUKFV, the DUKFI and the HDUKFI in the presence of large measurement noise.

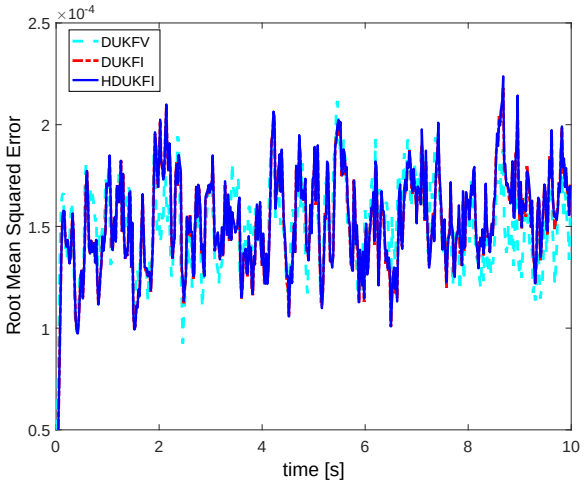


Figure 6.3: Root-mean-squared error of the DUKFV, the DUKFI and the HDUKFI with unknown gaussian process and measurement noises. They are assumed to be normally distributed with zero means and covariance matrices $\boldsymbol{Q} = 10^{-4}\boldsymbol{I}_{9\times 9}$ and $\boldsymbol{R} = 10^{-6}\boldsymbol{I}_{3\times 3}$, respectively.

the HDUKFI performs slightly better than the DUKFI with small noises. When the noise level increases, the DUKFI has an larger increase in estimation error than the HDUKFI. The latter has a bounded estimation error thanks to the H-infinity criterion.

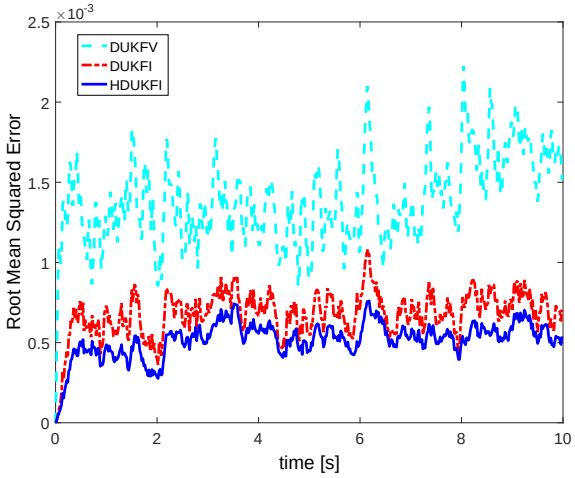


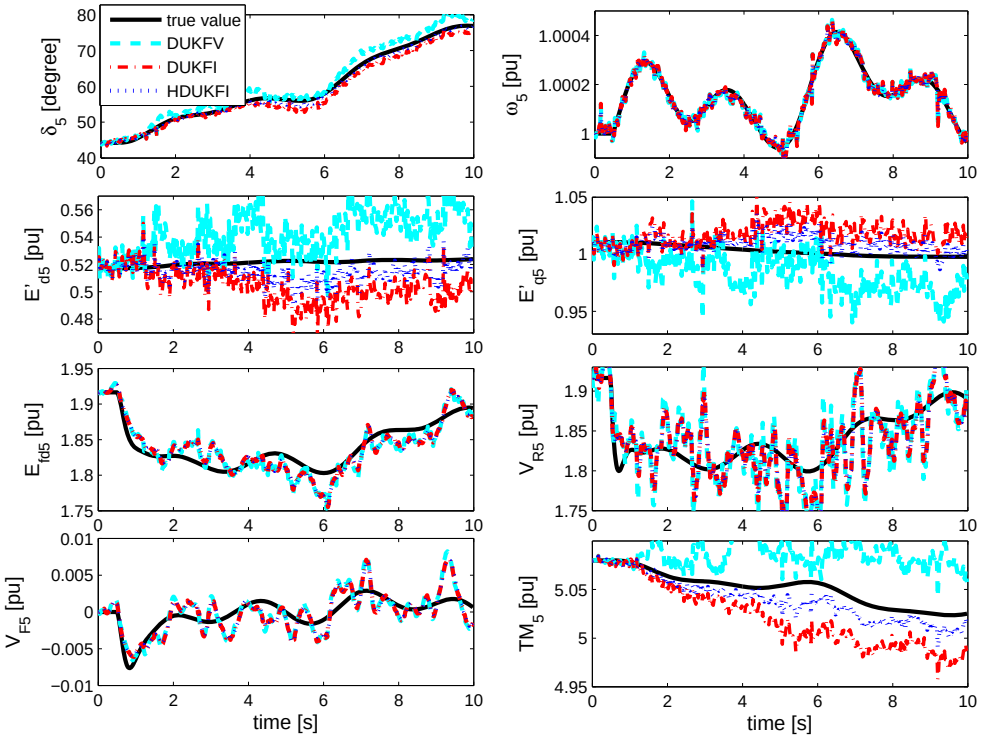
Figure 6.4: Root-mean-squared error of the DUKFV, the DUKFI and the HDUKFI with unknown gaussian process and measurement noises. They are assumed to be normally distributed with zero means and covariance matrices $\boldsymbol{Q} = 10^{-6}\boldsymbol{I}_{9\times 9}$ and $\boldsymbol{R} = 10^{-4}\boldsymbol{I}_{3\times 3}$, respectively.

6.4.2 Case 2: Unknown Gaussian Process and Measurement Noises

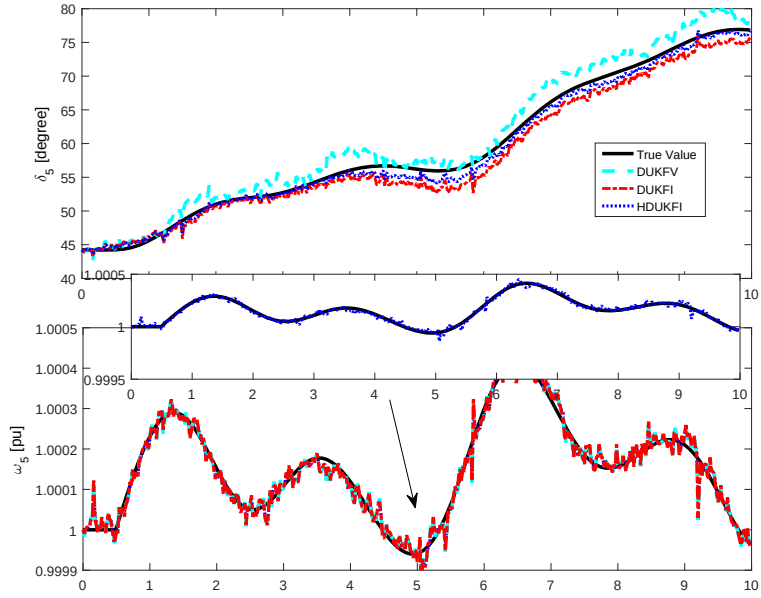
In practical power systems, the process and measurement noises can vary under different system operating conditions, yielding deviations from nominal values. In such cases, they are unknown to the operator. There are two scenarios: (i) the process noise is larger than the measurement noise and (ii) vice versa. In Scenario (i), the process and the measurement noises are assumed to be normally distributed with zero means and covariance matrices $\boldsymbol{Q} = 10^{-4}\boldsymbol{I}_{9\times 9}$ and $\boldsymbol{R} = 10^{-6}\boldsymbol{I}_{3\times 3}$, respectively, while in Scenario ii), $\boldsymbol{Q} = 10^{-6}\boldsymbol{I}_{9\times 9}$ and $\boldsymbol{R} = 10^{-4}\boldsymbol{I}_{3\times 3}$ are assumed for process and measurement noises, respectively. Note that the true covariance matrices are $\boldsymbol{Q} = 10^{-6}\boldsymbol{I}_{9\times 9}$ and $\boldsymbol{R} = 10^{-6}\boldsymbol{I}_{3\times 3}$. Simulation results of the DUKFV, the DUKFI and the HDUKFI for Scenarios i) and ii) are displayed in Figs. 6.3 and 6.4, respectively. From these figures, we conclude that all the three methods achieve comparable performances in Scenario (i) while the HDUKFI outperforms the other two methods in Scenario (ii). In summary, the HDUKFI can bound the influence of unknown process and measurement noises and provide good statistical efficiency.

6.4.3 Case 3: Non-Gaussian Process and Measurement Noise

Due to communication channel noises, the measurement noises associated with PMUs may deviate from the Gaussian assumption. In addition, since the PMU measurements are taken as system inputs, the system process noise is no longer Gaussian. To assess the sensitivity of each method to non-Gaussian process and measurement noise, we assume they follow a Gaussian mixture model, where 5% of the data are drawn with covariance matrices $\boldsymbol{Q} =$



(a)



(b)

Figure 6.5: (a) Estimated state variables of Generator 5 by the DUKFV, the DUKFI and the HDUKFI in the presence of non-Gaussian process and measurement noises; (b) Zoom in rotor angle and speed estimation.

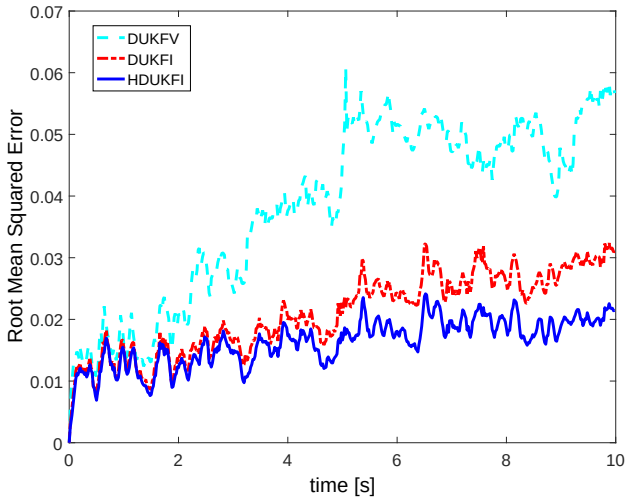


Figure 6.6: Root-mean-squared error of the DUKFV, the DUKFI and the HDUKFI in the presence of non-Gaussian process and measurement noises. They are assumed to follow a Gaussian mixture model, where 5% of the data are drawn with covariance matrices $\boldsymbol{Q} = 10^{-5}\boldsymbol{I}_{9\times 9}$ and $\boldsymbol{R} = 10^{-5}\boldsymbol{I}_{3\times 3}$ while the true covariance matrices are $\boldsymbol{Q} = 10^{-6}\boldsymbol{I}_{9\times 9}$ and $\boldsymbol{R} = 10^{-6}\boldsymbol{I}_{3\times 3}$

.

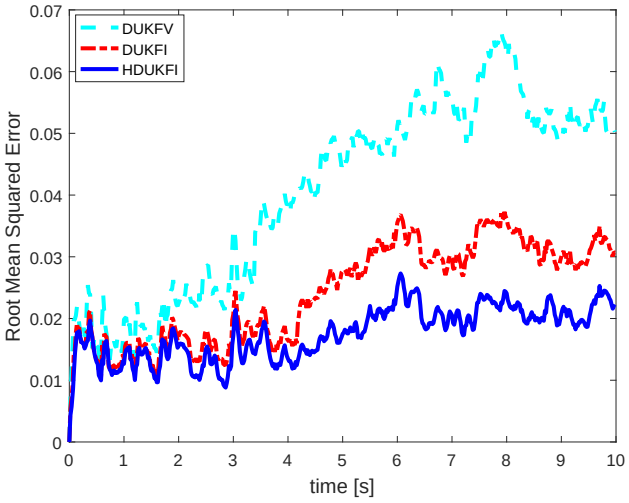


Figure 6.7: Root-mean-squared error of the DUKFV, the DUKFI and the HDUKFI when process and measurement noises follow a Gaussian mixture model, where 10% of the data are drawn with covariance matrices $\boldsymbol{Q} = 10^{-5}\boldsymbol{I}_{9\times 9}$ and $\boldsymbol{R} = 10^{-5}\boldsymbol{I}_{3\times 3}$ while the true covariance matrices are $\boldsymbol{Q} = 10^{-6}\boldsymbol{I}_{9\times 9}$ and $\boldsymbol{R} = 10^{-6}\boldsymbol{I}_{3\times 3}$

.

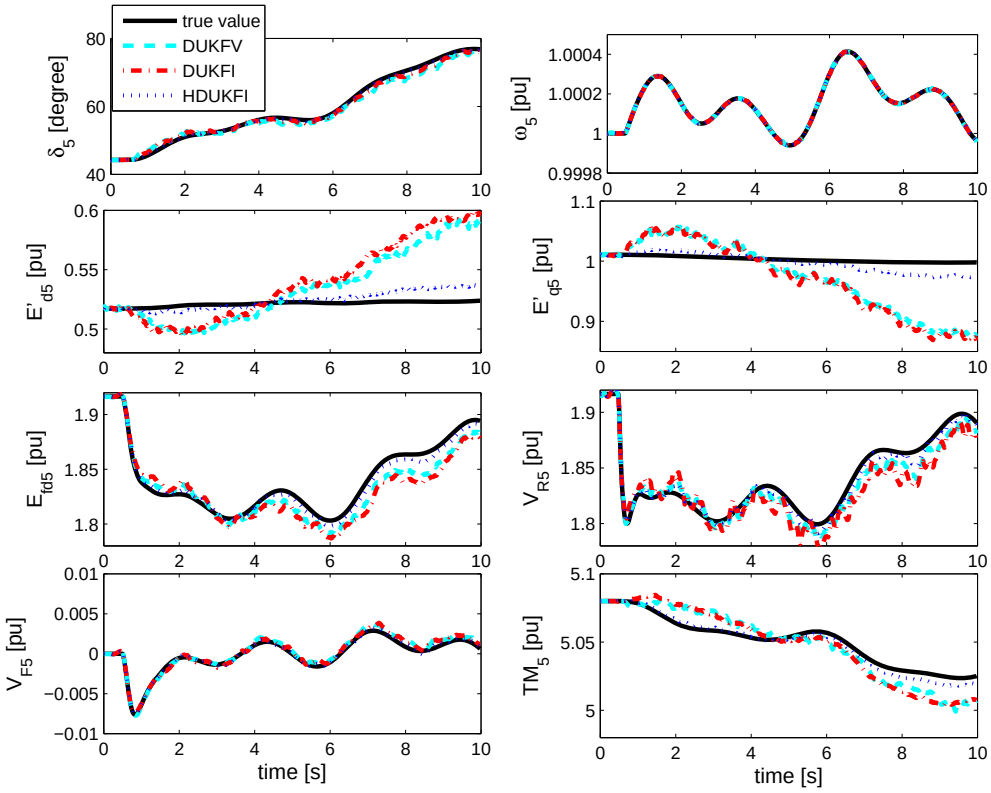


Figure 6.8: Estimated state variables of Generator 5 by the DUKFV, the DUKFI and the HDUKFI with model uncertainties.

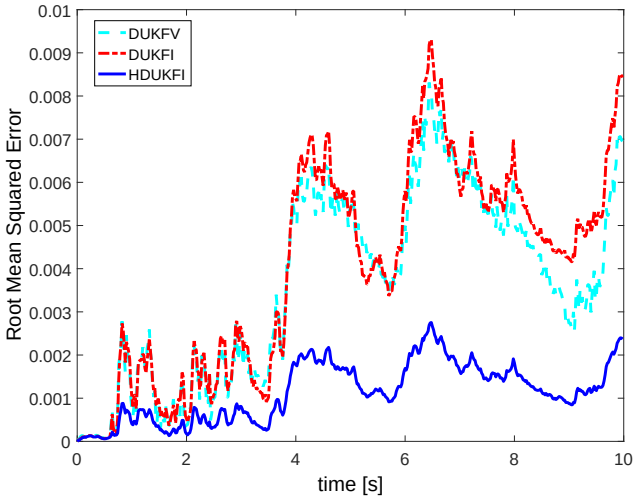


Figure 6.9: Root-mean-squared error of the DUKFV, the DUKFI and the HDUKFI with model uncertainties. Here, it is assumed that after the disturbance is applied, the transient reactance of Generator 5 deviate from the nominal values by a percentage of 10%.

$10^{-5}\boldsymbol{I}_{9\times9}$ and $\boldsymbol{R} = 10^{-5}\boldsymbol{I}_{3\times3}$ while the true covariance matrices are $\boldsymbol{Q} = 10^{-6}\boldsymbol{I}_{9\times9}$ and $\boldsymbol{R} = 10^{-6}\boldsymbol{I}_{3\times3}$. The estimated state variables of Generator 5 as well as the root-mean-squared errors of each method are displayed in Figs. 6.5-6.6. It can be observed that the DUKFV is very sensitive to non-Gaussian noises and provides the worst estimation results. By contrast, the DUKFI performs better than the DUKFV, but its statistical efficiency is much lower than our HDUKFI. For example, from the zoomed in figure that represents the rotor angle and speed estimates, we find that the difference can be several degrees, which is a lot and can significantly affect the intended applications. When a higher percentage of data deviate from the Gaussian assumption, say 10% data is drawn with covariance matrices $\boldsymbol{Q} = 10^{-5}\boldsymbol{I}_{9\times9}$ and $\boldsymbol{R} = 10^{-5}\boldsymbol{I}_{3\times3}$ while the true covariance matrices are $\boldsymbol{Q} = 10^{-6}\boldsymbol{I}_{9\times9}$ and $\boldsymbol{R} = 10^{-6}\boldsymbol{I}_{3\times3}$, the root-mean-squared error of each method is shown in Fig. 6.7. It is observed that with the increased degree of deviation from the Gaussian assumption, the estimation error of the DUKFI increases significantly, which demonstrates its sensitivity to non-Gaussian noise. By contrast, the increased estimation error of our HDUKFI remains small thanks to the H-infinity criterion. In summary, our HDUKFI is able to bound the influence of non-Gaussian process and measurement noise and provides the best state estimates.

6.4.4 Case 4: Robustness to Dynamical Model Uncertainties

The main scope of the paper is to address unknown non-Gaussian noise and model uncertainties that are induced by uncertain generator or control parameters. The results shown in Figs. 6.1-6.4 are to demonstrate the benefits of the proposed over others under other various

operation conditions. In this section, the robustness of the proposed estimator to dynamical model uncertainties is tested.

Due to aging processes, variations of the machine temperature during its operation, the effect of saturation on generator inductances, to cite a few, the reactance and transient reactance of a synchronous generator can change with time. In our simulations, we assume that after the disturbances, the transient reactance of the generators deviate from the nominal values by a percentage of 10%. Figs. 6.8-6.9 display the root-mean-squared errors of the DUKFV, the DUKFI and the HDUKFI with model uncertainties and the estimated state variables of Generator 5, respectively. Unlike the cases of unknown noises, both the DUKFV and the DUKFI exhibit a high sensitivity to dynamical model uncertainties and provide significantly biased estimation results; the reader is referred to the estimated d- and q-axis voltages of Generator 5 as an example. By contrast, our HDUKFI bounds the uncertainties and provides reasonably good state estimates. Further explanations of these conclusions can be stated as follows:

- In theory, both EKF and UKF assume that the dynamical model is accurate and the system noise statistics, and system inputs are known. Then, they can produce good estimation results. If these assumptions are violated, the state estimates can be significantly biased [121, 124]. Our results are consistent with these theoretical conclusions;

- According to the work [43, 132, 133, 143] in terms of generator model validation and calibration, it is found that only a few key parameters play the major role in the responses of the generator. In other words, changes among these parameters can affect the generator response significantly while the changes on other parameters induce negligible difference. Furthermore, the sensitivity of parameters can change in presence of different system disturbances. This has been validated and emphasized in NERC standard [133] when performing parameter calibrations. Last but not the least, both generator and controller parameters can have high sensitivities. Following [43, 132, 133, 143], we carry out trajectory sensitivity analysis to Generator 5 with and without parameter uncertainties on generator d- and q axis transient reactance, exciter gain K_A and governor T_{SV}, respectively. Due to space limitation, the detailed results are not shown here. The results show that the former two groups of parameters play the major role in the responses of the generator. This demonstrates why uncertainties on generator d- and q axis transient reactance can significantly affect EKF and UKF.

Since the controller parameters can change from time to time [128, 129] due to aging process, inappropriate tuning and controller failures, to name a few, we will investigate here the influence of the exciter gain K_A uncertainties on the state estimates. Fig. 6.10 displays the root-mean-squared errors of the DUKFV, the DUKFI and the HDUKFI when the exciter is with 10% uncertainties after the disturbance. It can be found that our HDUKFI still

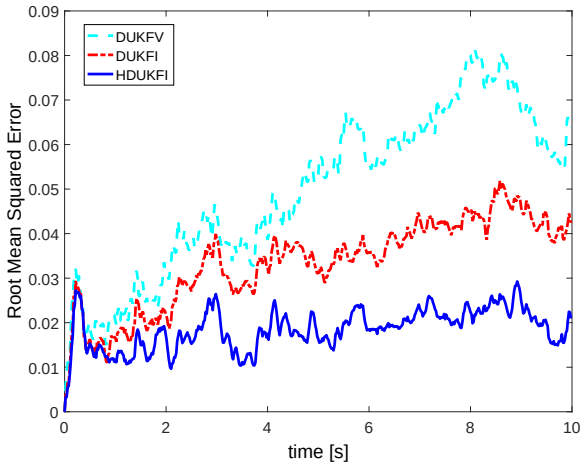


Figure 6.10: Root-mean-squared error of the DUKFV, the DUKFI and the HDUKFI with model uncertainties. Here, it is assumed that after the disturbance is applied, the exciter gain of Generator 5, K_A, deviates from the nominal values by a percentage of 10%.

outperforms the DUKFV and the DUKFI, which is consistent with the case that generator d- and q axis transient reactances are subject to uncertainties.

Furthermore, it can be found that the estimation accuracy of the internal state variables, i.e., rotor speeds and angles, depends on the operating conditions. For example, all the investigated approaches yield reasonable estimations of the generator rotor speeds and angles in presence of model uncertainties while the DUKFV and the DUKFI produce quite poor results for other state variables (see Fig. 6.8). On the other handle, the DUKFV and the DUKFI yield relatively large estimation errors of rotor speeds and angles in presence of unknown non-Gaussian noise (see Fig. 6.5). By contrast, our proposed HDUKFI outperforms all other approaches, yielding quite reasonable estimates in different conditions. The estimation of TM_5 is acceptable for our proposed HDUKFI while other approaches produce large biases. Note that the speed governor has a large time constant and will slowly respond to system disturbances. This may cause a slow drift of the estimation slowly. However, this is not a problem for our proposed HDUKFI since it can quickly re-track TM_5 (see Figs. 6.5 and 6.8).

6.4.5 Comparisons with Other Approaches

In this section, we carry out several representative tests to demonstrate the advantages of our HDUKFI over other existing robust estimation approaches, such as the GM-IEKF [38] and the H-infinity EKF (HEKF) [125]. Due to space limitation, we only test the GM-IEKF and the HEKF with model uncertainties, where the simulations in Case 4 are used. The results are shown in Figs. 6.11-6.12. It can be found that compared with the traditional UKF, the

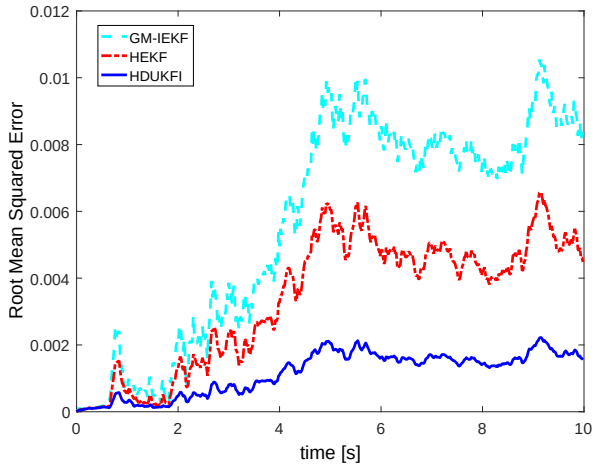


Figure 6.11: Root-mean-squared error of the GM-IEKF, the HEKF and the HDUKFI with model uncertainties. Here, it is assumed that after the disturbance is applied, the transient reactance of Generator 5 deviates from the nominal value by a percentage of 10%.

GM-IEKF and the HEKF have better capability of handling the model uncertainties. Thanks to the H-infinity criteria from robust control theory in bounding the influence of model and noise uncertainties, the HEKF achieves better statistical efficiency than the robust GM-IEKF. However, our HDUKFI still outperforms the GM-IEKF and the HEKF. It worth pointing out that, similar to Figs. 6.5 and 6.8, the proposed HDUKFI outperforms GM-IEKF and HEKF for each state variable. Due to space limitation, they are not shown here.

6.4.6 Assessment of Computational Efficiency

To assess the computational efficiency of each method, the average computing times of the DUKFV, the DUKFI and the HDUKFI for Cases 1-4 at each PMU scan are presented in Table. 6.1. All the tests are performed on a PC with Intel Core i5, 2.50 GHz, 8GB of RAM. From the table, we find that the DUKFV and the DUKFI have comparable computational efficiency in all cases. The proposed HDUKFI is the most time consuming approach as it proceeds with more matrix calculations. However, its computing time is much smaller than the PMU scan rate, which is 20ms for 50 sample/s. In addition, the computing times of all three methods vary slightly in different cases. In summary, the HDUKFI is able to keep up with the PMU scan rate and suitable for large-scale power system online applications.

6.4.7 Practical Values of the Proposed HUKF

For practical power systems, the estimated dynamic state variables can be used for many different applications, such as development of local and wide-area oscillation damping con-

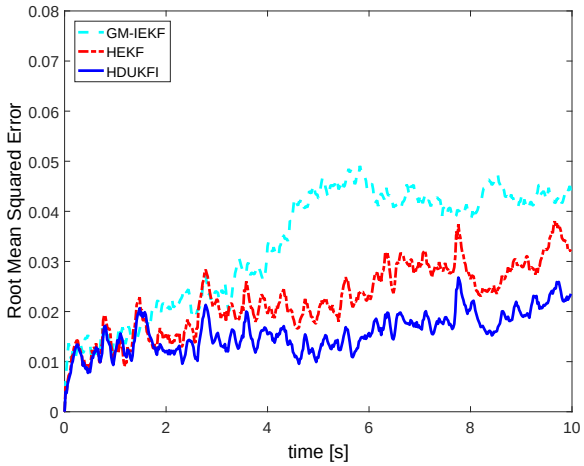


Figure 6.12: Root-mean-squared error of the GM-IEKF, the HEKF and the HDUKFI with model uncertainties. Here, it is assumed that after the disturbance is applied, the exciter gain of Generator 5, K_A, deviates from the nominal values by a percentage of 10%.

Table 6.1: Average Computing Times of The Three Decentralized Unscented Kalman Filters At Each PMU Scan

Cases	DUKFV	DUKFI	HDUKFI
Case 1	0.37ms	0.32ms	1.19ms
Case 2	0.36ms	0.31ms	1.15ms
Case 3	0.38ms	0.35ms	1.20ms
Case 4	0.39ms	0.37ms	1.22ms

trollers [91], design of adaptive out-of-step protection scheme [134], model validation [43] and performing dynamic security assessment, to name a few. Note that many dynamic state variables are not measured by PMUs, such as field voltage, generator rotor angle, mechanical power and so on. With the dynamic state estimator, all of them can be estimated online and used for further power grid monitoring, control and protection.

In principle, the more accurate the estimated state variables are, the more effective the developed damping controllers, adaptive out-of-step protection scheme, etc, will be. For example, with poorly estimated rotor speeds and angles, the decentralized damping controller may be ineffective and sometimes may induce adverse affects [91]. In addition, they may lead to incorrect conclusions of the rotor angle stability and the failure of the out-of-step protection scheme [134]. We would like to emphasize that there can be large differences for the intended applications if the accuracy of the estimated dynamic state variables show differences. For instance, the design of adaptive out-of-step protection schemes and damping controllers may have much higher requirements on the state estimation accuracy than the generator model parameter calibrations. On the other hand, our proposed H-infinity UKF

significantly outperforms other approaches in presence of large model uncertainties, demonstrating its capability for wide-area applications. The design of power system applications based on the estimated dynamic state variables is beyond the scope of the paper. We will investigate that in our future work.

6.5 Conclusions

In this chapter, a decentralized H-infinity unscented Kalman filter is proposed for power system DSE with various types of uncertainties, including unknown process and measurement noises, deviation from Gaussian assumptions and uncertain generator parameters. By applying the statistical linerization approach and using a batch-mode regression model, we derive a linear H-infinity Kalman filter formula along with an H-infinity UKF. We then utilize model decoupling approaches to enable the decentralized estimation of generator state variables through local PMU measurements. Extensive simulation results carried out on the IEEE 39-bus system demonstrate that our H-infinity UKF outperforms the existing decentralized UKF-based DSE in terms of statistical efficiency and robustness to uncertainties. In the future work, we will investigate how to further improve the robustness of our H-infinity UKF by resorting to the robust statistical theory.

Chapter 7

Robust H-infinity Unscented Kalman Filter: Theory and Application to Power System Dynamic State Estimation

7.1 Introduction

7.1.1 Motivation

The enhancement of the reliability, security, and resiliency of electric power systems depends on the availability of a fast, accurate, and robust dynamic state estimator (DSE) that processes both model information and online measurements obtained from phasor measurement units (PMUs). A DSE provides a better wide-area situation awareness of the system dynamics, leading to improved system oscillation monitoring and controls, enhanced dynamic security assessment and system protection schemes [91, 134, 135], among others. Therefore, it is of a critical importance that any power system DSE is robust to gross errors on the measurements and the model parameter values while providing good state estimates in the presence of large dynamical system model uncertainties and non-Gaussian thick-tailed process and observation noises.

It turns out that most of the existing DSEs assume an exact dynamical system model, accurate measurements and known Gaussian system process and measurement noise, etc. However, due to unknowable system inputs, including noise, parameter variations and actuator failures [136] and inaccuracies of the model parameter values of the synchronous generators, the loads, the lines, and the transformers, to name a few, the system is always subject to uncertainties. On the other hand, the process and the observation noise of the

system nonlinear dynamic models are non-Gaussian as verified by the Pacific Northwest National Lab (PNNL) [35, 81]. Finally, outliers can occur that may corrupt the state estimates [37]. Two types of outliers are commonly seen in power systems, namely innovation and observation outliers. Observation outliers may result from large biases in PMU measurements due to infrequent calibration, or instrument failures, or impulsive communication noise [40,41]. As for innovation outliers, they may occur in several different ways. For example, some of the generator models may not be well calibrated, resulting in highly inaccurate model outputs that are inconsistent with the measurements. This was precisely the case in the 1996 blackout, where the model being used predicted system stability while in reality the system was undergoing numerous cascading failures, which resulted in a rapid system collapse that occurred within minutes [42, 43]. Innovation outliers may also be induced by the approximations in the state prediction model or by a system process impulsive noise.

To bound the influence of model uncertainties and unknown system process and measurement noise, the first-order Taylor series approximations-based H-infinity extended Kalman filter and the unscented transformation-based H-infinity unscented Kalman filter (UKF) have been proposed [125,126]. However, they are vulnerable to non-Gaussian noise and any types of outliers. To address gross errors on the PMU measurements, several robust DSEs are developed. For instance, in [27], Rouhani and Abur propose a distributed two-stage robust UKF-based DSE using the least-absolute-value (LAV) estimator that can handle observation outliers. However, the authors do not address the vulnerability of the DSE to innovation outliers. To handle both types of outliers, generalized-maximum likelihood type iterated EKF and UKF are developed in [38, 88, 120]. However, their statistical efficiencies are low in the presence of non-Gaussian system process and measurement noise. To effectively suppress outliers while achieving a high statistical efficiency under a broad range of non-Gaussian process and observation noise, the generalized-maximum likelihood type UKF (GM-UKF) is proposed [79,92]. Its robustness and statistical efficiency for different noise distributions are analyzed analytically. The main weakness of the GM-UKF is that it may yield biased state estimates when the dynamical system has large uncertainties that are induced by inaccurate model parameters and system inputs.

7.1.2 Contributions and Paper Organization

To address the aforementioned challenges, this chapter presents a general theoretical framework that integrates both robust statistics and robust control theory to develop a robust dynamic state estimator for power systems. Specifically, it yields the following contributions:

- The GM-estimator, the UKF, and the H-infinity filter are integrated into a unified framework to yield the general robust H-infinity filter for nonlinear systems in the Krein space. Since the GM-estimator and the H-infinity filter stem from robust statistics and robust control theory, respectively, other combinations of robust estimators and filters

can be easily developed to obtain a good balance between estimation efficiency and robustness to outliers and model uncertainties.

- The proposed robust H-infinity UKF is able to handle large system uncertainties as well as suppress outliers while achieving good statistical efficiency under a broad range of non-Gaussian process and observation noises. Specifically, it leverages the H-infinity criterion to bound system uncertainties while relying on the robustness of GM-estimator to suppress outliers and filter out non-Gaussian noise.

- The statistical linerization is used to build a linear-like batch-mode regression form in the Krein space. We show that the H-infinity UKF is just the Krein space Kalman filter that minimizes the least squares criterion and has a bounded estimation error in presence of system uncertainties. However, it is vulnerable to outliers and non-Gaussian noise. It is worth pointing out that this is the first attempt to derive the H-infinity UKF in the Krein space.

- A theoretical analysis is performed to demonstrate that in the absence of non-Gaussian noise and outliers, the proposed robust H-infinity UKF reduces to the H-infinity UKF; on the other hand, if the tuning parameter of the H-infinity criterion tends to infinity, the proposed robust H-infinity UKF resembles the GM-UKF. This provides the justification why both the benefits of the H-infinity filter and the GM-estimator are maintained in the proposed robust H-infinity UKF.

- A comparative analysis is carried out, which shows that our proposed robust H-infinity UKF outperforms the H-infinity UKF and the GM-UKF in terms of statistical efficiency and robustness to outliers, non-Gaussian noise and model uncertainties.

The remainder of the paper is organized as follows. Section II presents the problem formulation. Section III describes the theory of the proposed theoretical framework and Section IV shows and analyzes the simulation results carried out on the IEEE 39-bus system. Finally, Section V concludes the paper.

7.2 Problem Formulation

In general, a dynamical system can be described by a set of continuous-time nonlinear differential and algebraic equations. These equations can be further discretized into the following discrete-time state space form:

$$\boldsymbol{x}_k = \boldsymbol{f}\left(\boldsymbol{x}_{k-1}, \boldsymbol{u}_k\right) + \boldsymbol{w}_k, \tag{7.1}$$

$$\boldsymbol{z}_k = \boldsymbol{h}\left(\boldsymbol{x}_k, \boldsymbol{u}_k\right) + \boldsymbol{v}_k, \tag{7.2}$$

where $\boldsymbol{x}_k \in \mathbb{R}^{n\times 1}$ and $\boldsymbol{z}_k \in \mathbb{R}^{m\times 1}$ are the state vector and the measurement/observation vector at time sample k, respectively; $\boldsymbol{f}$ and $\boldsymbol{h}$ are vector-valued nonlinear functions; $\boldsymbol{w}_k$

and $\boldsymbol{v}_k$ are the system process and observation noise, respectively; they are assumed to be independent and identically distributed with zero mean and covariance matrices $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$, respectively; $\boldsymbol{u}_k$ is the system input vector.

As an illustrative example, let us consider the synchronous generator model. If the two-axis generator model with the IEEE-DC1A exciter and the TGOV1 turbine-governor is considered, its dynamical model can be expressed by the following differential and algebraic equations [70]:

Differential equations:

$$T'_{do}\frac{dE'_q}{dt} = -E'_q - \left(X_d - X'_d\right) I_d + E_{fd}, \tag{7.3}$$

$$T'_{qo}\frac{dE'_d}{dt} = -E'_d - \left(X_q - X'_q\right) I_q, \tag{7.4}$$

$$\frac{d\delta}{dt} = \omega - \omega_s, \tag{7.5}$$

$$\frac{2H}{\omega_s}\frac{d\omega}{dt} = T_M - P_e - D\left(\omega - \omega_s\right), \tag{7.6}$$

$$T_E\frac{dE_{fd}}{dt} = -\left(K_E + S_E\left(E_{fd}\right)\right) E_{fd} + V_R, \tag{7.7}$$

$$T_{\mathrm{F}}\frac{dV_F}{dt} = -V_F + \frac{K_F}{T_E}V_R - \frac{K_F}{T_E}\left(K_E + S_E\left(E_{fd}\right)\right) E_{fd}, \tag{7.8}$$

$$T_A\frac{dV_R}{dt} = -V_R + K_A\left(V_{ref} - V_F - V\right), \tag{7.9}$$

$$T_{CH}\frac{dT_M}{dt} = -T_M + P_{SV}, \tag{7.10}$$

$$T_{SV}\frac{dP_{SV}}{dt} = -P_{SV} + P_C - \frac{1}{R_D}\left(\frac{\omega}{\omega_s} - 1\right), \tag{7.11}$$

Algebraic equations:

$$V_d = V\sin\left(\delta - \theta\right), V_q = V\cos\left(\delta - \theta\right), \tag{7.12}$$

$$I_d = \frac{E'_q - V_q}{X'_d}, I_q = \frac{V_d - E'_d}{X'_q}, \tag{7.13}$$

$$P_e = V_d I_d + V_q I_q, Q_e = -V_d I_q + V_q I_d, \tag{7.14}$$

where T'_{do}, T'_{qo}, T_E, T_F, T_A, T_{CH} and T_{SV} are time constants, in seconds; K_E, K_F and K_A are controller gains; V_{ref} and P_C are known control inputs; E'_q, E'_d, E_{fd}, V_F, V_R, T_M and P_{SV} are the q-axis and d-axis transient voltages, field voltage, scaled output of the stabilizing transformer and scaled output of the amplifier, synchronous machine mechanical torque and steam valve position, respectively; X_d, X'_d, X_q and X'_q are generator parameters; V and θ are

the terminal bus voltage magnitude and phase angle, respectively; P_e and Q_e are the active and reactive electrical power outputs; I_d and I_q are the d and q axis currents, respectively.

By applying time discretization to (7.3)-(7.14) using numerical approaches, such as the 4th-order Ruger-Kutta method, we get (7.1)-(7.2), yielding the state vector given by $\boldsymbol{x}_k = [\delta\ \omega\ E'_d\ E'_q\ E_{fd}\ V_F\ V_R\ T_M\ P_{SV}]$. The relationships given by (7.3)-(7.11) and (7.12)-(7.14) are represented in compact forms by the vector-valued function $\boldsymbol{f}(\cdot)$ and by $\boldsymbol{h}(\cdot)$, respectively. The system input vector is denoted by $\boldsymbol{u}_k = [V_{ref}\ P_C]^T$. The measurement vector $\boldsymbol{z}_k$ contains a collection of voltage phasor measurements $V\angle\theta$ and real and reactive power injections P_e and Q_e, which are obtained by the PMUs.

To estimate the system dynamic state vector or model parameters using Kalman-type filters, a two-step procedure is applied, namely a prediction step using (7.1), which is a Markov model, and a filtering/update step using (7.2). Specifically, given a state estimate at time step $k-1$, $\widehat{\boldsymbol{x}}_{k-1|k-1}$, with its covariance matrix, $\boldsymbol{P}_{k-1|k-1}$, the predicted state vector is calculated from (7.1) directly or through a set of points drawn from the probability distribution of $\widehat{\boldsymbol{x}}_{k-1|k-1}$, which is dependent on the assumed probability distributions of $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$. As for the filtering step, the predictions are used together with the observations at time sample k to estimate the state vector and its covariance matrix.

The Kalman-type filters work well only under the validity of certain assumptions [121, 122]. First, the system process and observation noise, $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$, are assumed to have at each time instant zero means and known covariance matrices $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$, respectively. Secondly, they are assumed to follow a Gaussian distribution, at which the filter is optimal with minimum variances. Finally, the system model is assumed to be known exactly. However, for most practical dynamical systems, these assumptions do not hold. Indeed, $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$ are difficult to obtain in practice; the process and observation noise do not follow a Gaussian distribution as verified in [35, 81]; the functions $\boldsymbol{f}$ and $\boldsymbol{h}$ are approximate; for instance, they may not account for all the nonlinearities of the system; some model parameter values may be unknown or incorrect [87, 125]; and the received measurements may be strongly biased due to impulsive communication noise, cyber attacks, to cite a few [40, 41]. To address these challenges, this chapter presents a novel general theoretical framework that integrates both robust statistics and robust control theory for robust dynamic state estimation. The proposed robust H-infinity UKF within this framework is able to handle large system uncertainties as well as suppress outliers while achieving good statistical efficiency under a broad range of non-Gaussian process and observation noise.

7.3 Theory of Proposed Robust H-infinity UKF

In this section, the Krein space UKF will be derived first. Then its equivalence to the H-infinity UKF will be proved. It will be further shown that Krein space UKF has a bounded estimation error in presence of system uncertainties. By carrying out a theoretical analysis of the weakness of the Krein space UKF to non-Gaussian noise and outliers, we propose

a robust Krein space UKF, i.e., a robust H-infinity UKF, which combines the robustness of the H-infinity criterion to model uncertainties and of the GM-estimator to outliers and non-Gaussian noise.

7.3.1 Derivation of the Krein Space UKF

The statistical linerization is used to convert the traditional nonlinear UKF into an equivalent linear-like regression form. The latter is further derived into the Krein space batch-regression model that allows us to resort to the Kalman filter framework for the development of Krein space UKF.

Derivation of the Linear-Like Regression Form of the Nonlinear UKF

The main idea of UKF is to use a deterministic sampling technique known as the unscented transform to choose a set of sample points, termed sigma points, which have the same means and covariances as the a priori state statistics under the Gaussian assumption [113]. These sigma points are then propagated through the non-linear functions $\boldsymbol{f}$ and $\boldsymbol{h}$, yielding an estimation of the a posteriori state statistics by using the Kalman filter structure, i.e., the mean and covariance estimates. Consequently, no calculation of Jacobian matrices is required, which for complex functions can be a difficult task by itself, requiring complicated derivatives if done analytically or being computationally costly if done numerically. Specifically, given a state estimate with mean $\widehat{\boldsymbol{x}}_{k-1|k-1} \in \mathbb{R}^{n\times 1}$ and covariance matrix $\boldsymbol{P}^{xx}_{k-1|k-1}$ at time step k-1, $2n$ weighted sigma points are generated through

$$\boldsymbol{\chi}^{j}_{k-1|k-1} = \widehat{\boldsymbol{x}}_{k-1|k-1} \pm \left(\sqrt{n\boldsymbol{P}^{xx}_{k-1|k-1}}\right)_{j}, \tag{7.15}$$

where the weights $w_j = 1/2n, j = 1, ..., 2n$ and the term $\left(\sqrt{n\boldsymbol{P}^{xx}_{k-1|k-1}}\right)_{j}$ represents its jth column vector. These sigma points are propagated through the nonlinear system process model (7.1) to obtain the following transformed sigma points

$$\boldsymbol{\chi}^{j}_{k|k-1} = \boldsymbol{f}\left(\boldsymbol{\chi}^{j}_{k-1|k-1}\right). \tag{7.16}$$

Then, the predicted state vector $\widehat{\boldsymbol{x}}_{k|k-1}$ and its covariance matrix are approximated by the weighted sample mean and sample covariance matrix of the transformed sigma points, respectively, yielding

$$\widehat{\boldsymbol{x}}_{k|k-1} = \sum_{j=1}^{2n} w_j \boldsymbol{\chi}^{j}_{k|k-1}, \tag{7.17}$$

$$\boldsymbol{P}^{xx}_{k|k-1} = \sum_{j=1}^{2n} w_j (\boldsymbol{\chi}^{j}_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})(\boldsymbol{\chi}^{j}_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})^T + \boldsymbol{Q}_k. \tag{7.18}$$

Next, the sigma points are updated to capture the information of system process noise through

$$\boldsymbol{\chi}^{j}_{k|k-1} = \widehat{\boldsymbol{x}}_{k|k-1} \pm \left(\sqrt{n\boldsymbol{P}^{xx}_{k|k-1}}\right)_j. \tag{7.19}$$

Then, the predicted measurement vector is given by $\widehat{\boldsymbol{z}}_{k|k-1} = \sum_{j=1}^{2n} w_j \boldsymbol{h}(\boldsymbol{\chi}^{j}_{k|k-1})$ and its associated error covariance matrix is estimated via

$$\boldsymbol{P}^{zz}_{k|k-1} = \sum_{j=1}^{2n} w_j (\boldsymbol{z}^{j}_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1})(\boldsymbol{z}^{j}_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1})^T + \boldsymbol{R}_k, \tag{7.20}$$

where $\boldsymbol{z}^{j}_{k|k-1} = \boldsymbol{h}(\boldsymbol{\chi}^{j}_{k|k-1})$.

It is worth pointing out that the statistical linerization is equivalent to the unscented transformation in terms of propagating the probability distribution of the state vector through nonlinear function [54,79,92]. Thus, by applying statistical linerization to the nonlinear state transition and measurement/observation functions around $\widehat{\boldsymbol{x}}_{k-1|k-1}$ and $\widehat{\boldsymbol{x}}_{k|k-1}$, respectively, we get

$$\boldsymbol{x}_k = \boldsymbol{F}_k \left(\boldsymbol{x}_{k-1} - \widehat{\boldsymbol{x}}_{k-1|k-1}\right) + \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{e}_k + \boldsymbol{w}_k, \tag{7.21}$$

$$\boldsymbol{z}_k = \boldsymbol{H}_k \left(\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k-1}\right) + \widehat{\boldsymbol{z}}_{k|k-1} + \boldsymbol{\varepsilon}_k + \boldsymbol{v}_k, \tag{7.22}$$

where $\boldsymbol{F}_k = (\boldsymbol{P}^{x\chi}_{k|k-1})^T (\boldsymbol{P}^{xx}_{k-1|k-1})^{-1}$;

$$\boldsymbol{P}^{x\chi}_{k|k-1} = \sum_{j=1}^{2n} w_j (\boldsymbol{\chi}^{j}_{k-1|k-1} - \widehat{\boldsymbol{x}}_{k-1|k-1})(\boldsymbol{\chi}^{j}_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})^T, \tag{7.23}$$

and $\boldsymbol{e}_k$ is the statistical linearization error term that is normally distributed with zero mean and covariance matrix $\boldsymbol{L}_k = \boldsymbol{P}^{xx}_{k|k-1} - (\boldsymbol{P}^{x\chi}_{k|k-1})^T (\boldsymbol{P}^{xx}_{k-1|k-1})^{-1} \boldsymbol{P}^{x\chi}_{k|k-1}$; $\boldsymbol{H}_k = (\boldsymbol{P}^{xz}_{k|k-1})^T (\boldsymbol{P}^{xx}_{k|k-1})^{-1}$;

$$\boldsymbol{P}^{xz}_{k|k-1} = \sum_{i=1}^{2n} w_j (\boldsymbol{\chi}^{j}_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})(\boldsymbol{z}^{j}_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1})^T, \tag{7.24}$$

and $\boldsymbol{\varepsilon}_k$ is the statistical linearization error term on the nonlinear measurement function; it is normally distributed with zero mean and covariance matrix $\boldsymbol{\Pi}_k = \boldsymbol{P}^{zz}_{k|k-1} - (\boldsymbol{P}^{xz}_{k|k-1})^T (\boldsymbol{P}^{xx}_{k|k-1})^{-1} \boldsymbol{P}^{xz}_{k|k-1}$.

Therefore, the nonlinear dynamical model expressed by (7.1) and (7.2) is converted into the following equivalent linear-like regression form:

$$\boldsymbol{x}_k = \boldsymbol{F}_k \boldsymbol{x}_{k-1} + \widehat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{F}_k \widehat{\boldsymbol{x}}_{k-1|k-1} + \boldsymbol{e}_k + \boldsymbol{w}_k, \tag{7.25}$$

$$\boldsymbol{z}_k = \boldsymbol{H}_k \boldsymbol{x}_k + \widehat{\boldsymbol{z}}_{k|k-1} - \boldsymbol{H}_k \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{\varepsilon}_k + \boldsymbol{v}_k, \tag{7.26}$$

where $\boldsymbol{F}_k$ and $\boldsymbol{H}_k$ are no longer the Jacobian matrices, but $\boldsymbol{e}_k$ and $\boldsymbol{\varepsilon}_k$ are the statistical linearization errors that are used to preserve the nonlinearities of the state transition and the measurement functions, respectively.

Remark 1: by using the Linear Kalman filter framework, the recursive state estimation form can be derived as:

$$\boldsymbol{K}_k = \boldsymbol{P}^{xz}_{k|k-1}\left(\boldsymbol{P}^{zz}_{k|k-1}\right)^{-1}, \tag{7.27}$$

$$\widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\left(\boldsymbol{z}_k - \widehat{\boldsymbol{z}}_{k|k-1}\right), \tag{7.28}$$

$$\boldsymbol{P}^{xx}_{k|k} = \boldsymbol{P}^{xx}_{k|k-1} - \boldsymbol{K}_k\boldsymbol{P}^{zz}_{k|k-1}\boldsymbol{K}_k^T, \tag{7.29}$$

which is precisely the standard UKF.

Remark 2: by taking the expectation on both side of (7.25) and using some mathematical manipulations, we can get the predicted state vector $\widehat{\boldsymbol{x}}_{k|k-1}$ and its covariance matrix $\boldsymbol{P}^{xx}_{k|k-1}$

$$\widehat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{x}_k - \boldsymbol{\delta}_k, \tag{7.30}$$

where $\boldsymbol{\delta}_k$ is the prediction error and $\mathbb{E}\left[\boldsymbol{\delta}_k\boldsymbol{\delta}_k^T\right] = \boldsymbol{P}^{xx}_{k|k-1}$. By processing (7.30) and (7.26) simultaneously, we get the following batch-mode regression form:

$$\begin{bmatrix} \boldsymbol{z}_k + \boldsymbol{H}_k\widehat{\boldsymbol{x}}_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1} \\ \widehat{\boldsymbol{x}}_{k|k-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_k \\ \boldsymbol{I} \end{bmatrix}\boldsymbol{x}_k + \begin{bmatrix} \boldsymbol{v}_k + \boldsymbol{\varepsilon}_k \\ -\boldsymbol{\delta}_k \end{bmatrix} \tag{7.31}$$

which can be rewritten in a compact form as

$$\widetilde{\boldsymbol{z}}_k = \widetilde{\boldsymbol{H}}_k\boldsymbol{x}_k + \widetilde{\boldsymbol{e}}_k. \tag{7.32}$$

The error covariance matrix is given by

$$\boldsymbol{W}_k = \mathbb{E}\left[\widetilde{\boldsymbol{e}}_k\widetilde{\boldsymbol{e}}_k^T\right] = \begin{bmatrix} \boldsymbol{\Sigma}_{k|k-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}^{xx}_{k|k-1} \end{bmatrix}, \tag{7.33}$$

$\boldsymbol{\Sigma}_{k|k-1} = \mathbb{E}[(\boldsymbol{v}_k + \boldsymbol{\varepsilon}_k)(\boldsymbol{v}_k + \boldsymbol{\varepsilon}_k)^T] = \boldsymbol{R}_k + \boldsymbol{\Pi}_k$. By using the weighted least squares criterion to (7.32), that is, $min\ (\widetilde{\boldsymbol{z}}_k - \widetilde{\boldsymbol{H}}_k\boldsymbol{x}_k)^T{\boldsymbol{W}_k}^{-1}(\widetilde{\boldsymbol{z}}_k - \widetilde{\boldsymbol{H}}_k\boldsymbol{x}_k)$, we obtain

$$\widehat{\boldsymbol{x}}_{k|k} = \left(\widetilde{\boldsymbol{H}}_k^T{\boldsymbol{W}_k}^{-1}\widetilde{\boldsymbol{H}}_k\right)^{-1}\widetilde{\boldsymbol{H}}_k^T{\boldsymbol{W}_k}^{-1}\widetilde{\boldsymbol{z}}_k, \tag{7.34}$$

with estimation error covariance matrix

$$\boldsymbol{P}^{xx}_{k|k} = \left(\widetilde{\boldsymbol{H}}_k^T{\boldsymbol{W}_k}^{-1}\widetilde{\boldsymbol{H}}_k\right)^{-1}. \tag{7.35}$$

By using the matrix inversion Lemma and mathematical manipulations, it can be verified that the results shown in (7.34)-(7.35) are the same as (7.27)-(7.29).

Derivation of the Krein Space UKF

Before the derivation of the Krein space UKF, let us review the Krein space linear Kalman filter. Following the main results shown in [137, 138], we reorganize the Krein space linear Kalman filter into the following Lemma:

Lemma 5. *For a Krein space discrete-time system,*

$$\begin{cases} \boldsymbol{x_k} = \boldsymbol{A_k}\boldsymbol{x}_{k-1} + \boldsymbol{\eta}_k \\ \boldsymbol{y}_k = \boldsymbol{C}_k\boldsymbol{x}_k + \boldsymbol{\zeta}_k \end{cases} \tag{7.36}$$

with the Gramian matrix

$$\left\langle \begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{\eta}_k \\ \boldsymbol{\zeta}_k \end{bmatrix}, \begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{\eta}_k \\ \boldsymbol{\zeta}_k \end{bmatrix} \right\rangle = diag[\ \boldsymbol{P}_{0|0} \quad \boldsymbol{Q}_k \quad \boldsymbol{R}_k\ ] \tag{7.37}$$

both of which can be obtained from Krein space mapping corresponding to the indefinite quadratic function

$$J = \left\|\boldsymbol{x}_0 - \widehat{\boldsymbol{x}}_{0|0}\right\|^2_{\boldsymbol{P}^{-1}_{0|0}} + \sum_{k=0}^{N-1} \left\|\boldsymbol{\eta}_k\right\|^2_{\boldsymbol{Q}^{-1}_k} + \sum_{k=0}^{N} \left\|\boldsymbol{\zeta}_k\right\|^2_{\boldsymbol{R}^{-1}_k}. \tag{7.38}$$

If $\boldsymbol{P}_{0|0} \succ 0$, $\boldsymbol{Q}_k \succ 0$ and $\boldsymbol{R}_k$ is invertible and $[\boldsymbol{A}_k\ \boldsymbol{C}_k]$ has full rank for all k, the existence condition for the Krein space Kalman filter is provided by

$$(\boldsymbol{P}^{xx}_{k|k})^{-1} = (\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{C}^T_k\boldsymbol{R}^{-1}_k\boldsymbol{C}_k \succ 0, \tag{7.39}$$

where $\widehat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{A}_k\widehat{\boldsymbol{x}}_{k-1|k-1}$, $\boldsymbol{P}^{xx}_{k|k-1} = \boldsymbol{A}_k\boldsymbol{P}^{xx}_{k-1|k-1}\boldsymbol{A}^T_k + \boldsymbol{Q}_k$ and the state vector is updated by the following equations

$$\boldsymbol{P}^{xz}_{k|k-1} = \boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{C}^T_k, \quad \boldsymbol{P}^{zz}_{k|k-1} = \boldsymbol{C}^T_k\boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{C}_k + \boldsymbol{R}_k, \tag{7.40}$$

$$\boldsymbol{K}_k = \boldsymbol{P}^{xz}_{k|k-1}\left(\boldsymbol{P}^{zz}_{k|k-1}\right)^{-1}, \tag{7.41}$$

$$\widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\left(\boldsymbol{y}_k - \boldsymbol{C}_k\widehat{\boldsymbol{x}}_{k|k-1}\right), \tag{7.42}$$

$$\boldsymbol{P}^{xx}_{k|k} = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\boldsymbol{P}^{xx}_{k|k-1} = \boldsymbol{P}^{xx}_{k|k-1} - \boldsymbol{K}_k\boldsymbol{P}^{zz}_{k|k-1}\boldsymbol{K}^T_k \tag{7.43}$$

Remark 3: it can be observed from the above equations that there is little difference between the Hilbert space linear Kalman filter and the Krein space Kalman filter, except for the condition that $(\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{C}^T_k\boldsymbol{R}^{-1}_k\boldsymbol{C}_k \succ 0$. The latter is in fact required by the Hilbert space Kalman filter as well because an estimation error covariance matrix must be positive-definite. On the other hand, it is worth pointing out that $(\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{C}^T_k\boldsymbol{R}^{-1}_k\boldsymbol{C}_k$ can be

indefinite for the Krein space Kalman filter [137, 138]. However, under this condition, it is unable to achieve the minimum error covariance for the state estimates any more.

According to remark 2, if we define $\widehat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{A}_k\widehat{\boldsymbol{x}}_{k-1|k-1}$, $\boldsymbol{P}^{xx}_{k|k-1} = \boldsymbol{A}_k\boldsymbol{P}^{xx}_{k-1|k-1}\boldsymbol{A}_k^T + \boldsymbol{Q}_k$ and $\widehat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{x}_k - \boldsymbol{o}_k$, where $\mathbb{E}\left[\boldsymbol{o}_k\boldsymbol{o}_k^T\right] = \boldsymbol{P}^{xx}_{k|k-1}$, (7.36) can be organized into the following batch-mode regression form:

$$\begin{bmatrix} \boldsymbol{y}_k \\ \widehat{\boldsymbol{x}}_{k|k-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_k \\ \boldsymbol{I} \end{bmatrix}\boldsymbol{x}_k + \begin{bmatrix} \boldsymbol{\zeta}_k \\ -\boldsymbol{o}_k \end{bmatrix} \tag{7.44}$$

which can be further rewritten as

$$\widetilde{\boldsymbol{y}}_k = \widetilde{\boldsymbol{C}}_k\boldsymbol{x}_k + \widetilde{\boldsymbol{o}}_k. \tag{7.45}$$

where $\widehat{\boldsymbol{W}}_k = \mathbb{E}\left[\widetilde{\boldsymbol{o}}_k\widetilde{\boldsymbol{o}}_k^T\right] = diag[\boldsymbol{R}_k \;\; \boldsymbol{A}_k\boldsymbol{P}^{xx}_{k-1|k-1}\boldsymbol{A}_k^T + \boldsymbol{Q}_k]$.

Corollary 5.1. *If $\boldsymbol{P}_{0|0} \succ 0$, $\boldsymbol{Q}_k \succ 0$, $\boldsymbol{R}_k \succ 0$, $rank[\boldsymbol{A}_k\ \boldsymbol{C}_k] = n$ and*

$$(\boldsymbol{P}^{xx}_{k|k})^{-1} = (\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{C}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{C}_k \succ 0, \tag{7.46}$$

the Krein space linear KF can be derived by applying the weighted least square estimator to (7.45).

Proof. Following the procedures shown in remak 2, we can get

$$\begin{aligned} \boldsymbol{P}^{xx}_{k|k} &= \left(\widetilde{\boldsymbol{C}}_k^T\widehat{\boldsymbol{W}}_k^{-1}\widetilde{\boldsymbol{C}}_k\right)^{-1} \\ &= \left[\boldsymbol{C}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{C}_k + \left(\boldsymbol{P}^{xx}_{k|k-1}\right)^{-1}\right]^{-1} \\ &= \boldsymbol{P}^{xx}_{k|k-1} - \boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{C}_k^T\left(\boldsymbol{C}_k\boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{C}_k^T + \boldsymbol{R}_k\right)^{-1}\boldsymbol{C}_k\boldsymbol{P}^{xx}_{k|k-1} \\ &= (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\,\boldsymbol{P}^{xx}_{k|k-1} = \boldsymbol{P}^{xx}_{k|k-1} - \boldsymbol{K}_k\boldsymbol{P}^{zz}_{k|k-1}\boldsymbol{K}_k^T \end{aligned} \tag{7.47}$$

where the gain matrix is expressed as

$$\boldsymbol{K}_k = \boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{C}_k^T(\boldsymbol{C}_k\boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{C}_k^T + \boldsymbol{R}_k)^{-1} = \boldsymbol{P}^{xz}_{k|k-1}(\boldsymbol{P}^{zz}_{k|k-1})^{-1}. \tag{7.48}$$

By following the similar rules and using the matrix inversion Lemma, we can derive the same state vector updating form as (7.42). □

Corollary 5.2. *If $\boldsymbol{P}_{0|0} \succ 0$, $\boldsymbol{Q}_k \succ 0$, $\boldsymbol{R}_k \succ 0$, $rank[\boldsymbol{F}_k\ \boldsymbol{H}_k] = n$ and*

$$(\boldsymbol{P}^{xx}_{k|k})^{-1} = (\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k \succ 0, \tag{7.49}$$

the Krein space UKF can be derived as follows:

$$\boldsymbol{K}_k = \boldsymbol{P}^{xz}_{k|k-1}\left(\boldsymbol{P}^{zz}_{k|k-1}\right)^{-1}, \tag{7.50}$$

$$\widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\left(\boldsymbol{z}_k - \widehat{\boldsymbol{z}}_{k|k-1}\right), \tag{7.51}$$

$$\boldsymbol{P}^{xx}_{k|k} = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\,\boldsymbol{P}^{xx}_{k|k-1}, \tag{7.52}$$

where $\boldsymbol{P}^{zz}_{k|k-1}$ and $\boldsymbol{P}^{xz}_{k|k-1}$ are expressed in (7.20) and (7.24), respectively.

Proof. The linear-like regression model of the UKF shown in (7.25)-(7.26) can be easily organized into the similar one as (7.36). Thus, following Lemma 5 and Corollary 5.1, we can get

$$\boldsymbol{P}^{xz}_{k|k-1} = \boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k^T, \;\; \boldsymbol{P}^{zz}_{k|k-1} = \boldsymbol{H}_k^T\boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k + \boldsymbol{R}_k. \tag{7.53}$$

Then, the gain $\boldsymbol{K}_k$ and state estimates $\widehat{\boldsymbol{x}}_{k|k}$ can be derived as (7.50) and (7.51). □

Remark 4: In fact, the positive-definiteness of $(\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k$ must be satisfied by the traditional Hilbert space UKF. This is because in the state prediction stage, the generation of new sigma points requires the square-root of $\boldsymbol{P}^{xx}_{k|k}$. The positive-definiteness of $\boldsymbol{P}^{xx}_{k|k}$ immediately implies $(\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k \succ 0$ according to the relationship between them shown in (7.49). On the other hand, motivated by the Krein space UKF, the Krein space H-infinity UKF can be derived as well. This is shown in the next section.

7.3.2 Derivation of the Krein Space H-infinity UKF

The H-infinity criterion aims to design a filter that achieves the smallest estimation error for all possible disturbances with bounded energy. Specifically, the filter is designed such that the following criterion is satisfied:

$$\frac{\sum_{k=0}^{N}\|\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k}\|^2_{\boldsymbol{P}^{-1}_{k|k}}}{\|\boldsymbol{x}_0 - \widehat{\boldsymbol{x}}_{0|0}\|^2_{\boldsymbol{P}^{-1}_{0|0}} + \sum_{k=0}^{N-1}\|\boldsymbol{w}_k\|^2_{\boldsymbol{Q}_k^{-1}} + \sum_{k=0}^{N}\|\boldsymbol{v}_k\|^2_{\boldsymbol{R}_k^{-1}}} \leq \gamma^2, \tag{7.54}$$

where $\boldsymbol{x}_0$ and $\boldsymbol{P}_{0|0}$ are the initial state vector and its covariance matrix, respectively; γ is a positive scalar parameter that bounds the uncertainties.

Theorem 6. *If rank*$[\boldsymbol{F}_k \; \boldsymbol{H}_k] = n$*, there exists a filter that achieves the H-infinity criterion shown in (7.54) if and only if the estimation error covariance matrix* $\boldsymbol{P}^{xx}_{k|k}$ *for all* k *satisfies*

$$(\boldsymbol{P}^{xx}_{k|k})^{-1} = (\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k - \gamma^{-2}\boldsymbol{I} \succ 0, \tag{7.55}$$

and the recursive H-infinity UKF can be expressed as

$$\widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\left(\boldsymbol{z}_k - \widehat{\boldsymbol{z}}_{k|k-1}\right), \tag{7.56}$$

$$\boldsymbol{K}_k = \boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k^T(\boldsymbol{H}_k\boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k^T + \boldsymbol{R}_k)^{-1}, \tag{7.57}$$

$$= \boldsymbol{P}^{xz}_{k|k-1}(\boldsymbol{P}^{zz}_{k|k-1})^{-1} \tag{7.58}$$

$$\boldsymbol{P}^{xx}_{k|k} = (\boldsymbol{I} - \boldsymbol{P}^{xx}_{k|k-1}[\boldsymbol{H}_k^T \; \boldsymbol{I}]\boldsymbol{R}^{-1}_{e,k}[\boldsymbol{H}_k^T \; \boldsymbol{I}]^T)\boldsymbol{P}^{xx}_{k|k-1}, \tag{7.59}$$

$$= \boldsymbol{P}^{xx}_{k|k-1} - [\boldsymbol{P}^{xz}_{k|k-1} \; \boldsymbol{P}^{xx}_{k|k-1}]\boldsymbol{R}^{-1}_{e,k}[\boldsymbol{P}^{xz}_{k|k-1} \; \boldsymbol{P}^{xx}_{k|k-1}]^T, \tag{7.60}$$

$$\boldsymbol{R}_{e,k} = \begin{bmatrix} \boldsymbol{R}_k + \boldsymbol{H}_k \boldsymbol{P}^{xx}_{k|k-1} \boldsymbol{H}_k^T & \left(\boldsymbol{P}^{xx}_{k|k-1} \boldsymbol{H}_k^T\right)^T \\ \boldsymbol{P}^{xx}_{k|k-1} \boldsymbol{H}_k^T & -\gamma^2 \boldsymbol{I} + \boldsymbol{P}^{xx}_{k|k-1} \end{bmatrix}; \tag{7.61}$$

$$= \begin{bmatrix} \boldsymbol{R}_k + \boldsymbol{P}^{zz}_{k|k-1} & [\boldsymbol{P}^{xz}_{k|k-1}]^T \\ \boldsymbol{P}^{xz}_{k|k-1} & -\gamma^2 \boldsymbol{I} + \boldsymbol{P}^{xx}_{k|k-1} \end{bmatrix}, \tag{7.62}$$

and where $\boldsymbol{I}$ is an identity matrix.

Proof. The first step is to convert the suboptimal H-infinity filtering problem (7.54) to an indefinite form so that the Krein space Kalman filter can be used. Formally, we get

$$\begin{aligned} J_\infty &= \left\| \boldsymbol{x}_0 - \widehat{\boldsymbol{x}}_{0|0} \right\|^2_{\boldsymbol{P}^{-1}_{0|0}} + \sum_{k=0}^{N-1} \|\boldsymbol{w}_k\|^2_{\boldsymbol{Q}_k^{-1}} + \sum_{k=0}^{N} \|\boldsymbol{v}_k\|^2_{\boldsymbol{R}_k^{-1}} \\ &\quad - \gamma^{-2} \sum_{k=0}^{N} \|\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k}\|^2_{\boldsymbol{P}^{-1}_{k|k}} \\ &= \|\boldsymbol{x}_0 - \widehat{\boldsymbol{x}}_{0|0}\|^2_{\boldsymbol{P}^{-1}_{0|0}} + \sum_{k=0}^{N-1} \|\boldsymbol{w}_k\|^2_{\boldsymbol{Q}_k^{-1}} \\ &\quad + \sum_{k=0}^{N} \begin{bmatrix} \boldsymbol{z}_k - \boldsymbol{h}(\boldsymbol{x}_k) \\ \boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k} \end{bmatrix}^T \begin{bmatrix} \boldsymbol{R}_k & \boldsymbol{0} \\ \boldsymbol{0} & -\gamma^2 \boldsymbol{I} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{z}_k - \boldsymbol{h}(\boldsymbol{x}_k) \\ \boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k} \end{bmatrix} \end{aligned} \tag{7.63}$$

Define $\boldsymbol{c}_k = \widehat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{F}_k \widehat{\boldsymbol{x}}_{k-1|k-1} + \boldsymbol{e}_k$, $\boldsymbol{m}_k = \widehat{\boldsymbol{z}}_{k|k-1} - \boldsymbol{H}_k \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{\varepsilon}_k$, $\boldsymbol{y}_k = \boldsymbol{z}_k - \boldsymbol{m}_k$, $\widetilde{\boldsymbol{C}}_k^T = [\boldsymbol{H}_k^T \ \boldsymbol{I}]$, $\widetilde{\boldsymbol{m}}_k = [\boldsymbol{y}_k^T \ \widehat{\boldsymbol{x}}_{k|k}^T]^T$ and use the statistical linerization results shown in (7.25)-(7.26), (7.63) can be rewritten as:

$$\begin{aligned} J_\infty &= \|\boldsymbol{x}_0 - \widehat{\boldsymbol{x}}_{0|0}\|^2_{\boldsymbol{P}^{-1}_{0|0}} + \sum_{k=0}^{N-1} \|\boldsymbol{w}_k\|^2_{\boldsymbol{Q}_k^{-1}} \\ &\quad + \sum_{k=0}^{N} \begin{bmatrix} \boldsymbol{y}_k - \boldsymbol{H}_k \boldsymbol{x}_k \\ \boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k} \end{bmatrix}^T \begin{bmatrix} \boldsymbol{R}_k & \boldsymbol{0} \\ \boldsymbol{0} & -\gamma^2 \boldsymbol{I} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{y}_k - \boldsymbol{H}_k \boldsymbol{x}_k \\ \boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k} \end{bmatrix} \end{aligned} \tag{7.64}$$

that resembles the Krein space Kalman filter form (7.36). Therefore, according to Lemma 5 and Corollary 5.1, the estimation error covariance matrix is derived as

$$\begin{aligned} (\boldsymbol{P}^{xx}_{k|k})^{-1} &= (\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \widetilde{\boldsymbol{C}}_k \boldsymbol{P}^{xx}_{k|k-1} \widetilde{\boldsymbol{C}}_k^T \\ &= (\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{H}_k^T \boldsymbol{R}_k^{-1} \boldsymbol{H}_k - \gamma^{-2} \boldsymbol{I} \end{aligned} \tag{7.65}$$

It is clear that $(\boldsymbol{P}_{k|k}^{xx})^{-1} \succ 0$ must hold for true for the existence of H-infinity UKF. This is because the positive-definiteness of $\boldsymbol{P}_{k|k}^{xx}$ is needed for the generation of sigma points at the next time instant. By applying the matrix inversion Lemma to (7.65) and using tedious algebraic manipulations, we are able to derive

$$\begin{aligned}\boldsymbol{P}_{k|k}^{xx} &= (\boldsymbol{I} - \boldsymbol{P}_{k|k-1}^{xx}[\boldsymbol{H}_k^T \ \boldsymbol{I}]\boldsymbol{R}_{e,k}^{-1}[\boldsymbol{H}_k^T \ \boldsymbol{I}]^T)\boldsymbol{P}_{k|k-1}^{xx} \\ &= \boldsymbol{P}_{k|k-1}^{xx} - [\boldsymbol{P}_{k|k-1}^{xz} \ \boldsymbol{P}_{k|k-1}^{xx}]\boldsymbol{R}_{e,k}^{-1}[\boldsymbol{P}_{k|k-1}^{xz} \ \boldsymbol{P}_{k|k-1}^{xx}]^T \end{aligned} \tag{7.66}$$

where the expression of $\boldsymbol{R}_{e,k}$ is the same as (7.62).

In the meantime, according to the formula of the Krein space Kalman filter, the filtered state vector is derived as

$$\begin{aligned}\widehat{\boldsymbol{x}}_{k|k} &= \widehat{\boldsymbol{x}}_{k|k-1} \\ &\quad + \boldsymbol{P}_{k|k-1}^{xx}\widetilde{\boldsymbol{C}}_k^T(\widetilde{\boldsymbol{C}}_k\boldsymbol{P}_{k|k-1}^{xx}\widetilde{\boldsymbol{C}}_k^T + \boldsymbol{R}_k)^{-1}(\widetilde{\boldsymbol{m}}_k - \widetilde{\boldsymbol{C}}_k\widehat{\boldsymbol{x}}_{k|k-1}) \\ &= \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{P}_{k|k-1}^{xx}[\boldsymbol{H}_k^T \ \boldsymbol{I}]\begin{bmatrix} \boldsymbol{I} & -\widehat{\boldsymbol{R}}_k\boldsymbol{H}_k\boldsymbol{P}_{k|k-1}^{xx} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \\ &\quad \times \begin{bmatrix} \widehat{\boldsymbol{R}}_k & \boldsymbol{0} \\ \boldsymbol{0} & -\gamma^2\boldsymbol{I} + ((\boldsymbol{P}_{k|k-1}^{xx})^{-1} + \boldsymbol{H}_k^T\boldsymbol{H}_k)^{-1} \end{bmatrix}^{-1} \\ &\quad \times \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ -\boldsymbol{P}_{k|k-1}^{xx}\boldsymbol{H}_k^T\widehat{\boldsymbol{R}}_k^{-1} & \boldsymbol{I} \end{bmatrix}\begin{bmatrix} \boldsymbol{z}_k - \widehat{\boldsymbol{z}}_{k|k-1} \\ \widehat{\boldsymbol{x}}_{k|k} - \widehat{\boldsymbol{x}}_{k|k-1} \end{bmatrix}^T \end{aligned} \tag{7.67}$$

where $\widehat{\boldsymbol{R}}_k = \boldsymbol{R}_k + \boldsymbol{H}_k\boldsymbol{P}_{k|k-1}^{xx}\boldsymbol{H}_k^T = \boldsymbol{P}_{k|k-1}^{zz}$. Using the matrix inversion Lemma and mathematical manipulations, we can finally get

$$\widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\left(\boldsymbol{z}_k - \widehat{\boldsymbol{z}}_{k|k-1}\right), \tag{7.68}$$

$$\boldsymbol{K}_k = \boldsymbol{P}_{k|k-1}^{xz}(\boldsymbol{P}_{k|k-1}^{zz})^{-1} \tag{7.69}$$

□

It is interesting to notice that the H-infinity UKF shares the similar structure to that of the the original UKF. One apparent difference is the updating of estimation error covariance matrix. The latter drives the H-infinity UKF to achieve the smallest estimation error for all possible disturbances. The bounded error performance of the H-infinity filter subject to uncertainties can be proved following the procedures shown in [121]. On the other hand, γ can be seen as a tuning parameter to balance the tradeoff between the H-infinity and the minimum mean-square error performance. Indeed, when γ tends to infinity, the H-infinity UKF will reduce to the traditional UKF. This indicates that the H-infinity norm of the traditional UKF may be quite large, leading to poor robustness against uncertainties.

Furthermore, indefinite covariance matrix $\begin{bmatrix} \boldsymbol{R}_k & \boldsymbol{0} \\ \boldsymbol{0} & -\gamma^2\boldsymbol{I} \end{bmatrix}^{-1}$ is involved during the proof when deriving the Krein space H-infinity UKF, which is very different from that of the Hilbert space H-infinity Kalman filter. The latter requires the positive-definiteness of a covariance matrix.

Theorem 7. *The H-infinity UKF derived in the Krein space is based on the weighted least square estimator and thus lacks of robustness to non-Gaussian noise and any types of outliers.*

Proof. Define $\widehat{\boldsymbol{x}}_{k|k} = \boldsymbol{x}_k - \boldsymbol{n}_k$, where $\mathbb{E}[\boldsymbol{n}_k] = \boldsymbol{0}$ and $\mathbb{E}[\boldsymbol{n}_k\boldsymbol{n}_k^T] = -\gamma^2\boldsymbol{I}$. According to corollaries 5.1 and 5.2, the predicted state information can be used to construct the batch-mode Krein space UKF instead of the form shown in (7.25). Thus, the state prediction error form $\widehat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{x}_k - \boldsymbol{\delta}_k$ is used, where $\boldsymbol{\delta}_k$ is the prediction error and $\mathbb{E}\left[\boldsymbol{\delta}_k\boldsymbol{\delta}_k^T\right] = \boldsymbol{P}^{xx}_{k|k-1}$. On the other hand, during the proof of Theorem 6, the following Krein space regression form is actually used,

$$\begin{bmatrix} \boldsymbol{z}_k + \boldsymbol{H}_k\widehat{\boldsymbol{x}}_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1} - \boldsymbol{\varepsilon}_k \\ \widehat{\boldsymbol{x}}_{k|k-1} \\ \widehat{\boldsymbol{x}}_{k|k} \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_k \\ \boldsymbol{I} \\ \boldsymbol{I} \end{bmatrix}\boldsymbol{x}_k + \begin{bmatrix} \boldsymbol{v}_k \\ -\boldsymbol{\delta}_k \\ -\boldsymbol{n}_k \end{bmatrix} \tag{7.70}$$

that can be rewritten as

$$\widetilde{\boldsymbol{z}}_k = \widetilde{\boldsymbol{H}}_k\boldsymbol{x}_k + \boldsymbol{\varsigma}_k. \tag{7.71}$$

where $\widetilde{\boldsymbol{W}}_k = \mathbb{E}[\boldsymbol{\varsigma}_k\boldsymbol{\varsigma}_k^T] = diag[\boldsymbol{R}_k\ \boldsymbol{P}^{xx}_{k|k-1}\ -\gamma^2\boldsymbol{I}]$. Applying the weighted least square estimator to (7.71) yields

$$\begin{aligned} \widehat{\boldsymbol{x}}_{k|k} &= \mathbb{E}[\left(\widetilde{\boldsymbol{H}}_k^T\widetilde{\boldsymbol{W}}_k^{-1}\widetilde{\boldsymbol{H}}_k\right)^{-1}\widetilde{\boldsymbol{H}}_k^T\widetilde{\boldsymbol{W}}_k^{-1}\widetilde{\boldsymbol{z}}_k] \\ &= ([\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\ (\boldsymbol{P}^{xx}_{k|k-1})^{-1}\ -\gamma^2\boldsymbol{I}]\begin{bmatrix} \boldsymbol{H}_k \\ \boldsymbol{I} \\ \boldsymbol{I} \end{bmatrix})^{-1}\widetilde{\boldsymbol{H}}_k^T\boldsymbol{W}_k^{-1}\mathbb{E}[\widetilde{\boldsymbol{z}}_k] \\ &= [\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k + (\boldsymbol{P}^{xx}_{k|k-1})^{-1} - \gamma^2\boldsymbol{I}]^{-1} \\ &\times \begin{bmatrix} \boldsymbol{R}_k^{-T}\boldsymbol{H}_k \\ (\boldsymbol{P}^{xx}_{k|k-1})^{-T} \\ -\gamma^2\boldsymbol{I} \end{bmatrix}^T \begin{bmatrix} \boldsymbol{z}_k + \boldsymbol{H}_k\widehat{\boldsymbol{x}}_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1} \\ \widehat{\boldsymbol{x}}_{k|k-1} \\ \widehat{\boldsymbol{x}}_{k|k} \end{bmatrix} \end{aligned} \tag{7.72}$$

Moving the term $\widehat{\boldsymbol{x}}_{k|k}$ from the right-hand side to the left-hand side and applying matrix manipulations, we can finally arrive

$$\widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\left(\boldsymbol{z}_k - \widehat{\boldsymbol{z}}_{k|k-1}\right), \tag{7.73}$$

$$\boldsymbol{K}_k = \boldsymbol{P}^{xz}_{k|k-1}(\boldsymbol{P}^{zz}_{k|k-1})^{-1} \tag{7.74}$$

where $\boldsymbol{P}^{xz}_{k|k-1} = \boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k^T$ and $\boldsymbol{P}^{zz}_{k|k-1} = \boldsymbol{H}_k^T\boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k + \boldsymbol{R}_k$. In terms of the estimation error covariance matrix, it is expressed as

$$\begin{aligned}\boldsymbol{P}^{xx}_{k|k} &= (\widetilde{\boldsymbol{H}}_k^T\widetilde{\boldsymbol{W}}_k^{-1}\widetilde{\boldsymbol{H}}_k)^{-1} \\ &= [\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k + (\boldsymbol{P}^{xx}_{k|k-1})^{-1} - \gamma^2\boldsymbol{I}]^{-1} \\ &= \boldsymbol{P}^{xx}_{k|k-1} - [\boldsymbol{P}^{xz}_{k|k-1}\ \boldsymbol{P}^{xx}_{k|k-1}]\boldsymbol{R}^{-1}_{e,k}[\boldsymbol{P}^{xz}_{k|k-1}\ \boldsymbol{P}^{xx}_{k|k-1}]^T \end{aligned} \tag{7.75}$$

where $\boldsymbol{R}_{e,k}$ is the same as (7.60). Therefore, it is clear that H-infinity UKF is based on the least square estimator in the Krein space. According to the robust statistics [67, 82], it is well-known that its influence function is unbounded and therefore lacks of robustness to non-Gaussian noise and outliers. □

Remark 5: It is worth emphasizing that the proposed nonlinear H-infinity filter is very general. The UKF is just used as an representative example. Other filters that share the similar structure of UKF can be leveraged to derive their corresponding H-infinity filters in the Krein space, such as divided difference filter, quadrature Kalman filter, cubature Kalman filter, EnKF and so on [139–141]. On the other hand, instead of propagating statistics of the state vector using sigma points, the first-order Taylor series expansion-based methods can be used, yielding H-infinity EKF.

7.3.3 Derivation of the Robust H-infinity UKF

It is shown in Theorem 2 that the H-infinity filter is actually based on the Krein space weighted least squares estimator. Despite of its bounded performance against system uncertainties, it lacks of robustness to non-Gaussian noise and outliers. By contrast, the GM-estimator derived from robust statistics is able to handle them but it may yield large estimation errors in presence of system uncertainties. Therefore, to suppress the outliers and filter out thick-tailed non-Gaussian measurement noise while bounding the system uncertainties, we propose to apply GM-estimator to (7.71) instead of the weighted least squares, yielding the following objective function:

$$J(\boldsymbol{x}_k) = \sum_{i=1}^{m+2n} \varpi_i^2 \rho(r_{S_i}), \tag{7.76}$$

where ϖ_i are the weights to downweight outliers calculated by applying projection statistics [38, 88] to innovation matrix $\boldsymbol{Z}_k$. The latter is defined as follows:

$$\boldsymbol{Z}_k = \begin{bmatrix} \boldsymbol{z}_{k-1} - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k-1|k-2}) & \boldsymbol{z}_k - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k|k-1}) \\ \widehat{\boldsymbol{x}}_{k-1|k-2} & \widehat{\boldsymbol{x}}_{k|k-1} \\ \widehat{\boldsymbol{x}}_{k-2|k-2} & \widehat{\boldsymbol{x}}_{k-1|k-1} \end{bmatrix}, \tag{7.77}$$

where $\boldsymbol{z}_{k-1} - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k-1|k-2})$ and $\boldsymbol{z}_k - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k|k-1})$ are the innovation vectors while $\widehat{\boldsymbol{x}}_{k-1|k-2}$ and $\widehat{\boldsymbol{x}}_{k|k-1}$ are the predicted state vectors at time instants k-1 and k, respectively; $\widehat{\boldsymbol{x}}_{k-2|k-2}$

and $\widehat{\boldsymbol{x}}_{k-1|k-1}$ are the filtered state vectors at time instants k-2 and k-1, respectively; the mathematical expression of projection statistics is defined as [38, 88]:

$$PS_j = \max_{\|\boldsymbol{\ell}\|=1} \frac{\left|\boldsymbol{l}_j^T\boldsymbol{\ell} - med_i\left(\boldsymbol{l}_i^T\boldsymbol{\ell}\right)\right|}{1.4826\, med_\kappa \left|\boldsymbol{l}_\kappa^T\boldsymbol{\ell} - med_i\left(\boldsymbol{l}_i^T\boldsymbol{\ell}\right)\right|}, \tag{7.78}$$

for $i, j, \kappa = 1, 2, ..., m{+}2n$. where $\boldsymbol{l}_j^T$ is the jth row vector of $\boldsymbol{Z}_k$; $\boldsymbol{\ell}$ represents a set of directions that originate from the coordinatewise medians of the $\boldsymbol{Z}_k$ and pass through every data point. The calculated PS values are compared to a statistical threshold to identify outliers. Extensive Monte Carlo simulations and Q-Q plots reveal that the probability distributions of the PS applied to $\boldsymbol{Z}_k$ follow chi-square distributions with degree of freedom 2. As a result, the points that satisfy $\mathrm{PS}_i > \chi^2_{2,0.975}$ are identified as outliers and downweighted through

$$\varpi_i = \min\left(1, d^2/PS_i^2\right), \tag{7.79}$$

where the parameter d is set equal to 1.5 to yield good statistical efficiency at different distributions.

In (7.76), $r_{S_i} = r_i/s\varpi_i$ is the standardized residual; $r_i = \widetilde{z}_i - \widetilde{\boldsymbol{h}}_i^T\widehat{\boldsymbol{x}}$ is the residual, where $\widetilde{\boldsymbol{h}}_i^T$ is the ith row vector of the matrix $\widetilde{\boldsymbol{H}}_k$; $s = 1.4826 \cdot b_m \cdot \mathrm{median}_i\,|r_i|$ is the robust scale estimate; b_m is a correction factor to achieve unbiasedness for a finite sample of size $m + 2n$ at a given probability distribution; $\rho(\cdot)$ is the nonlinear function of r_{S_i} expressed as

$$\rho\left(r_{S_i}\right) = \begin{cases} \frac{1}{2}r_{S_i}^2, & \text{for } |r_{S_i}| < \lambda \\ \lambda\left|r_{S_i}\right| - \lambda^2/2, & elsewhere \end{cases}, \tag{7.80}$$

where the parameter λ is typically chosen between 1.5 to 3 in the literature.

To minimize (7.76), we take its partial derivative with respect to $\boldsymbol{x}_k$ and set it equal to zero, yielding

$$\frac{\partial J\left(\boldsymbol{x}_k\right)}{\partial \boldsymbol{x}_k} = \sum_{i=1}^{m+2n} -\frac{\varpi_i\widetilde{\boldsymbol{h}}_i}{s\sigma_i}\psi\left(r_{S_i}\right) = \boldsymbol{0}, \tag{7.81}$$

where $\psi\left(r_{S_i}\right) = \partial\rho\left(r_{S_i}\right)/\partial r_{S_i}$; σ_i is the square-root of the ith diagonal element of matrix $\widetilde{\boldsymbol{W}}_k$. By dividing and multiplying the standardized residual r_{S_i} to both sides of (7.81) and putting it in a matrix form, we get

$$\widetilde{\boldsymbol{H}}_k^T\widetilde{\boldsymbol{W}}_k^{-1}\widehat{\boldsymbol{Q}}\left(\widetilde{\boldsymbol{z}}_k - \widetilde{\boldsymbol{H}}_k\boldsymbol{x}_k\right) = \boldsymbol{0}, \tag{7.82}$$

where $\widehat{\boldsymbol{Q}}$ =diag$(q\left(r_{S_i}\right))$ and $q\left(r_{S_i}\right) = \psi\left(r_{S_i}\right)/r_{S_i}$.

By using the iteratively reweighted least squares (IRLS) algorithm [38, 82], the state vector correction at the j iteration is calculated through

$$\Delta\widehat{\boldsymbol{x}}_{k|k}^{(j+1)} = \left(\widetilde{\boldsymbol{H}}_k^T\widetilde{\boldsymbol{W}}_k^{-1}\widehat{\boldsymbol{Q}}^{(j)}\widetilde{\boldsymbol{H}}_k\right)^{-1}\widetilde{\boldsymbol{H}}_k^T\widetilde{\boldsymbol{W}}_k^{-1}\widehat{\boldsymbol{Q}}^{(j)}\widetilde{\boldsymbol{z}}_k^j, \tag{7.83}$$

where $\Delta\widehat{\boldsymbol{x}}_{k|k}^{(j+1)} = \widehat{\boldsymbol{x}}_{k|k}^{(j+1)} - \widehat{\boldsymbol{x}}_{k|k}^{(j)}$. It should be noted that $\widetilde{\boldsymbol{z}}_k$ needs to be updated at each jth iteration as it contains $\widehat{\boldsymbol{x}}_{k|k}$. The latter changes at each iteration. Following similar steps of Theorem 7, we move $\widehat{\boldsymbol{x}}_{k|k}$ from the right-hand side of (7.83) to its left-hand side, yielding

$$\Delta\widehat{\boldsymbol{x}}_{k|k}^{(j+1)} = \left(\boldsymbol{\Gamma}_k^T\boldsymbol{\Lambda}_k^{-1}\boldsymbol{\Xi}^{(j)}\boldsymbol{\Gamma}_k\right)^{-1}\boldsymbol{\Gamma}_k^T\boldsymbol{\Lambda}_k^{-1}\boldsymbol{\Xi}^{(j)}\boldsymbol{\xi}_k, \tag{7.84}$$

where $\boldsymbol{\Gamma} = [\boldsymbol{H}_k^T \ \boldsymbol{I}]^T \in \mathbb{R}^{(m+n)\times n}$; $\boldsymbol{\Xi} \in \mathbb{R}^{(m+n)\times(m+n)}$ and $\boldsymbol{\Lambda}_k \in \mathbb{R}^{(m+n)\times(m+n)}$ are diagonal matrices whose diagonal elements are the previous $m+n$ diagonal ones of $\widehat{\boldsymbol{Q}}$ and $\widetilde{\boldsymbol{W}}_k$, respectively; $\boldsymbol{\xi}_k = [\widetilde{\boldsymbol{z}}_k^T \ \widehat{\boldsymbol{x}}_{k|k-1}^T]^T \in \mathbb{R}^{(m+n)\times 1}$. As a result, it resembles the GM-UKF. In fact, (7.84) is recommended to obtain the state estimates. This is because matrices $\boldsymbol{\Gamma}$ and $\boldsymbol{\Xi}$ have much lower dimensions compared with $\widetilde{\boldsymbol{H}}_k$ and $\widehat{\boldsymbol{Q}}$, leading to improved computational efficiency. Note that the algorithm converges when $\left\|\Delta\widehat{\boldsymbol{x}}_{k|k}^{(j+1)}\right\|_\infty \leq 10^{-2}$.

For the asymptotic error covariance matrix of the robust H-infinity UKF at time sample k, we can derive it using the influence function of the GM-estimator. According to our previous work [38], the influence function of the robust H-infinity UKF can be derived as:

$$\boldsymbol{IF}(\boldsymbol{x};\Phi,\boldsymbol{T}) = [\int \frac{1}{s}\psi'(r_{S_i})\widetilde{\boldsymbol{H}}_k\widetilde{\boldsymbol{H}}_k^T|_{T(\Phi)}\, d\Phi]^{-1}\varpi\widetilde{\boldsymbol{H}}_k\psi(r_{S_i}) \tag{7.85}$$

where Φ is the cpdf of the standardized residual r_{S_i}; $T(\cdot)$ is the functional form of the robust H-infinity UKF. Finally, the asymptotic error covariance matrix can be updated through

$$\begin{aligned}\boldsymbol{P}_{k|k}^{xx} &= \mathbb{E}[\boldsymbol{IF}(\boldsymbol{x};\Phi,\boldsymbol{T})\cdot\boldsymbol{IF}(\boldsymbol{x};\Phi,\boldsymbol{T})^T] \\ &= \alpha(\widetilde{\boldsymbol{H}}_k^T\widetilde{\boldsymbol{W}}_k^{-1}\widetilde{\boldsymbol{H}}_k)^{-1}(\widetilde{\boldsymbol{H}}_k^T\widehat{\boldsymbol{Q}}_\varpi\widetilde{\boldsymbol{W}}_k^{-1}\widetilde{\boldsymbol{H}}_k)(\widetilde{\boldsymbol{H}}_k^T\widetilde{\boldsymbol{W}}_k^{-1}\widetilde{\boldsymbol{H}}_k)^{-1}\end{aligned} \tag{7.86}$$

where $\alpha = \frac{\mathbb{E}_\Phi[\psi^2(r_{S_i})]}{\{\mathbb{E}_\Phi[\psi'(r_{S_i})]\}^2}$ and $\boldsymbol{Q}_\varpi = diag(\varpi_i^2)$.

Remark 6: In the absence of non-Gaussian noise and outliers, no predicted state and measurement are downweighted. Therefore, $\widehat{\boldsymbol{Q}} = \boldsymbol{I}$ and

$$\widehat{\boldsymbol{x}}_{k|k} = \left(\widetilde{\boldsymbol{H}}_k^T\widetilde{\boldsymbol{W}}_k^{-1}\widetilde{\boldsymbol{H}}_k\right)^{-1}\widetilde{\boldsymbol{H}}_k^T\widetilde{\boldsymbol{W}}_k^{-1}\widetilde{\boldsymbol{z}}_k, \tag{7.87}$$

$$\boldsymbol{P}_{k|k}^{xx} = \frac{\mathbb{E}_\Phi[\psi^2(r_{S_i})]}{\{\mathbb{E}_\Phi[\psi'(r_{S_i})]\}^2}(\widetilde{\boldsymbol{H}}_k^T\widetilde{\boldsymbol{W}}_k^{-1}\widetilde{\boldsymbol{H}}_k)^{-1}, \tag{7.88}$$

where α is very close to 1. For example, under Gaussian noise, this value can be calculated as 1.0369. As a result, (7.87) and (7.88) are the same of (7.72) and (7.75), respectively. Thus, the robust H-infinity UKF reduces to the H-infinity UKF. On the other hand, when γ tends to infinity, the H-infinity UKF reduces to the UKF. It is thus easy to conclude that the robust H-infinity UKF reduces to the GM-UKF. Now, it becomes clear that the proposed robust H-infinity UKF leverages the robustness of the GM-estimator to filter out non-Gaussian noise and suppress outliers while relying on the H-infinity criterion to bound system uncertainties.

7.3.4 Application to Power System Dynamic State Estimation

With regard to power system dynamic state estimation using PMU measurements, both centralized and decentralized versions have been proposed [38, 88]. The former one requires accurate dynamical system model of each components, including synchronous generators, dynamic loads, etc, and wide-area PMU measurements, which may be hard to achieve for practical power system. This motivates the development of decentralized DSE that is implemented at each local synchronous generator. The proposed robust H-infinity UKF can be implemented in both ways. However, only the decentralized one is used for demonstration in this chapter. For dynamic state estimation, the state vector is $\boldsymbol{x}_k = [\delta\ \omega\ E'_d\ E'_q\ E_{fd}\ V_F\ V_R\ T_M\ P_{SV}]$. The system input vector is denoted by $\boldsymbol{u}_k = [V_{ref}\ P_C\ \boldsymbol{V}\ \boldsymbol{\theta}]^T$, where $\boldsymbol{V}$ and $\boldsymbol{\theta}$ are the generator terminal voltage magnitudes and angles obtained by PMUs. The measurement vector $\boldsymbol{z}_k$ contains a collection of real and reactive power injections P_e and Q_e obtained by the PMUs.

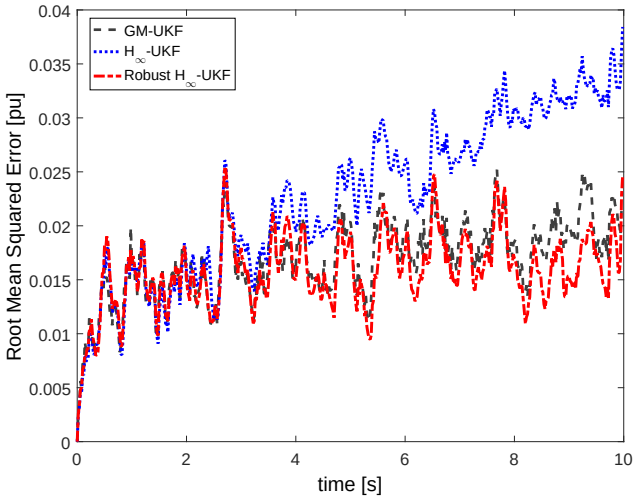


Figure 7.1: Root-mean-squared errors of the GM-UKF, the H-infinity UKF and the robust H-infinity UKF in presence of unknown non-Gaussian system process and measurement noise.

7.4 Numerical Results

Extensive simulations are carried out on the IEEE 39-bus system to assess the performance of the proposed robust H-infinity UKF under various scenarios. Each synchronous generator is assumed to be the two-axis model equipped with the IEEE-DC1A exciter and the TGOV1 turbine-governor. The parameters of the generator model can be found in [70]. The transient stability time domain simulations are performed to generate measurements and true state variables using the Matlab-based software PST [142] with some modifications. The 4th order

Ruger-Kutta approach is adopted with an integration step of t=1/120 s to solve differential and algebraic equations. The simulations consist of the following steps: Line 15-16 is tripped at t=0.5s to simulate a system disturbance; the voltage phasor, current phasor and frequency at each generator's terminal bus are corrupted by additive noise to simulate realistic PMU measurements; the sampling rate of the PMU measurements is assumed to be 60 samples/s. The maximal number of iterations allowed for the IRLS algorithm is 20. The parameters λ and d are set to 1.5. The convergence tolerance threshold of the IRLS algorithm is 0.01. The tuning parameter of the H-infinity filter is 10; the root-mean-squared error (RMSE) of all estimated generator state variables is used as the overall performance index while the estimated state variables of Generator 5 are taken for illustration. The H-infinity UKF proposed in this chapter and the GM-UKF [92] are used for comparison.

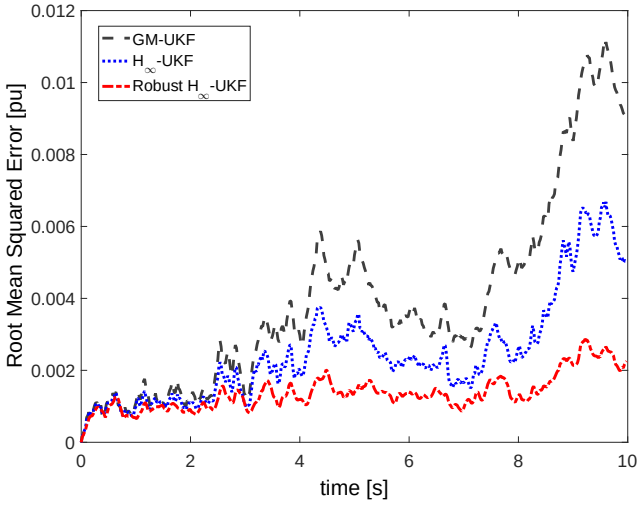


Figure 7.2: Root-mean-squared error of the GM-UKF, the H-infinity UKF and the robust H-infinity UKF with model uncertainties. Here, it is assumed that after the disturbance is applied, the transient reactances of Generator 5 deviate from the nominal values by a percentage of 10%.

7.4.1 Case 1: Non-Gaussian Process and Measurement Noise

Due to communication channel noises and GPS synchronization errors, the measurement noises associated with PMUs may deviate from the Gaussian assumption. Therefore, if part of the PMU measurements are taken as system inputs, the system process noise is no longer Gaussian. Furthermore, due to the changes of system operation conditions and non-stationary ambient environment, the assumed model to approximate the true one may yield unknown characteristics. In order to assess the sensitivity of each approach to the deviations from model and measurement assumptions, both system process and measurement noise are assumed to be non-Gaussian. Since the Gaussian-mixture model can be used to

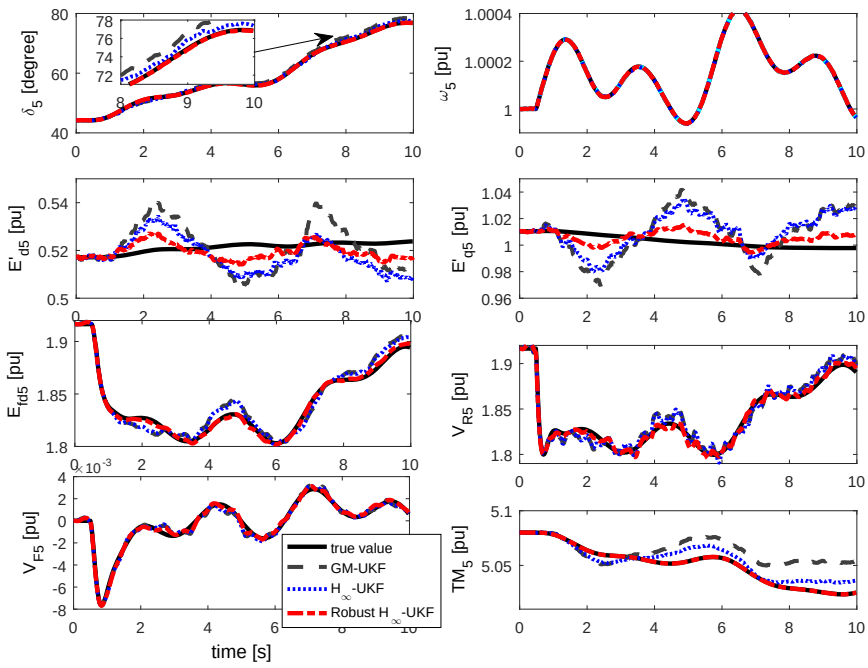


Figure 7.3: Estimated state variables by the GM-UKF, the H-infinity UKF and the robust H-infinity UKF with model uncertainties.

approximate any distribution, it is assumed in this chapter. Specifically, we assume they follow a Gaussian mixture model, where 10% of the data are drawn with covariance matrices $\boldsymbol{Q} = 5\times10^{-5}\boldsymbol{I}_{9\times9}$ and $\boldsymbol{R} = 5\times10^{-5}\boldsymbol{I}_{2\times2}$ while the true covariance matrices are $\boldsymbol{Q} = 10^{-6}\boldsymbol{I}_{9\times9}$ and $\boldsymbol{R} = 10^{-6}\boldsymbol{I}_{2\times2}$. The root-mean-squared errors of each method are displayed in Fig. 7.1. It can be observed that the H-infinity UKF is sensitive to non-Gaussian noises and provides the worst estimation results. The reason is that it is based on the weighted least squares criterion in the Krein space and thus achieves very low statistical efficiency. By contrast, both the GM-UKF and the proposed robust H-infinity UKF are able to withstand the non-Gaussian noise thanks to robustness of GM-estimator. However, the proposed robust H-infinity UKF outperforms the GM-UKF slightly. This is because the non-Gaussian noise is actually unknown and the proposed robust H-infinity UKF first leverages the robustness of GM-estimator to suppress it and then relies on the H-infinity criterion to further bound these uncertainties, yielding improved statistical efficiency.

7.4.2 Case 2: Dynamical Model Uncertainties

Because of aging processes, variations of the machine temperature during its operation, the effect of saturation on generator inductances, the reactance and transient reactance of a synchronous generator may change significantly. In this chapter, it is assumed that after the system event, the transient reactances of the generators deviate from the nominal values by a percentage of 10% random error simulated by the Gaussian distribution. The rational of choosing transient reactance for simulation is as follows: according to the work [143, 144], a few critical parameters of the generators are in charge of the system response. In other words, the changes of these parameter values are able to affect the generator response significantly while the changes on other parameters induce negligible difference. By using the trajectory sensitivity analysis approach [143, 144], the transient reactances and the gain of excitor are identified as these critical parameters. We anticipate that the uncertainties of them would impose huge challenges to the dynamic state estimator.

Figs. 7.2-7.3 display the root-mean-squared errors of the GM-UKF, the H-infinity UKF and the robust H-infinity UKF with model uncertainties and the estimated state variables of Generator 5, respectively. It is observed from the results that the GM-UKF shows the highest sensitivity to large dynamical model uncertainties and provides biased estimation results; by contrast, thanks to the H-infinity criterion, the H-infinity UKF is able to bound the uncertainties to a certain degree and achieves better results than the GM-UKF. Here, we would like to emphasize that the estimation error covariance matrix of the GM-UKF is robust thanks to the weights provided by the projection statistics as well as the GM-estimator. In other words, the GM-UKF modifies the estimation error covariance matrix at each iteration to achieve some robustness to system uncertainties, which is similar to the H-infinity UKF. This justifies that why the H-infinity UKF does not outperform significantly the GM-UKF. The proposed robust H-infinity UKF leverages both the robustness of estimation error covariance matrix and the H-infinity criterion, yielding the best estimation results. Note that the estimation error of rotor angle by the GM-UKF is about several degrees, which is a lot and may lead to large errors to the rotor angle-based applications, such as rotor angle stability assessment, out-of-step protection, causing serious concerns to the system security. However, this is not the case for our robust H-infinity filter that can reliably always track the system dynamic states.

7.4.3 Case 3: Observation Outliers

Due to imperfect phasor synchronization, the saturation of metering current transformers and cyber attacks, to name a few, gross errors can occur in the PMU measurements [27, 38]. To test the robustness of three methods to observation outliers, the measured real and reactive powers of Generator 5 is contaminated with 20% error from t=4s to t=6s. The root-mean-squared errors and the estimated state variables of Generator 5 for the three methods are shown in Fig. 7.4 and Fig. 7.5, respectively. It can be found that the Krein space weighted

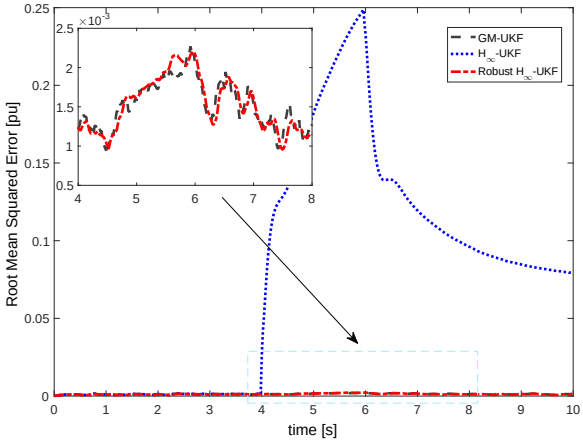


Figure 7.4: Root-mean-squared error of the GM-UKF, the H-infinity UKF and the robust H-infinity UKF in presence of observation outliers. Here, it is assumed that the real and reactive power measurements of Generator 5 are corrupted with 20% errors from t=4s to t=6s.

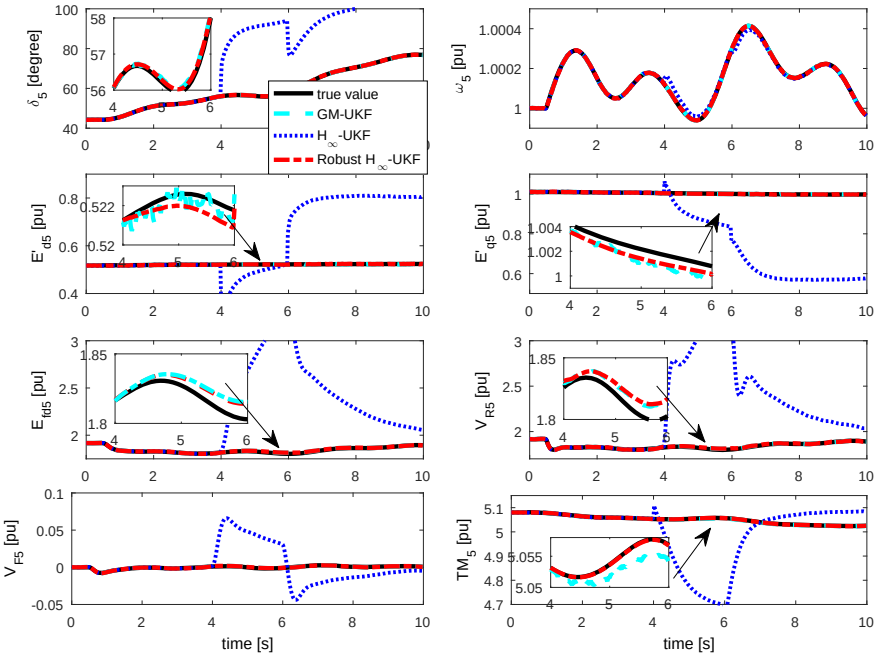


Figure 7.5: Estimated state variables of the GM-UKF, the H-infinity UKF and the robust H-infinity UKF in the presence of observation outliers from t=4s to t=6s.

Table 7.1: Average Computing Times of the GM-UKF, the H-infinity UKF and the Robust H-infinity UKF at Each PMU Scan

Cases	GM-UKF	H-infinity UKF	Robust H-infinity UKF
Case 1	1.34ms	1.20ms	1.46ms
Case 2	1.42ms	1.23ms	1.56ms
Case 3	1.55ms	1.30ms	1.83ms

least squares estimator-based H-infinity lacks of robustness to outliers, yielding significantly biased state estimates. By contrast, thanks to the weights provided by projection statistics and the GM-estimator, both the GM-UKF and the robust H-infinity UKF are able to suppress outliers, achieving comparable performance. According to the results shown in Cases 1-3, it is now clear that the proposed robust H-infinity UKF achieves the desired performance, that is, leveraging the H-infinity criterion to bound system uncertainties while relying on the robustness of GM-estimator to filter out non-Gaussian noise and suppress outliers.

7.4.4 Computational Efficiency

To assess the computational efficiency of each method, the average computing times of the GM-UKF, the H-infinity UKF and the robust H-infinity UKF for Cases 1-3 at each PMU scan are presented in Table. 7.1. All the tests are performed on a PC with Intel Core i5, 2.50 GHz, 8GB of RAM. It is found from this table that the H-infinity UKF is the most computational efficient approach, followed by the GM-UKF. Although the robust H-infinity UKF is the most time consuming one, its difference with other two is negligible. Furthermore, all three methods spend much less time than the PMU scan rate, which is 16.7ms. Thus, they can be implemented for power system online applications.

7.5 Conclusions

In this chapter, a novel theoretical framework for robust dynamic state estimation is proposed that integrates both robust statistics and robust control theory. The GM-estimator, the unscented Kalman filter (UKF), and the H-infinity filter are integrated into a unified framework to yield the general robust H-infinity UKF. The latter is able to handle large system uncertainties as well as suppress outliers while achieving good statistical efficiency under a broad range of non-Gaussian process and observation noise. Specifically, it leverages the H-infinity criterion to bound system uncertainties while relying on the robustness of GM-estimator to filter out non-Gaussian noise and suppress outliers. We also show that the H-infinity UKF is based on the Krein space least squares estimator and thus lacks robustness to outliers and non-Gaussian noise. Comparative results show that our proposed

robust H-infinity UKF outperforms the H-infinity UKF and the GM-UKF in terms of statistical efficiency and robustness to outliers, non-Gaussian noise and model uncertainties. As a future work, we will improve the statistical efficiency of the proposed robust H-infinity UKF by resorting to other robust control strategies such as sliding mode control and probabilistic robust control. The breakdown point of the robust H-infinity UKF will be investigated as well.

Chapter 8

Conclusions and Future Work

In this dissertation, we have pioneered a general theoretical framework that advances both robust statistics and robust control theory for robust dynamic state and parameter estimation of a cyber-physical system. Specifically, the generalized maximum-likelihood-type (GM)-estimator, the extended and unscented Kalman filter, and the H-infinity filter are integrated into a unified framework to yield various centralized and decentralized robust dynamic state estimators. These new estimators are able to handle large system uncertainties as well as suppress three types of outliers while achieving good statistical efficiency under a broad range of non-Gaussian process and observation noise. In the chapter 2, a centralized robust GM-IEKF is proposed for estimating power system state dynamics when subjected to disturbances. The proposed GM-IEKF dynamic state estimator is able to track system transients in a faster and more reliable way than the conventional EKF and the UKF thanks to its batch-mode regression form and its robustness to innovation and observation outliers, even in position of leverage. However, it achieves low statistical efficiency for non-Gaussian process and measurement noise. To this end, in chapter 3, the theory of a new robust GM-UKF is developed that is able to suppress observation and innovation outliers while filtering out non-Gaussian process and measurement noise. The state estimates and residuals of our GM-UKF are proved to be roughly Gaussian, allowing the sigma points to reliably approximate the mean and the covariance matrices of the predicted and corrected state vectors. The asymptotic error covariance matrix of the GM-UKF state estimates is derived from the total influence function. The robustness our GM-UKF to outliers and cyber attacks is investigated; it is shown that it can handle *at least* 25% of incorrect PMU measurements and predicted state variables. The GM-UKF is able to handle two types of outliers, namely the innovation and observation outliers, but is fails to suppress structural outliers. To address this challenge, in Chapter 4, a fast and robust DSE using multiple hypothesis testing is proposed to detect, identify and suppress three types of outliers. Specifically, three hypotheses corresponding to the occurrence of three types of outliers are assumed by constructing three innovation matrices; these matrices are made up by time-correlated innovation vectors and/or predicted states and/or measurements; then projection statistics are applied to each

of the innovation matrix and its calculated projection values are checked by a statistical test to validate the assumed hypothesis. The identified outliers are further suppressed by a generalized maximum-likelihood type (GM)-estimator. The proposed robust decentralized DSE is further applied for the estimation of bus frequency in Chapter 5. It is worth pointing out that although GM-UKF achieves good robustness to outliers and non-Gaussian process and measurement noise, it produces biased state estimates in presence of system uncertainties. To address this issue, the Chapter 6 presents a decentralized H-infinity UKF that leverages the strength of the H-infinity criteria developed in robust control for handling system uncertainties with the advantage of the UKF for addressing strong model nonlinearities. However, the decentralized H-infinity UKF still lacks of robustness to outliers and non-Gaussian noise. To this end, the GM-estimator, the unscented Kalman filter (UKF), and the H-infinity filter are integrated into a unified framework to yield the general robust H-infinity UKF. The latter is able to handle large system uncertainties as well as suppress outliers while achieving good statistical efficiency under a broad range of non-Gaussian process and observation noise. Specifically, it leverages the H-infinity criterion to bound system uncertainties while relying on the robustness of GM-estimator to filter out non-Gaussian noise and suppress outliers. We also show that the H-infinity UKF is based on the Krein space least squares estimator and thus lacks robustness to outliers and non-Gaussian noise.

Based on the results presented in this dissertation, several extensions are identified to be pursued in the future. They are briefly discussed below:

- **Enhancement of the Proposed Theoretical Framework**: the combination of GM-estimator, the H-infinity filter and the UKF opens the door for the development of more advanced robust dynamic state estimators that achieve good trade-off between robustness to outliers, non-Gaussian noise and system uncertainties and statistical efficiency. For example, other filters that share the similar structure of UKF can be leveraged to derive their corresponding robust H-infinity filters, such as divided difference filter, quadrature Kalman filter, cubature Kalman filter, EnKF and so on [139–141]. We will also investigate the concepts of finite-sample local and global breakdown points for nonlinear dynamical structured regression models and determine the breakdown points of the developed dynamic state estimators.

- **Extension to Model Parameter Error Identification and Correction**: the developed robust dynamic state estimators mainly focus on synchronous generators. They can be extended for the model validation and calibration of various system components, such as inverter-based renewable generation units, dynamic loads, transformers, etc. In addition, the work can be used to detect the failures of various control devices, including exciters, power system stabilizers, transformers and so on.

- **Design of Power Oscillation Damping Controller**: both local and wide-area dynamic state variables obtained from a robust DSE are available, enabling the design of effective local and wide-area controllers; for instance, the estimated rotor speeds and

other estimated state variables can be used as input control signals of FACTS devices to ensure global asymptotic stability of a power system by satisfying the Lyapunov sufficient conditions for stability; the coherent generator groups can be identified using estimated rotor angles so that effective wide-area control can be implemented; the gains of the exciter/PSS can be tuned optimally online through the proposed robust dynamic state estimator, contributing to the enhancement of system stability.

- **Detection of Cyber Attacks**: thanks to the robustness of the developed approaches, they are naturally resistant to cyber attacks on the system inputs and online measurements. We can extend the proposed estimators to detect and distinguish different types of attacks on an cyber-physical system. In addition, novel detection approaches can be proposed by appropriately combining the robust statistics with robust control theory.
- **Dynamic State Estimation of Active Distribution System and Micro-grids**: the concepts, tools and methodologies will be extended from transmission system to the active distribution system and the micro-grid. It is of interest to design robust controllers for the enhancement of micro-grid stability in both grid-connected and island modes. This will promote a new energy management system for micro-grid operation and control.

Chapter 9

Publications of This Dissertation

9.1 Published Journal Papers

- **J. B. Zhao**, L. Mili, "A Robust Generalized-Maximum Likelihood Unscented Kalman Filter for Power System Dynamic State Estimation," ***IEEE Journal of Selected Topics in Signal Processing***, 2018, in press.
- **J. B. Zhao**, L. Mili, "Power System Robust Decentralized Dynamic State Estimation Based on Multiple Hypothesis Testing," ***IEEE Transactions on Power Systems***, 2017, in press.
- **J. B. Zhao**, L. Mili, F. Milano, "Robust Frequency Divider for Power System Online Monitoring and Control," ***IEEE Transactions on Power Systems***, 2017, in press.
- **J. B. Zhao**, L. Mili, "Robust Unscented Kalman Filter for Power System Dynamic State Estimation with Unknown Noise Statistics," ***IEEE Transactions on Smart Grid***, 2017, in press.
- **J. B. Zhao**, L. Mili, "A Decentralized H-infinity Unscented Kalman Filter for Dynamic State Estimation Against Uncertainties," ***IEEE Transactions on Smart Grid***, 2018, in press.
- **J. B. Zhao**, L. Mili, "Vulnerability of the Largest Normalized Residual Statistical Test to Leverage Points," ***IEEE Transactions on Power Systems***, 2018, in press.
- **J. B. Zhao**, L. Mili, R. Pires, "Statistical and Numerical Robust State Estimator for Heavily Loaded Power Systems," ***IEEE Transactions on Power Systems***, 2018, in press.

- **J. B. Zhao**, L. Mili, M. Wang, "A Generalized False Data Injection Attacks Against Power System Nonlinear State Estimator and Countermeasures," ***IEEE Transactions on Power Systems***, 2018, in press.
- **J. B. Zhao**, S. Wang, R. Huang, Z. Huang, L. Mili, "A Robust State Estimation Framework Considering Measurement Correlations and Imperfect Synchronization," ***IEEE Transactions on Power Systems***, 2018, in press.
- **J. B. Zhao**, L. Mili, "A Framework for Robust Hybrid State Estimation with Unknown Measurement Noise Statistics," ***IEEE Transactions on Industrial Informatics***, 2017, in press.
- **J. B. Zhao**, "Power System Dynamic State Estimation Considering Measurement Correlations," ***IEEE Transactions on Energy Conversion***, vol. 32, no. 4, pp. 1630–1632, 2017.
- S. Wang, **J. B. Zhao**, Z. Huang, R. Diao, "Assessing Gaussian Assumption of PMU Measurement Error Using Field Data," ***IEEE Transactions on Power Delivery***, 2017, in press.
- **J. B. Zhao**, "Dynamic State Estimation with Model Uncertainties Using H-infinity Extended Kalman Filter. ***IEEE Transactions on Power Systems***, vol. 33, no. 1, pp. 1099-1100, 2018.
- X. Wei, **J. B. Zhao**, T. Huang, "A Novel Cascading Fault Network for Assessing Transmission Network Vulnerability," ***IEEE Transactions on Power Systems***, vol. 33, no. 3, pp. 2995-3000, 2018.
- **J. B. Zhao**, L. Mili, "Sparse State Recovery vs Generalized Maximum-likelihood Estimator of a Power System," ***IEEE Transactions on Power Systems***, vol. 33, no. 1, pp. 1104-1106, 2018.
- **J. B. Zhao**, N. Marcos, L. Mili, "A Robust Iterated Extended Kalman Filter for Power System Dynamic State Estimation using PMU Measurements," ***IEEE Transactions on Power Systems***, vol. 32, no. 4, pp. 3205-3216, 2017.

9.2 Submitted Journal Papers

- **J. B. Zhao**, S. Fliscounakis, P. Panciatici, L. Mili, "Robust Parameter Estimation of the French Power System Using Field Data," ***IEEE Transactions on Smart Grid*** (Under review).
- **J. B. Zhao**, L. Mili, "Robust Power System Dynamic State Estimator Against Cyber Attacks," ***IEEE Transactions on Power Systems*** (to be submitted).

- Y. Xu, L. Mili, **J. B. Zhao**, " Propagating Uncertainty in Power System Dynamic Simulations Using Polynomial Chaos," ***IEEE Transactions on Power Systems*** (Second round of review).

- **J. B. Zhao**, L. Mili, "Robust H-infinity Unscented Kalman Filter: Theory and Application to Power Systems Dynamic State Estimation," ***IEEE Transactions on Signal Processing*** (Under review).

9.3 Peer Reviewed Conference Papers

- **J. B. Zhao**, L. Mili, "State Estimation for Heavily Loaded System: A Comparative Study," *in Proc. of IEEE PES General Meeting*, Portland, pp. 1-5, August 2018.

- **J. B. Zhao**, L. Mili, A. Abdelhadi, "Robust Dynamic State Estimator to Outliers and Cyber Attacks," *in Proc. of IEEE PES General Meeting*, Chicago, pp. 1-5, July 2017.

- N. Marcos, **J. B. Zhao**, L. Mili, "Robust Dynamic State Estimation using PMU Measurements," *in Proc. of IEEE PES General Meeting*, Boston, CO, pp. 1-5, July 2016.

Bibliography

[1] Y. Susuki and I. Mezić, "Nonlinear Koopman modes and coherency identification of coupled swing dynamics," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 1894-1904, Nov. 2011.

[2] Y. Susuki and I. Mezić, "Nonlinear Koopman modes and a precursor to power system swing instabilities," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1182-1191, Aug. 2012.

[3] E. Barocio, B. C. Pal, N. F. Thornhill and A. R. Messina, "A dynamic mode decomposition framework for global power system oscillation analysis," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 2902-2912, Nov. 2015.

[4] J. Pierre, D. Trudnowski, et. al, "Identification of electromechanical modes in power systems," IEEE Task Force Rep., Jun. 2012.

[5] J. F. Hauer, C. J. Demeure and L. L. Scharf, "Initial results in Prony analysis of power system response signals," *IEEE Transactions on Power Systems*, vol. 5, no. 1, pp. 80-89, Feb 1990.

[6] J. B. Zhao, G. X. Zhang, "A robust prony method against synchrophasor measurement noise and outliers," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2484-2486, 2017.

[7] L. Jiang, Q. H. Wu, J. Y. Wen, "Decentralized nonlinear adaptive control for multi-machine power systems via high-gain perturbation observer," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 51, no. 10, pp. 2052-2059, Oct. 2004.

[8] H. Liu, Z. Hu and Y. Song, "Lyapunov-based decentralized excitation control for global asymptotic stability and voltage regulation of multi-machine power systems," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2262-2270, Nov. 2012.

[9] Hui Ni, G. T. Heydt and L. Mili, "Power system stability agents using robust wide area control," *IEEE Transactions on Power Systems*, vol. 17, no. 4, pp. 1123-1131, Nov 2002.

[10] R. Yousefian, S. Kamalasadan, "Energy function inspired value priority based global wide-area control of power grid," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 552-563, 2018.

[11] C. W. Taylor, D. C. Erickson, K. E. Martin, R. E. Wilson and V. Venkatasubramanian, "WACS-wide-area stability and voltage control system: R&D and online demonstration," *Proceedings of the IEEE*, vol. 93, no. 5, pp. 892-906, May 2005.

[12] A. P. Meliopoulos, *et al.*, "Dynamic state estimation based protection: status and promise," *IEEE Transactions on Power Delivery*, vol. 32, no. 1, pp. 320-330, Feb. 2017.

[13] E. Farantatos, R. Huang, G. J. Cokkinides and A. P. Meliopoulos, "A predictive generator out-of-step protection and transient stability monitoring scheme enabled by a distributed dynamic state estimator," *IEEE Transactions on Power Delivery*, vol. 31, no. 4, pp. 1826-1835, Aug. 2016.

[14] P. Kundur, "*Power system stability and control*," New York: McGraw-Hill, 1994.

[15] M. A. M. Ariff, B. C. Pal, "Adaptive protection and control in the power system for wide-area blackout prevention," *IEEE Transactions on Power Delivery*, vol. 31, no. 4, pp. 1815–1825, 2016.

[16] B. Shrestha, G. Ramakrishna, S. Mohindar, "Out-of-step protection using state-plane trajectories analysis," *IEEE Transactions on Power Delivery*, vol. 28, no. 2, pp. 1083-1093, 2013.

[17] H. Modir, R. Schlueter, "A dynamic state estimator for dynamic security assessment," *IEEE Transactions on Power Systems*, vol. 11, no. 100, pp. 4644–4652, 1981.

[18] M. He, J. Zhang, V. Vittal, "Robust online dynamic security assessment using adaptive ensemble decision-tree learning," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4089–4098, 2013.

[19] "Next generation on-line dynamic security assessment," Final report, Power Systems Engineering Research Center, 2011.

[20] "Power plant dynamic model verification using PMUs," Reliability Guideline, NERC, 2016.

[21] Z. Huang, K. Schneider, and J. Nieplocha, "Feasibility studies of applying Kalman filter techniques to power system dynamic state estimation," *Proceedings of 8th International Power Engineering Conference*, Singapore, Dec. 2007, pp. 376–382.

[22] E. Ghahremani, I. Kamwa, "Dynamic state estimation in power system by applying the extended Kalman filter with unknown inputs to phasor measurements," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2556–2566, Nov. 2011.

[23] E. Ghahremani, I. Kamwa, "Local and wide-area PMU-based decentralized dynamic state estimation in multi-machine power systems," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 547–562, 2016.

[24] L. Fan and Y. Wehbe, "Extended Kalman filtering based real-time dynamic state and parameter estimation using PMU data," *Electrical Power System Research*, vol. 103, pp. 168–177, Oct. 2013.

[25] S. Wang, W. Gao, A. P. S. Meliopoulos, "An alternative method for power system dynamic state estimation based on unscented transform," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 942–950, May 2012.

[26] A. K. Singh, B. C. Pal, "Decentralized dynamic state estimation in power systems using unscented transformation," *IEEE Transactions on Power Systems*, vol. 29, no. 2, pp. 794–804, Sep 2014.

[27] A. Rouhani, A. Abur, "Linear phasor estimator assisted dynamic state estimation," *IEEE Transactions on Smart Grid*, vol. 9, no. 1, pp. 211- 219, Jan. 2018.

[28] N. Zhou, D. Meng, S. Lu, "Estimation of the dynamic states of synchronous machines using an extended particle filter," *IEEE Transactions on Power Systems*, vol. 28, no. 4, 4152–4161, 2013.

[29] K. Emami, T. Fernando, H. H. C. Iu, H. Trinh, K. P. Wong, "Particle filter approach to dynamic state estimation of generators in power systems," *IEEE Transactions on Power Systems*, vol. 30, no. 5, pp. 2665-2675, Sept. 2015.

[30] Y. Cui, R. Kavasseri, "A particle filter for dynamic state estimation in multi-machine systems with detailed models," *IEEE Transactions on Power Systems*, vol. 30, no. 6, 3377–3385, 2015.

[31] J. Chen, R. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems*, Boston, MA: Kluwer Academic Publishers, 1999.

[32] D. Simon, *Optimal State Estimation. Hoboken*, NJ: Wiley, 2006.

[33] M. Grimble and M. Johnson, "H_∞robust control design–A tutorial review," *Computing & Control Engineering Journal*, pp. 275–282, 1991.

[34] H. Xu, S. Mannor, "A kalman filter design based on the performance/robustness tradeoff," *IEEE Transactions on Automatic Control*, vol. 54, no. 5, pp. 1171–1175, 2009.

[35] Z. Huang, N. Zhou, R. Diao, S. Wang, S. Elbert, D. Meng, S. Lu, "Capturing real-time power system dynamics: opportunities and challenges," *Proceedings of IEEE Power Engineering Soceity General Meeting*, July 2015.

[36] N. Zhou, Z. Huang, D. Meng, "Capturing dynamics in the power grid: formulation of dynamic state estimation through data assimilation. Technical Report PNNL-23213, Pacific Northwest National Laboratory, 2014.

[37] M. Gandhi, L. Mili, "Robust Kalman filter based on a generalized maximum-likelihood-type estimator, *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2509–2520, 2010.

[38] J. B. Zhao, M. Netto, L. Mili, "A Robust iterated extended Kalman filter for power system dynamic state estimation," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3205-3216, 2017.

[39] K. Martin, J. Hauer, T. Faris, "PMU testing and installation considerations at the Bonneville power administration," *Proceedings of IEEE Power Engineering Soceity General Meeting*, Jun. 2007, pp. 1–6.

[40] K. D. Jones, A. Pal, J. S. Thorp, "Methodology for performing synchrophasor data conditioning and validation," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1121–1130, May. 2015.

[41] L. Vanfretti, J. H. Chow, S. Sarawgi, B. Fardanesh, "A phasor-data-based state estimator incorporating phase bias correction," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 11–119, 2011.

[42] D. N. Kosterev, C. W. Taylor, W. A. Mittelstadt, "Model validation for the August 10, 1996 WSCC system outage," *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 967–979, 1999.

[43] Z. Huang, P. Du, D. Kosterev, S. Yang, "Generator dynamic model validation and parameter calibration using phasor measurements at the point of connection," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1939–1949, 2013.

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

[45] N. Zhou, D. Meng, Z. Huang, G. Welch, "Dynamic state estimation of a synchronous machine using PMU data: A comparative study," *IEEE Transactions on Smart Grid*, vol. 6, no. 1, pp: 450–460, 2015.

[46] A. Teixeira, D. Perez, H. Sandberg, K. H. Johansson, "Attack models and scenarios for networked control systems," *Proceedings of the 1st ACM international conference on High Confidence Networked Systems*, pp: 55-64, 2012.

[47] L. Mili, C. W. Coakley, "Robust estimation in structured linear regression," *The Annals of Statistics*, vol. 24, no. 6, pp. 2593–2607, 1996.

[48] A. J. Stromberg, D. Ruppert, "Breakdown in nonlinear regression," *Journal of the American Statistical Association*, vol. 87, no. 420, pp. 991–997, 1992.

[49] S. Sakata, H. White, "An alternative definition of finite-sample breakdown point with applications to regression model estimators," *Journal of the American Statistical Association*, vol. 90, no. 431, pp. 1099–1106, 1995.

[50] E. A. Wan, R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," *Adaptive Systems for Signal Processing, Communications, and Control Symposium*, pp. 153-158, 2000.

[51] S. S. Haykin. *Kalman filtering and neural networks*. New York: Wiley, 2001.

[52] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automactic Control*, vol. 45, no. 3, pp. 477–482, Mar. 2000.

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

[54] T. Lefebvre, H. Bruyninckx, J. De Schuller, "Comment on "A new method for the nonlinear transformation of means and covariances in filters and estimators"," *IEEE Transactions on Automatic Control*, vol. 47, no. 8, pp. 1406-1409, 2002.

[55] J. B. Zhao, G. X. Zhang, M. L. Scala, "PMU based robust dynamic state estimation method for power systems," in *Proceedings of IEEE Power Engineering Soceity General Meeting*, pp. 26–30, 2015.

[56] E. Farantatos, G. K. Stefopoulos, G. J. Cokkinides, A. P. Meliopoulos, "PMU-based dynamic state estimation for electric power systems," in *Proc. IEEE Power Eng. Soc. General Meeting*, pp. 1–8, 2009.

[57] J. L. Crassidis, F. L. Markley, Y. Cheng, "Survey of nonlinear attitude estimation methods," *Journal of guidance, control, and dynamics*, vol. 30, no. 1, pp. 12–28, 2007.

[58] F. Gustafsson, G. Hendeby, "Some relations between extended and unscented Kalman filters," *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp: 545-555, 2012.

[59] E. Ghahremani, I. Kamwa, "Online state estimation of a synchronous generator using unscented Kalman filter from phasor measurements units," *IEEE Transactions on Energy Conversion*, vol. 26, no. 4, pp. 1099–1108, Dec. 2011.

[60] M. A. Pai, *Energy function analysis for power system stability*. Springer Science & Business Media, 1989.

[61] J. B. Zhao, G. X. Zhang, K. Das, Ge. N. Korres, N. M. Manousakis, A. K. Sinha, Z. Y. He, "Power system real-time monitoring by using PMU-based robust state estimation method," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 300-309, 2016.

[62] J. B. Zhao, G. X. Zhang, M. L. Scala, Z. Wang, "Enhanced robustness of state estimator to bad data processing through multi-innovation analysis", *IEEE Transactions on Industrail Informatics*, vol. 13, no. 4, pp. 1610-1619, 2017.

[63] N. Zhou, S. Lu, R. Singh, et al, "Calibration of reduced dynamic models of power systems using phasor measurement unit (PMU) data," *North American Power Symposium (NAPS)*, pp: 1–7, 2011.

[64] N. Zhou, Z. Huang, D. Meng, et al. "Capturing dynamics in the power grid: formulation of dynamic state estimation through data assimilation," Technical Report PNNL-23213, Pacific Northwest National Laboratory, 2014.

[65] L. Mili, M. Cheniae, N. Vichare, and P. Rousseeuw, "Robust state estimation based on projection statistics," *IEEE Transactions on Power Systems*, vol. 11, no. 2, p. 1118–1127, 1996.

[66] L. Thomas, L. Mili, "A robust GM-estimator for the automated detection of external defects on barked hardwood logs and stems," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3568–3576, 2007.

[67] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust Statistics: The Approach Based on Influence Functions*. New York: John Wiley & Sons, Inc., 1986.

[68] U. Hammes, E. Wolsztynski, A. M. Zoubir, "Robust tracking and geolocation for wireless networks in NLOS environments," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 5, pp. 889–901, 2009.

[69] U. Hammes, A. M. Zoubir, "Robust MT tracking based on M-estimation and interacting multiple model algorithm," *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3398–3409, 2011.

[70] IEEE PES TF on Benchmark System for Stability Controls, "Benchmark systems for small-signal stability analysis and control," Aug. 2015.

[71] L. Mili, G. Steeno, F. Dobraca, D. French, "A robust estimation method for topology error identification," *IEEE Transactions on Power Systems*, vol. 14, no. 4, p. 1469–1476, 1999.

[72] C. Rao, *Linear Statistical Inference and its Applications*. New York: Wiley-Interscience, 1973.

[73] I. Kamwa, R. Grondin, and Y. Hebert, "Wide-area measurement based stabilizing control of large power systems-a decentralized/hierarchical approach," *IEEE Transactions on Power Systems*, vol. 16, no. 1, pp. 136–153, Feb. 2001.

[74] G. Evensen, "Sequential data assimilation with a nonlinear quasigeostrophic model using Monte Carlo methods to forecast error statistics," Journal of Geophysical Research, vol. 99, no. 5, pp. 143-162, May 1994.

[75] X. Wang, N. Cui, J. Guo, "Huber-based unscented filtering and its application to vision-based relative navigation," *IET radar, sonar and navigation*, vol. 4, no. 1, pp. 134–141, 2010.

[76] L. Y. Chang, B. Hu, G. K. Chang, et al, "Huber-based novel robust unscented Kalman filter," *IET Science, Measurement & Technology*, vol. 6, no. 6, pp. 502–509, 2012.

[77] L. B. Chang, K. Li, "Unified form for the robust Gaussian information filtering based on M-estimate," *IEEE Signal Processing Letters*, vol. 24, no. 4, pp. 412-416, 2017.

[78] W. Li, S. Sun, Y. Jia, J. Du, "Robust unscented Kalman filter with adaptation of process and measurement noise covariances," *Digital Signal Processing*, vol. 48, pp. 93-103, 2016.

[79] J. B. Zhao, L. Mili, "A robust generalized-maximum likelihood unscented Kalman filter for power system dynamic state estimation," *IEEE Journal of Selected Topics in Signal Processing*, 2018.

[80] J. B. Zhao, L. Mili, "Robust unscented Kalman filter for power system dynamic state estimation with unknown noise statistics," *IEEE Transactions on Smart Grid*, DOI: 10.1109/TSG.2017.2761452, 2017.

[81] S. Wang, J. B. Zhao, Z. Huang, R. Diao, "Assessing Gaussian assumption of PMU measurement error using field data," *IEEE Transactions on Power Delivery*, 2017, in press.

[82] P. J. Huber, *Robust Statistics*. New York: Wiley, 1981.

[83] L. Fernholz, "Von mises calculus for statistical functionals," Lecture Notes in Statistics. New York: Springer-Verlag, 1983, vol. 19.

[84] P. W. Sauer, M. A. Pai, Power System Dynamics and Stability, Prentice-Hall, 1998.

[85] "Modeling of gas turbines and steam turbines in combined cycle power plants," CIGRE report, Task Force C4.02.25, 2003.

[86] Z. Huang, P. Du, D. Losterev, B. Yang, "Application of extended Kalman filter techniques for dynamic model parameter calibration," in *Proceedings of IEEE Power Energy Soceity General Meeting*, 2009, pp. 1–6.

[87] G. Anagnostou, B. C. Pal, "Derivative-free Kalman filtering based approaches to dynamic state estimation for power systems with unknown inputs," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 116-130, 2018.

[88] J. B. Zhao, L. Mili, "Power system robust decentralized dynamic state estimation based on multiple hypothesis testing," *IEEE Transactions on Power Systems*, DOI: 10.1109/TPWRS.2017.2785344, 2017, in press.

[89] A. Rouhani, A. Abur, "Observability analysis for dynamic state estimation of synchronous machines," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3168-3175, 2017.

[90] NERC Reliability Guideline, "PMU Placement and Installation," 2016.

[91] A. K. Singh, B. C Pal, "Decentralized control of oscillatory dynamics in power systems using an extended LQR," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 1715-1728, 2016.

[92] J. B. Zhao, L. Mili, "Robust unscented Kalman filter for power system dynamic state estimation with unknown noise statistics," *IEEE Transactions on Smart Grid*, 2017.

[93] E. Scholtz, V. D. Donde, and J. C. Tournier, "Parallel computation of dynamic state estimation for power system," U.S. Patent Application 13/832,670, Mar. 15, 2013.

[94] Setting-less Protection: Laboratory Testing-Final Report, Power Systems Engineering Research Center, 2014.

[95] IEEE Std 1110-2002, IEEE Guide for Synchronous Generator Modeling Practices and Applications in Power System Stability Analyses, IEEE Power and Engineering Society.

[96] National Grid Frequency Response Working Group, Frequency Response Technical Sub-Group Report, Nov. 2011, Tech. Rep. [Online].

[97] J. Van de Vyver, J. D. M. De Kooning, B. Meersman, L. Vandevelde, T. L. Vandoorn "Droop control as an alternative inertial response strategy for the synthetic inertia on wind turbines," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1129-1138, 2016.

[98] H. Ye, P. Pei, Z. Qi, "Analytical modeling of inertial and droop responses from a wind farm for short-term frequency regulation in power systems," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3414–3423, 2016.

[99] J. B. Zhao, S. Wang, L. Mili, R. Huang, Z. Huang, "A robust state estimation framework considering measurement correlations and imperfect synchronization," *IEEE Transactions on Power Systems*, DOI: 10.1109/TPWRS.2018.2790390, 2018, in press.

[100] C. Zhao, U. Topcu, S. H. Low, "Optimal load control via frequency measurement and neighborhood area communication," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3576-3587, 2013.

[101] Y. Wen, W. Li, G. Huang, X. Liu, "Frequency dynamics constrained unit commitment with battery energy storage," *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 5115–5125, 2016.

[102] IEEE Task Force on Load Representation for Dynamic Performance, "Load representation for dynamic performance analysis of power systems," *IEEE Transactions on Power Systems*, vol. 8, no. 2, pp. 472–482, 1993.

[103] L. Wang, D. Z. Fang, T. S. Chung, "New techniques for enhancing accuracy of EMTP/TSP hybrid simulation algorithm," in *Proc. IEEE International Conference Electrical Utility Deregulation, Restructuring Power Technology*, pp. 734–739, 2004.

[104] G. Radman, M. A. Tabrizi, "Simulation of wide area frequency measurement from phasor measurement units (PMUs) or frequency disturbance recorders (FDRs)," Oct. 2010 [Online].

[105] M. Lai, M. Nakano, G. Hsieh, "Application of fuzzy logic in the phase-locked loop speed control of induction motor drive," *IEEE Transactions on Industrial Electronics*, vol. 43, no. 6, pp. 630-639, 1996.

[106] J. Nutaro, V. Protopopescu, "Calculating frequency at loads in simulations of electromechanical transients," *IEEE Transactions on Smart Grid*, vol. 3, no. 1, pp. 233–240, 2012.

[107] F. Milano, A. Ortega, "Frequency divider," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1493-1501, 2017.

[108] J. B. Zhao, L. Mili, F. Milano, "Robust frequency divider for power system online monitoring and controls," *IEEE Transactions on Power Systems*, DOI: 10.1109/TPWRS.2017.2785348, 2017, in press.

[109] Y. Zhang, P. Markham, *et al.*, "Wide-area frequency monitoring network (FNET) architecture and applications," *IEEE Transactions on Smart Grid*, vol. 1, no. 2, pp. 159-167, 2010.

[110] A. G. Phadke, B. Kasztenny, "Synchronized phasor and frequency measurementunder transient conditions," *IEEE Transactions on Power Delivery*, vol. 24, no. 1, pp. 89-95, 2009.

[111] E. Ahad, M. Kezunovic, "Impact of electromechanical wave oscillations propagation on protection schemes," *Electric Power Systems Research*, vol. 138, pp. 85-91, 2016.

[112] NERC Reliability Guideline, "Power Plant Dynamic Model Verification using PMUs," 2016.

[113] S. Julier, J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[114] J. H. Chow, K. W. Cheung, "A toolbox for power system dynamics and control engineering education and research," *IEEE Transactions on Power Systems*, vol. 7, no. 4, pp. 1559-1564, 1992.

[115] P. Gao, M. Wang, et. al, "Missing data recovery by exploiting low-dimensionality in power systems synchrophasor measurements," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1006-1013, 2016.

[116] Y. Chakhchoukh, V. Vittal, G. T. Heydt, "PMU based state estimation by integrating correlation," *IEEE Transactions on Power Systems*, vol. 29, no. 2, pp. 617–626, 2014.

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

[118] S. Prasad, S. Purwar, N. Kishor, "H-infinity based non-linear sliding mode controller for frequency regulation in interconnected power systems with constant and time-varying delays," *IET Generation, Transmission & Distribution*, vol. 10, no. 11, pp. 2771-2784, 2016.

[119] T. Ramachandran, M. H. Nazari, S. Grijalva, M. Egerstedt, "Overcoming communication delays in distributed frequency regulation," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2965-2973, 2016.

[120] J. B. Zhao, L. Mili, "Robust dynamic state estimator to outliers and cyber attacks," *Proceedings of IEEE Power Energy Soceity General Meeting*, 2017, pp. 1-5.

[121] D. Simon, *Optimal State Estimation. Hoboken*, NJ: Wiley, 2006.

[122] K. Zhou, J. C. Doyle, K. Glover, *Robust and Optimal Control*, Prentice-Hall, NewJersey, 1996.

[123] W. Li, Y. Jia, "H-infinity filtering for a class of nonlinear discrete-time systems based on unscented transform," *Signal Processing*, vol. 90, no. 12, pp. 3301-3307, 2010.

[124] G. A. Einicke, L. B. White, "Robust extended Kalman filtering," *IEEE Transactions on Signal Processing*, vol. 47, no. 9, pp. 2596–2599, 1999.

[125] J. B. Zhao, "Dynamic state estimation with model uncertainties using H-infinity extended Kalman filter," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 1099-1100, 2018.

[126] J. B. Zhao, L. Mili, "A decentralized H-infinity unscented Kalman filter for power system dynamic state estimation with uncertainties," *IEEE Transactions on Power Systems*, 2018.

[127] E. Kyriakides, *et. al*, "Online estimation of synchronous generator parameters using a damper current observer and a graphic user interface," *IEEE Transactions on Energy Conversion*, vol. 19, no. 3, pp.499-507, 2004.

[128] J. Chow, *et al*, "Generator and exciter parameter estimation of Fort Patrick Henry hydro unit 1," *IEEE Transactions on Energy Conversion*, vol. 14, no. 4, pp. 923-929, 1999.

[129] D. Kosterev, "Hydro turbine-governor model validation in Pacific Northwest," *IEEE Transactions on Power Systems*, vol. 19, no. 2, pp. 1144-1149, 2004.

[130] S. Akhlaghi, N. Zhou, Z. Huang, "A multi-step adaptive interpolation approach to mitigating the impact of nonlinearity on dynamic state estimation," *IEEE Transations on Smart Grid*, 2016.

[131] S. Benchluch, J. Chow, "A trajectory sensitivity method for the identification of nonlinear excitation system models," *IEEE Transations on Energy Conversion*, vol. 8, no. 2, pp. 159-164, 1993.

[132] I. A. Hiskens, "Nonlinear dynamic model evaluation from disturbance measurements," *IEEE Transactions on Power Systems*, vol. 16, no. 4, pp. 702-710, 2001.

[133] Reliability Guideline of Power Plant Dynamic Model Verification using PMUs, North American Electric Reliability Corporation, 2016.

[134] A. P. S. Meliopoulos, G. J. Cokkinides, *et. al*, "Dynamic state estimation-based protection: status and promise," *IEEE Transactions on Power Delivery*, vol. 32, no. 1, pp. 320-330, 2017.

[135] H. Modir, Robert A. Schlueter, "A dynamic state estimator for dynamic security assessmeng," *IEEE Transactions on Power Systems*, vol. 100, no. 11, pp. 4644–4652, 1981.

[136] J. Chen, R. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems*, Boston, MA: Kluwer Academic Publishers, 2012.

[137] B. Hassibi, Ali H. Sayed, T. Kailath, "Linear estimation in Krein spaces-Part I: Theory," *IEEE Transactions on Automatic Control*, vol. 41, no. 1, pp. 18-33, 1996.

[138] B. Hassibi, Ali H. Sayed, T. Kailath, "Linear estimation in Krein spaces-Part II: Applications," *IEEE Transactions on Automatic Control*, vol. 41, no. 1, pp. 34-49, 1996.

[139] I. Arasaratnam, S. Haykin, Robert J. Elliott, "Discrete-time nonlinear filtering algorithms using GaussCHermite quadrature," *Proceedings of the IEEE*, vol. 95, no. 5, pp. 953-977, 2007.

[140] I. Arasaratnam, S. Haykin, Thomas R. Hurd, "Cubature Kalman filtering for continuous-discrete systems: theory and simulations," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 4977-4993, 2010.

[141] G. Evensen, "The ensemble Kalman filter: theoretical formulation and practical implementation," *Ocean Dynamics*, vol. 53, no. 4, pp. 343C367, 2003.

[142] J. H. Chow, K. W. Cheung, "A toolbox for power system dynamics and control engineering education and research," *IEEE Transactions on Power Systems*, vol. 7, no. 4, pp. 1559-1564, 1992.

[143] S. Benchluch, J. Chow, "A trajectory sensitivity method for the identification of nonlinear excitation system models," *IEEE Transactions on Energy Conversion*, vol. 8, no. 2, pp. 159-164, 1993.

[144] Z. Huang, P. Du, D. Kosterev, S. Yang, "Generator dynamic model validation and parameter calibration using phasor measurements at the point of connection," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1939–1949, 2013.