A Theoretical Framework of Robust H-infinit Unscented Kalman Filter and Its Application to Power System Dynamic State Estimation

Junbo Zhao, *Member, IEEE*, Lamine Mili, *Life Fellow, IEEE*

***Abstract*—This paper presents a new theoretical framework that, by integrating robust statistics and robust control theory, allows us to develop a robust dynamic state estimator of a cyber-physical system. This state estimator combines the generalized maximum-likelihood-type (GM)-estimator, the unscented Kalman filte (UKF), and the H-infinit filte into a robust H-infinit UKF filte in the Krein space, which is able to handle large system uncertainties as well as suppress outliers while achieving a good statistical efficienc under Gaussian and non-Gaussian process and observation noises. Specificall , we firs use the statistical linerization approach to build a linear-like regression model in the Krein space. Then, we show that the H-infinit UKF is just the Krein space Kalman filte that exhibits a bounded estimation error in presence of system uncertainties while minimizing the least squares criterion; consequently, it suffers from a lack of robustness to outliers and non-Gaussian noise. Because the GM-estimator is able to handle outliers, but it may yield large estimation errors in presence of system uncertainties, we propose to combine it with the H-infinit UKF in a robust H-infinit UKF. We carry out a theoretical analysis to demonstrate the connections that our filte has with the H-infinit UKF and the GM-UKF. The good performance of the new filte is demonstrated via extensive simulation performed on the IEEE 39-bus power system.**

***Index Terms*—Dynamic state estimation, robust statistics, model uncertainties, unscented Kalman filte , non-Gaussian noise, H-infinit filte , power system estimation, robustness, phasor measurement units.**

I. Introduction

A. 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 [1]–[4], 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 noises. However, due to unknowable system inputs, including noise, parameter variations and actuator failures [5] and inaccuracies of the model parameter values of the synchronous generators, the loads, the lines, and the transformers, to name a few, the system model is subject to uncertainties. Furthermore, the process and the observation noises of the system nonlinear dynamic models are non-Gaussian as verifie by the Pacifi Northwest National Lab (PNNL) [6], [7]. Finally, outliers can occur that may corrupt the state estimates [8]. 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 [9], [10]. 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 [11], [12]. Innovation outliers may also be induced by the approximations in the state prediction model or by process impulsive noise.

To bound the influenc of the model uncertainties and large system process and measurement noises, the first-orde Taylor series approximations-based H-infinit extended Kalman filte and the unscented transformation-based H-infinit unscented Kalman filte (UKF) have been proposed in [13], [14]. However, they are vulnerable to any types of outliers and exhibit large biases and variances in presence of non-Gaussian noise. To suppress gross errors in the PMU measurements, several robust DSEs are developed. For instance, in [15], 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, a generalized maximum-likelihood-type iterated EKF and UKF are developed in [16]–[18]. However, their statistical efficiencie are low in the presence of non-Gaussian system process and measurement noises. To effectively suppress outliers while achieving a high statistical

This work is supported in part by the U.S. National Science Foundation under Grant ECCS-1711191 and the Advanced Grid Modeling program of the U.S. Department of Energy (DOE) Offic of Electricity Delivery & Energy Reliability.

Junbo Zhao and Lamine Mili are with the Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Falls Church, VA 22043, USA (e-mail: zjunbo@vt.edu, lmili@vt.edu).

efficien y under a broad range of non-Gaussian process and observation noises, the generalized maximum-likelihood-type UKF (GM-UKF) is proposed in [19], [20]. Its robustness and statistical efficien y 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.

B. Contributions and Paper Organization

To address the aforementioned challenges, this paper presents a novel theoretical framework that integrates both robust statistics and robust control theory to develop a robust dynamic state estimator. Specificall , it yields the following contributions:

- The GM-estimator, the UKF, and the H-infinit filte are for the firs time integrated into a unifie framework to yield the robust H-infinit filte for general nonlinear systems in the Krein space.
- The statistical linerization is used to build a linear-like batch-mode regression form in the Krein space. Specificall , we show that the H-infinit UKF is just the Krein space Kalman filte 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 the H-infinit unscented Kalman filte was firstl proposed in [21] following the framework of linear H-infinit filte . This paper provides an alternative way of deriving the H-infinit unscented Kalman filte in the Krein space.
- A robust H-infinit UKF is developed in the Krein space by leveraging the H-infinit criterion to bound system uncertainties with the robustness of GM-estimator to suppress outliers and filte out non-Gaussian noise. It is shown that it is able to handle large system uncertainties as well as suppress outliers while achieving good statistical efficien y under a broad range of non-Gaussian process and observation noises.
- A theoretical analysis is performed to demonstrate that, in the absence of non-Gaussian noise and outliers, the proposed robust H-infinit UKF reduces to the H-infinit UKF [14], [21]; on the other hand, if the tuning parameter of the H-infinit criterion tends to infinit , the proposed robust H-infinit UKF resembles the GM-UKF. This provides the justificatio why both the benefit of the H-infinit filte and the GM-estimator are maintained in the proposed robust H-infinit UKF.
- A comparative analysis is carried out; it shows that our proposed robust H-infinit UKF outperforms the H-infinit UKF and the GM-UKF in terms of statistical efficien y 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.

II. 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:

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

$$\boldsymbol{z}_k = \boldsymbol{h}\left(\boldsymbol{x}_k, \boldsymbol{u}_k\right) + \boldsymbol{v}_k, \tag{2}$$

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 [22]:

Differential equations:

$$T'_{do}\frac{dE'_q}{dt} = -E'_q - \left(X_d - X'_d\right) I_d + E_{fd}, \tag{3}$$

$$T'_{qo}\frac{dE'_d}{dt} = -E'_d - \left(X_q - X'_q\right) I_q, \tag{4}$$

$$\frac{d\delta}{dt} = \omega - \omega_s, \tag{5}$$

$$\frac{2H}{\omega_s}\frac{d\omega}{dt} = T_M - P_e - D\left(\omega - \omega_s\right), \tag{6}$$

$$T_E\frac{dE_{fd}}{dt} = -\left(K_E + S_E\left(E_{fd}\right)\right) E_{fd} + V_R, \tag{7}$$

$$T_{\mathrm{F}}\frac{dV_F}{dt} = -V_F + \frac{K_F}{T_E}V_R - \frac{K_F}{T_E}\left(K_E + S_E\left(E_{fd}\right)\right) E_{fd}, \tag{8}$$

$$T_A\frac{dV_R}{dt} = -V_R + K_A\left(V_{ref} - V_F - V\right), \tag{9}$$

$$T_{CH}\frac{dT_M}{dt} = -T_M + P_{SV}, \tag{10}$$

$$T_{SV}\frac{dP_{SV}}{dt} = -P_{SV} + P_C - \frac{1}{R_D}\left(\frac{\omega}{\omega_s} - 1\right), \tag{11}$$

Algebraic equations:

$$V_d = V\sin\left(\delta - \theta\right), V_q = V\cos\left(\delta - \theta\right), \tag{12}$$

$$I_d = \frac{E'_q - V_q}{X'_d}, I_q = \frac{V_d - E'_d}{X'_q}, \tag{13}$$

$$P_e = V_d I_d + V_q I_q, Q_e = -V_d I_q + V_q I_d, \tag{14}$$

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, fiel voltage, scaled output of the stabilizing transformer and scaled output of the amplifie , 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 a time discretization to (3)-(14) using numerical approaches, such as the 4th-order Ruger-Kutta method, we get (1)-(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 (3)-(11) and (12)-(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 (1), which is a Markov model, and a filtering/updat step using (2). Specificall , 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 (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 filterin 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 filter work well only under the validity of the following assumptions [23], [24]. 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 filte 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 difficul to obtain in practice; the process and observation noise do not obey a Gaussian distribution as verifie in [6], [7]; 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 [13], [25], [26]; and the received measurements may be strongly biased due to impulsive communication noise, cyber attacks, to cite a few [9], [10]. To address these challenges, this paper presents a new theoretical framework that integrates both robust statistics and robust control theory for DSE. The proposed robust H-infinit UKF within this framework is able to handle large system uncertainties and suppress outliers while achieving good statistical efficien y under a broad range of non-Gaussian process and observation noise.

III. Theoretical Framework of the Proposed Robust H-infinity UKF

In this section, the Krein space UKF will be derived first Then its equivalence to the H-infinit 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-infinit UKF, which combines the robustness of the H-infinit criterion to model uncertainties and of the GM-estimator to outliers and non-Gaussian noise.

A. Derivation of the Krein Space UKF

The statistical linearization 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 filte framework for the development of Krein space UKF.

1) Derivation of the Linear-Like Regression Form of the Nonlinear UKF: The main idea of the 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 covariance matrices as the a priori state statistics under the Gaussian assumption [27]. 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 filte structure, i.e., the mean and covariance estimates. Consequently, no calculation of Jacobian matrices is required, which for complex functions can be a difficul task by itself, requiring complicated derivatives if done analytically or being computationally costly if done numerically. Specificall , 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

$$\begin{aligned}\boldsymbol{\chi}^{j}_{k-1|k-1} &= \widehat{\boldsymbol{x}}_{k-1|k-1} + \left(\sqrt{n\boldsymbol{P}^{xx}_{k-1|k-1}}\right)_j,\\ \boldsymbol{\chi}^{j+n}_{k-1|k-1} &= \widehat{\boldsymbol{x}}_{k-1|k-1} - \left(\sqrt{n\boldsymbol{P}^{xx}_{k-1|k-1}}\right)_j,\end{aligned} \tag{15}$$

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

$$\boldsymbol{\chi}^{j}_{k|k-1} = \boldsymbol{f}\left(\boldsymbol{\chi}^{j}_{k-1|k-1}\right). \tag{16}$$

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

$$\widehat{\boldsymbol{x}}_{k|k-1} = \sum_{j=1}^{2n} w_j \boldsymbol{\chi}^{j}_{k|k-1}, \tag{17}$$

$$\boldsymbol{P}^{xx}_{k|k-1} = \sum_{j=1}^{2n} w_j (\boldsymbol{\chi}^{j}_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})(\boldsymbol{\chi}^{j}_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})^T + \boldsymbol{Q}_k, \tag{18}$$

where the weights $w_j = 1/2n$. Next, the sigma points are updated to capture the information of system process noise through

$$\begin{aligned}\boldsymbol{\chi}^{j}_{k|k-1} &= \widehat{\boldsymbol{x}}_{k|k-1} + \left(\sqrt{n\boldsymbol{P}^{xx}_{k|k-1}}\right)_j,\\ \boldsymbol{\chi}^{j+n}_{k|k-1} &= \widehat{\boldsymbol{x}}_{k|k-1} - \left(\sqrt{n\boldsymbol{P}^{xx}_{k|k-1}}\right)_j,\end{aligned} \tag{19}$$

Finally, 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

$$\boldsymbol{P}^{zz}_{k|k-1} = \sum_{j=1}^{2n} w_j (\boldsymbol{z}^j_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1})(\boldsymbol{z}^j_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1})^T + \boldsymbol{R}_k, \tag{20}$$

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 that it propagates the probability distribution of the state vector through nonlinear functions [19], [20], [28]. Thus, by applying a 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

$$\boldsymbol{x}_k = \boldsymbol{F}_k \left(\boldsymbol{x}_{k-1} - \widehat{\boldsymbol{x}}_{k-1|k-1} \right) + \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{e}_k + \boldsymbol{w}_k, \tag{21}$$

$$\boldsymbol{z}_k = \boldsymbol{H}_k \left(\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k-1} \right) + \widehat{\boldsymbol{z}}_{k|k-1} + \boldsymbol{\varepsilon}_k + \boldsymbol{v}_k, \tag{22}$$

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

$$\boldsymbol{P}^{x\chi}_{k|k-1} = \sum_{j=1}^{2n} w_j (\boldsymbol{\chi}^j_{k-1|k-1} - \widehat{\boldsymbol{x}}_{k-1|k-1})(\boldsymbol{\chi}^j_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})^T. \tag{23}$$

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

$$\boldsymbol{P}^{xz}_{k|k-1} = \sum_{i=1}^{2n} w_j (\boldsymbol{\chi}^j_{k|k-1} - \widehat{\boldsymbol{x}}_{k|k-1})(\boldsymbol{z}^j_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1})^T, \tag{24}$$

and where $\boldsymbol{\varepsilon}_k$ is the statistical linearization error term on the nonlinear measurement function; it is normally distributed with zero mean and covariance matrix given by $\boldsymbol{\Pi}_k = \boldsymbol{P}^{zz}_{k|k-1} - (\boldsymbol{P}^{xz}_{k|k-1})^T (\boldsymbol{P}^{xx}_{k|k-1})^{-1} \boldsymbol{P}^{xz}_{k|k-1}$.

Therefore, the nonlinear dynamical model expressed by (1) and (2) is converted into the following equivalent linear-like regression form:

$$\boldsymbol{x}_k = \boldsymbol{F}_k \boldsymbol{x}_{k-1} + \widehat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{F}_k \widehat{\boldsymbol{x}}_{k-1|k-1} + \boldsymbol{e}_k + \boldsymbol{w}_k, \tag{25}$$

$$\boldsymbol{z}_k = \boldsymbol{H}_k \boldsymbol{x}_k + \widehat{\boldsymbol{z}}_{k|k-1} - \boldsymbol{H}_k \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{\varepsilon}_k + \boldsymbol{v}_k, \tag{26}$$

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 filte framework, the recursive state estimation form can be derived as

$$\boldsymbol{K}_k = \boldsymbol{P}^{xz}_{k|k-1} \left(\boldsymbol{P}^{zz}_{k|k-1} \right)^{-1}, \tag{27}$$

$$\widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k \left(\boldsymbol{z}_k - \widehat{\boldsymbol{z}}_{k|k-1} \right), \tag{28}$$

$$\boldsymbol{P}^{xx}_{k|k} = \boldsymbol{P}^{xx}_{k|k-1} - \boldsymbol{K}_k \boldsymbol{P}^{zz}_{k|k-1} \boldsymbol{K}_k^T, \tag{29}$$

which is precisely the standard UKF.

Remark 2: By taking the expectation on both side of (25) and using some mathematical manipulations, we can get the predicted state vector $\widehat{\boldsymbol{x}}_{k|k-1}$ and its covariance matrix $\boldsymbol{P}^{xx}_{k|k-1}$

$$\widehat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{x}_k - \boldsymbol{\delta}_k, \tag{30}$$

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 (30) and (26) simultaneously, we get the following batch-mode regression form:

$$\begin{bmatrix} \boldsymbol{z}_k + \boldsymbol{H}_k \widehat{\boldsymbol{x}}_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1} \\ \widehat{\boldsymbol{x}}_{k|k-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_k \\ \boldsymbol{I} \end{bmatrix} \boldsymbol{x}_k + \begin{bmatrix} \boldsymbol{v}_k + \boldsymbol{\varepsilon}_k \\ -\boldsymbol{\delta}_k \end{bmatrix} \tag{31}$$

which can be rewritten in a compact form as

$$\widetilde{\boldsymbol{z}}_k = \widetilde{\boldsymbol{H}}_k \boldsymbol{x}_k + \widetilde{\boldsymbol{e}}_k. \tag{32}$$

The error covariance matrix is given by

$$\boldsymbol{W}_k = \mathbb{E}\left[\widetilde{\boldsymbol{e}}_k \widetilde{\boldsymbol{e}}_k^T\right] = \begin{bmatrix} \boldsymbol{\Sigma}_{k|k-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{P}^{xx}_{k|k-1} \end{bmatrix}, \tag{33}$$

$\boldsymbol{\Sigma}_{k|k-1} = \mathbb{E}[(\boldsymbol{v}_k + \boldsymbol{\varepsilon}_k)(\boldsymbol{v}_k + \boldsymbol{\varepsilon}_k)^T] = \boldsymbol{R}_k + \boldsymbol{\Pi}_k$. By using the weighted least squares criterion to (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

$$\widehat{\boldsymbol{x}}_{k|k} = \left(\widetilde{\boldsymbol{H}}_k^T \boldsymbol{W}_k^{-1} \widetilde{\boldsymbol{H}}_k \right)^{-1} \widetilde{\boldsymbol{H}}_k^T \boldsymbol{W}_k^{-1} \widetilde{\boldsymbol{z}}_k, \tag{34}$$

with estimation error covariance matrix

$$\boldsymbol{P}^{xx}_{k|k} = \left(\widetilde{\boldsymbol{H}}_k^T \boldsymbol{W}_k^{-1} \widetilde{\boldsymbol{H}}_k \right)^{-1}. \tag{35}$$

By applying the matrix inversion Lemma and performing mathematical manipulations, it can be verifie that the results shown in (34)-(35) are the same as (27)-(29).

2) Derivation of the Krein Space UKF: Before the derivation of the Krein space UKF, let us review the Krein space linear Kalman filte . Following the main results shown in [29], [30], we reorganize the Krein space linear Kalman filte into the following Lemma:

Lemma 1. *For a Krein space discrete-time system,*

$$\begin{cases} \boldsymbol{x}_k = \boldsymbol{A}_k \boldsymbol{x}_{k-1} + \boldsymbol{\eta}_k, \\ \boldsymbol{y}_k = \boldsymbol{C}_k \boldsymbol{x}_k + \boldsymbol{\zeta}_k, \end{cases} \tag{36}$$

with the Gramian matrix given by

$$\left\langle \begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{\eta}_k \\ \boldsymbol{\zeta}_k \end{bmatrix}, \begin{bmatrix} \boldsymbol{x}_0 \\ \boldsymbol{\eta}_k \\ \boldsymbol{\zeta}_k \end{bmatrix} \right\rangle = diag[\ \boldsymbol{P}_{0|0} \quad \boldsymbol{Q}_k \quad \boldsymbol{R}_k\ ], \tag{37}$$

both of which can be obtained from Krein space mapping corresponding to the indefinit quadratic function

$$J = \left\| \boldsymbol{x}_0 - \widehat{\boldsymbol{x}}_{0|0} \right\|^2_{\boldsymbol{P}_{0|0}^{-1}} + \sum_{k=0}^{N-1} \|\boldsymbol{\eta}_k\|^2_{\boldsymbol{Q}_k^{-1}} + \sum_{k=0}^{N} \|\boldsymbol{\zeta}_k\|^2_{\boldsymbol{R}_k^{-1}}. \tag{38}$$

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 filte is provided by

$$(\boldsymbol{P}^{xx}_{k|k})^{-1} = (\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{C}_k^T \boldsymbol{R}_k^{-1} \boldsymbol{C}_k \succ 0, \tag{39}$$

where $\widehat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{A}_k \widehat{\boldsymbol{x}}_{k-1|k-1}$, $\boldsymbol{P}^{xx}_{k|k-1} = \boldsymbol{A}_k \boldsymbol{P}^{xx}_{k-1|k-1} \boldsymbol{A}_k^T + \boldsymbol{Q}_k$ and the state vector is updated by the following equations:

$$\boldsymbol{P}^{xz}_{k|k-1} = \boldsymbol{P}^{xx}_{k|k-1} \boldsymbol{C}_k^T,\ \ \boldsymbol{P}^{zz}_{k|k-1} = \boldsymbol{C}_k^T \boldsymbol{P}^{xx}_{k|k-1} \boldsymbol{C}_k + \boldsymbol{R}_k, \tag{40}$$

$$\boldsymbol{K}_k = \boldsymbol{P}^{xz}_{k|k-1} \left(\boldsymbol{P}^{zz}_{k|k-1} \right)^{-1}, \tag{41}$$

$$\widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k \left(\boldsymbol{y}_k - \boldsymbol{C}_k \widehat{\boldsymbol{x}}_{k|k-1} \right), \tag{42}$$

$$\boldsymbol{P}^{xx}_{k|k} = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\boldsymbol{P}^{xx}_{k|k-1} = \boldsymbol{P}^{xx}_{k|k-1} - \boldsymbol{K}_k\boldsymbol{P}^{zz}_{k|k-1}\boldsymbol{K}_k^T, \tag{43}$$

Remark 3: It can be observed from the above equations that there is little difference between the Hilbert space linear Kalman filte and the Krein space Kalman filte , except for the condition that $(\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{C}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{C}_k \succ 0$. The latter is in fact required by the Hilbert space Kalman filte as well because an estimation error covariance matrix must be positive-definite On the other hand, it is worth pointing out that $(\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{C}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{C}_k$ can be indefinit for the Krein space Kalman filte [29], [30]. 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 defin $\widehat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{A}_k\widehat{\boldsymbol{x}}_{k-1|k-1}$, $\boldsymbol{P}^{xx}_{k|k-1} = \boldsymbol{A}_k\boldsymbol{P}^{xx}_{k-1|k-1}\boldsymbol{A}_k^T + \boldsymbol{Q}_k$ and $\widehat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{x}_k - \boldsymbol{o}_k$, where $\mathbb{E}\left[\boldsymbol{o}_k\boldsymbol{o}_k^T\right] = \boldsymbol{P}^{xx}_{k|k-1}$, (36) can be organized into the following batch-mode regression form:

$$\begin{bmatrix} \boldsymbol{y}_k \\ \widehat{\boldsymbol{x}}_{k|k-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_k \\ \boldsymbol{I} \end{bmatrix}\boldsymbol{x}_k + \begin{bmatrix} \boldsymbol{\zeta}_k \\ -\boldsymbol{o}_k \end{bmatrix}, \tag{44}$$

which can be further rewritten as

$$\widetilde{\boldsymbol{y}}_k = \widetilde{\boldsymbol{C}}_k\boldsymbol{x}_k + \widetilde{\boldsymbol{o}}_k, \tag{45}$$

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 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*

$$(\boldsymbol{P}^{xx}_{k|k})^{-1} = (\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{C}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{C}_k \succ 0, \tag{46}$$

the Krein space linear KF can be derived by applying the weighted least square estimator to (45).

Proof: Following the procedures described in Remark 2, we get

$$\begin{aligned}
\boldsymbol{P}^{xx}_{k|k} &= \left(\widetilde{\boldsymbol{C}}_k^T\widehat{\boldsymbol{W}}_k^{-1}\widetilde{\boldsymbol{C}}_k\right)^{-1} \\
&= \left[\boldsymbol{C}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{C}_k + \left(\boldsymbol{P}^{xx}_{k|k-1}\right)^{-1}\right]^{-1} \\
&= \boldsymbol{P}^{xx}_{k|k-1} - \boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{C}_k^T\left(\boldsymbol{C}_k\boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{C}_k^T + \boldsymbol{R}_k\right)^{-1}\boldsymbol{C}_k\boldsymbol{P}^{xx}_{k|k-1} \\
&= (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\,\boldsymbol{P}^{xx}_{k|k-1} = \boldsymbol{P}^{xx}_{k|k-1} - \boldsymbol{K}_k\boldsymbol{P}^{zz}_{k|k-1}\boldsymbol{K}_k^T,
\end{aligned} \tag{47}$$

where the gain matrix is expressed as

$$\boldsymbol{K}_k = \boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{C}_k^T(\boldsymbol{C}_k\boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{C}_k^T + \boldsymbol{R}_k)^{-1} = \boldsymbol{P}^{xz}_{k|k-1}(\boldsymbol{P}^{zz}_{k|k-1})^{-1}. \tag{48}$$

By following the similar rules and using the matrix inversion Lemma, we can derive the same state vector updating form as (42). ■

Corollary 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*

$$(\boldsymbol{P}^{xx}_{k|k})^{-1} = (\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k \succ 0, \tag{49}$$

the Krein space UKF can be derived as follows:

$$\boldsymbol{K}_k = \boldsymbol{P}^{xz}_{k|k-1}\left(\boldsymbol{P}^{zz}_{k|k-1}\right)^{-1}, \tag{50}$$

$$\widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\left(\boldsymbol{z}_k - \widehat{\boldsymbol{z}}_{k|k-1}\right), \tag{51}$$

$$\boldsymbol{P}^{xx}_{k|k} = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{C}_k)\,\boldsymbol{P}^{xx}_{k|k-1}, \tag{52}$$

where $\boldsymbol{P}^{zz}_{k|k-1}$ and $\boldsymbol{P}^{xz}_{k|k-1}$ are expressed in (20) and (24), respectively.

Proof: The linear-like regression model of the UKF shown in (25)-(26) can be easily organized into the similar one as (36). Thus, by virtue of Lemma 1 and Corollary 1, we get

$$\boldsymbol{P}^{xz}_{k|k-1} = \boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k^T, \;\; \boldsymbol{P}^{zz}_{k|k-1} = \boldsymbol{H}_k^T\boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k + \boldsymbol{R}_k. \tag{53}$$

Then, the gain $\boldsymbol{K}_k$ and state estimates $\widehat{\boldsymbol{x}}_{k|k}$ can be derived as (50) and (51). ■

Remark 4: In fact, the positive-definitenes of $(\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k$ must be satisfie 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-definitenes 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 (49). On the other hand, motivated by the Krein space UKF, the Krein space H-infinit UKF can be derived as well. This is shown in the next section.

B. Derivation of the Krein Space H-infinit UKF

The H-infinit criterion aims to design a filte that achieves the smallest estimation error for all possible disturbances with bounded energy. Specificall , the filte is designed such that the following criterion is satisfied

$$\frac{\sum\limits_{k=0}^{N}\|\boldsymbol{x}_k - \widehat{\boldsymbol{x}}_{k|k}\|^2_{\boldsymbol{P}^{-1}_{k|k}}}{\|\boldsymbol{x}_0 - \widehat{\boldsymbol{x}}_{0|0}\|^2_{\boldsymbol{P}^{-1}_{0|0}} + \sum\limits_{k=0}^{N-1}\|\boldsymbol{w}_k\|^2_{\boldsymbol{Q}_k^{-1}} + \sum\limits_{k=0}^{N}\|\boldsymbol{v}_k\|^2_{\boldsymbol{R}_k^{-1}}} \le \gamma^2, \tag{54}$$

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 1. *If $rank[\boldsymbol{F}_k \;\; \boldsymbol{H}_k] = n$, there exists a filte that achieves the H-infinit criterion shown in (54) if and only if the estimation error covariance matrix $\boldsymbol{P}^{xx}_{k|k}$ for all k satisfie*

$$(\boldsymbol{P}^{xx}_{k|k})^{-1} = (\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k - \gamma^{-2}\boldsymbol{I} \succ 0, \tag{55}$$

and the recursive H-infinit UKF can be expressed as

$$\widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\left(\boldsymbol{z}_k - \widehat{\boldsymbol{z}}_{k|k-1}\right), \tag{56}$$

$$\boldsymbol{K}_k = \boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k^T(\boldsymbol{H}_k\boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k^T + \boldsymbol{R}_k)^{-1}, \tag{57}$$

$$= \boldsymbol{P}^{xz}_{k|k-1}(\boldsymbol{P}^{zz}_{k|k-1})^{-1}, \tag{58}$$

$$\boldsymbol{P}^{xx}_{k|k} = (\boldsymbol{I} - \boldsymbol{P}^{xx}_{k|k-1}[\boldsymbol{H}_k^T \; \boldsymbol{I}]\boldsymbol{R}_{e,k}^{-1}[\boldsymbol{H}_k^T \; \boldsymbol{I}]^T)\boldsymbol{P}^{xx}_{k|k-1}, \tag{59}$$

$$= \boldsymbol{P}^{xx}_{k|k-1} - [\boldsymbol{P}^{xz}_{k|k-1} \; \boldsymbol{P}^{xx}_{k|k-1}]\boldsymbol{R}_{e,k}^{-1}[\boldsymbol{P}^{xz}_{k|k-1} \; \boldsymbol{P}^{xx}_{k|k-1}]^T, \tag{60}$$

$$\boldsymbol{R}_{e,k} = \begin{bmatrix} \boldsymbol{R}_k + \boldsymbol{H}_k\boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k^T & (\boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k^T)^T \\ \boldsymbol{P}^{xx}_{k|k-1}\boldsymbol{H}_k^T & -\gamma^2\boldsymbol{I} + \boldsymbol{P}^{xx}_{k|k-1} \end{bmatrix}; \tag{61}$$

$$= \begin{bmatrix} \boldsymbol{R}_k + \boldsymbol{P}^{zz}_{k|k-1} & [\boldsymbol{P}^{xz}_{k|k-1}]^T \\ \boldsymbol{P}^{xz}_{k|k-1} & -\gamma^2\boldsymbol{I} + \boldsymbol{P}^{xx}_{k|k-1} \end{bmatrix}, \tag{62}$$

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

Proof: See the proof given in Appendix A. ■

It is interesting to note that the H-infinit UKF shares a similar structure as that of the original UKF. One apparent difference is the updating of the estimation error covariance matrix. The latter drives the H-infinit UKF to achieve the smallest estimation error for all possible disturbances. The bounded error performance of the H-infinit filte subject to uncertainties can be proved following the procedures shown in [23]. On the other hand, γ can be seen as a tuning parameter to balance the tradeoff between the H-infinit and the minimum mean-square error performance. Indeed, when γ tends to infinit , the H-infinit UKF will reduce to the traditional UKF. This indicates that the H-infinit norm of the traditional UKF may be quite large, leading to a poor robustness against uncertainties. Furthermore, an indefinit 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-infinit UKF, which is very different from that of the Hilbert space H-infinit Kalman filte . The latter requires the positive-definitenes of a covariance matrix.

Theorem 2. *The H-infinit 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: Defin $\widehat{\boldsymbol{x}}_{k|k} = \boldsymbol{x}_k - \boldsymbol{n}_k$, 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 1 and 2, the predicted state information can be used to construct the batch-mode Krein space UKF instead of the form shown in (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 $\mathbb{E}\left[\boldsymbol{\delta}_k\boldsymbol{\delta}_k^T\right] = \boldsymbol{P}_{k|k-1}^{xx}$. On the other hand, during the proof of Theorem 1, the following Krein space regression form is actually used:

$$\begin{bmatrix} \boldsymbol{z}_k + \boldsymbol{H}_k\widehat{\boldsymbol{x}}_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1} - \boldsymbol{\varepsilon}_k \\ \widehat{\boldsymbol{x}}_{k|k-1} \\ \widehat{\boldsymbol{x}}_{k|k} \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_k \\ \boldsymbol{I} \\ \boldsymbol{I} \end{bmatrix}\boldsymbol{x}_k + \begin{bmatrix} \boldsymbol{v}_k \\ -\boldsymbol{\delta}_k \\ -\boldsymbol{n}_k \end{bmatrix}, \tag{63}$$

which can be rewritten as

$$\widetilde{\boldsymbol{z}}_k = \widetilde{\boldsymbol{H}}_k\boldsymbol{x}_k + \boldsymbol{\varsigma}_k. \tag{64}$$

where $\widetilde{\boldsymbol{W}}_k = \mathbb{E}[\varsigma_k\varsigma_k^T] = diag[\boldsymbol{R}_k \ \boldsymbol{P}_{k|k-1}^{xx} \ -\gamma^2\boldsymbol{I}]$. Applying the weighted least square estimator to (64) yields

$$\begin{aligned}
\widehat{\boldsymbol{x}}_{k|k} &= \mathbb{E}[\left(\widetilde{\boldsymbol{H}}_k^T\widetilde{\boldsymbol{W}}_k^{-1}\widetilde{\boldsymbol{H}}_k\right)^{-1}\widetilde{\boldsymbol{H}}_k^T\widetilde{\boldsymbol{W}}_k^{-1}\widetilde{\boldsymbol{z}}_k] \\
&= ([\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1} \ (\boldsymbol{P}_{k|k-1}^{xx})^{-1} \ -\gamma^2\boldsymbol{I}]\begin{bmatrix} \boldsymbol{H}_k \\ \boldsymbol{I} \\ \boldsymbol{I} \end{bmatrix})^{-1}\widetilde{\boldsymbol{H}}_k^T\boldsymbol{W}_k^{-1}\mathbb{E}[\widetilde{\boldsymbol{z}}_k] \\
&= [\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k + (\boldsymbol{P}_{k|k-1}^{xx})^{-1} - \gamma^2\boldsymbol{I}]^{-1} \\
&\times \begin{bmatrix} \boldsymbol{R}_k^{-T}\boldsymbol{H}_k \\ (\boldsymbol{P}_{k|k-1}^{xx})^{-T} \\ -\gamma^2\boldsymbol{I} \end{bmatrix}^T \begin{bmatrix} \boldsymbol{z}_k + \boldsymbol{H}_k\widehat{\boldsymbol{x}}_{k|k-1} - \widehat{\boldsymbol{z}}_{k|k-1} \\ \widehat{\boldsymbol{x}}_{k|k-1} \\ \widehat{\boldsymbol{x}}_{k|k} \end{bmatrix}.
\end{aligned} \tag{65}$$

Moving the term $\widehat{\boldsymbol{x}}_{k|k}$ from the right-hand side to the left-hand side and performing matrix manipulations, we can finall arrive

$$\widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\left(\boldsymbol{z}_k - \widehat{\boldsymbol{z}}_{k|k-1}\right), \tag{66}$$

$$\boldsymbol{K}_k = \boldsymbol{P}_{k|k-1}^{xz}(\boldsymbol{P}_{k|k-1}^{zz})^{-1}, \tag{67}$$

where $\boldsymbol{P}_{k|k-1}^{xz} = \boldsymbol{P}_{k|k-1}^{xx}\boldsymbol{H}_k^T$ and $\boldsymbol{P}_{k|k-1}^{zz} = \boldsymbol{H}_k^T\boldsymbol{P}_{k|k-1}^{xx}\boldsymbol{H}_k + \boldsymbol{R}_k$. In terms of the estimation error covariance matrix, it is expressed as

$$\begin{aligned}
\boldsymbol{P}_{k|k}^{xx} &= (\widetilde{\boldsymbol{H}}_k^T\widetilde{\boldsymbol{W}}_k^{-1}\widetilde{\boldsymbol{H}}_k)^{-1} \\
&= [\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k + (\boldsymbol{P}_{k|k-1}^{xx})^{-1} - \gamma^2\boldsymbol{I}]^{-1} \\
&= \boldsymbol{P}_{k|k-1}^{xx} - [\boldsymbol{P}_{k|k-1}^{xz} \ \boldsymbol{P}_{k|k-1}^{xx}]\boldsymbol{R}_{e,k}^{-1}[\boldsymbol{P}_{k|k-1}^{xz} \ \boldsymbol{P}_{k|k-1}^{xx}]^T,
\end{aligned} \tag{68}$$

where $\boldsymbol{R}_{e,k}$ is the same as (60). Therefore, it is clear that H-infinit UKF is based on the least square estimator in the Krein space. According to the robust statistics [31], [32], it is well-known that its influenc function is unbounded and therefore lacks any robustness to non-Gaussian noise and outliers. ■

Remark 5: It is worth emphasizing that the proposed nonlinear H-infinit filte is very general. The UKF is just used as an representative example. Other filter that share the similar structure of UKF can be leveraged to derive their corresponding H-infinit filter in the Krein space, such as the divided difference filte , the quadrature Kalman filte , the cubature Kalman filte , the EnKF, to cite a few [33]–[35]. On the other hand, instead of propagating the statistics of the state vector using sigma points, the first-orde Taylor series expansion-based methods can be used, yielding the H-infinit EKF.

C. Derivation of the Robust H-infinit UKF

It is shown in Theorem 2 that the H-infinit filte is actually based on the Krein space weighted least squares estimator. Despite of its bounded performance against system uncertainties, it lacks any 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 filte out thick-tailed non-Gaussian measurement noise while bounding system uncertainties, we propose to apply the GM-estimator to (64) instead of using the weighted least squares estimator, yielding the following objective function:

$$J(\boldsymbol{x}_k) = \sum_{i=1}^{m+2n}\varpi_i^2\rho\left(r_{S_i}\right), \tag{69}$$

where ϖ_i are the weights to downweight outliers calculated by applying projection statistics [17], [18] to the innovation matrix $\boldsymbol{Z}_k$. The latter is define as follows:

$$\boldsymbol{Z}_k = \begin{bmatrix} \boldsymbol{z}_{k-1} - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k-1|k-2}) & \boldsymbol{z}_k - \boldsymbol{h}(\widehat{\boldsymbol{x}}_{k|k-1}) \\ \widehat{\boldsymbol{x}}_{k-1|k-2} & \widehat{\boldsymbol{x}}_{k|k-1} \\ \widehat{\boldsymbol{x}}_{k-2|k-2} & \widehat{\boldsymbol{x}}_{k-1|k-1} \end{bmatrix}, \tag{70}$$

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 filtere state vectors at time instants k-2 and

k-1, respectively; the mathematical expression of projection statistics is define as [17], [18]:

$$PS_j = \max_{\|\boldsymbol{\ell}\|=1} \frac{\left| \boldsymbol{l}_j^T \boldsymbol{\ell} - med_i \left(\boldsymbol{l}_i^T \boldsymbol{\ell} \right) \right|}{1.4826 \; med_\kappa \left| \boldsymbol{l}_\kappa^T \boldsymbol{\ell} - med_i \left(\boldsymbol{l}_i^T \boldsymbol{\ell} \right) \right|}, \tag{71}$$

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 identifie as outliers and downweighted through

$$\varpi_i = \min\left(1, d^2/PS_i^2\right), \tag{72}$$

where the parameter d is set equal to 1.5 to yield good statistical efficien y at different distributions.

In (69), $r_{S_i} = r_i/s\varpi_i$ 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 finit sample of size $m + 2n$ at a given probability distribution; $\rho(\cdot)$ is the nonlinear function of r_{S_i} expressed as

$$\rho\left(r_{S_i}\right) = \begin{cases} \frac{1}{2} r_{S_i}^2, & \text{for } |r_{S_i}| < \lambda \\ \lambda \left|r_{S_i}\right| - \lambda^2/2, & elsewhere \end{cases}, \tag{73}$$

where the parameter λ is typically chosen between 1.5 to 3 in the literature.

To minimize (69), we take its partial derivative with respect to $\boldsymbol{x}_k$ and set it equal to zero, yielding

$$\frac{\partial J\left(\boldsymbol{x}_k\right)}{\partial \boldsymbol{x}_k} = \sum_{i=1}^{m+2n} - \frac{\varpi_i \widetilde{\boldsymbol{h}}_i}{s \sigma_i} \psi\left(r_{S_i}\right) = \boldsymbol{0}, \tag{74}$$

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 (74) and putting it in a matrix form, we get

$$\widetilde{\boldsymbol{H}}_k^T \widetilde{\boldsymbol{W}}_k^{-1} \widehat{\boldsymbol{Q}} \left(\widetilde{\boldsymbol{z}}_k - \widetilde{\boldsymbol{H}}_k \boldsymbol{x}_k \right) = \boldsymbol{0}, \tag{75}$$

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 [17], [31], the state vector correction at the j iteration is calculated through

$$\widehat{\boldsymbol{x}}_{k|k}^{(j+1)} = \left(\widetilde{\boldsymbol{H}}_k^T \widetilde{\boldsymbol{W}}_k^{-1} \widehat{\boldsymbol{Q}}^{(j)} \widetilde{\boldsymbol{H}}_k \right)^{-1} \widetilde{\boldsymbol{H}}_k^T \widetilde{\boldsymbol{W}}_k^{-1} \widehat{\boldsymbol{Q}}^{(j)} \widetilde{\boldsymbol{z}}_k^j, \tag{76}$$

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 2, we move $\widehat{\boldsymbol{x}}_{k|k}$ from the right-hand side of (76) to its left-hand side, yielding

$$\widehat{\boldsymbol{x}}_{k|k}^{(j+1)} = \left(\boldsymbol{\Gamma}_k^T \boldsymbol{\Lambda}_k^{-1} \boldsymbol{\Xi}^{(j)} \boldsymbol{\Gamma}_k \right)^{-1} \boldsymbol{\Gamma}_k^T \boldsymbol{\Lambda}_k^{-1} \boldsymbol{\Xi}^{(j)} \boldsymbol{\xi}_k, \tag{77}$$

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, (77) 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 efficien y. Note that the algorithm converges when $\left\| \Delta \widehat{\boldsymbol{x}}_{k|k}^{(j+1)} \right\|_\infty \leq 10^{-2}$.

To derive the asymptotic error covariance matrix of the robust H-infinit UKF at time sample k, we use the influenc function of the GM-estimator. According to our previous work [17], the influenc function of the robust H-infinit UKF is given by

$$\boldsymbol{IF}(\boldsymbol{x}; \Phi, \boldsymbol{T}) = [\int \frac{1}{s} \psi^{'}(r_{S_i}) \widetilde{\boldsymbol{H}}_k \widetilde{\boldsymbol{H}}_k^T |_{T(\Phi)} d\Phi]^{-1} \varpi \widetilde{\boldsymbol{H}}_k \psi(r_{S_i}) \tag{78}$$

where Φ is the cpdf of the standardized residual r_{S_i}; $T(\cdot)$ is the functional form of the robust H-infinit UKF. Finally, the asymptotic error covariance matrix is updated through

$$\begin{aligned} \boldsymbol{P}_{k|k}^{xx} &= \mathbb{E}[\boldsymbol{IF}(\boldsymbol{x}; \Phi, \boldsymbol{T}) \cdot \boldsymbol{IF}(\boldsymbol{x}; \Phi, \boldsymbol{T})^T] \\ &= \alpha (\widetilde{\boldsymbol{H}}_k^T \widetilde{\boldsymbol{W}}_k^{-1} \widetilde{\boldsymbol{H}}_k)^{-1} (\widetilde{\boldsymbol{H}}_k^T \widehat{\boldsymbol{Q}}_\varpi \widetilde{\boldsymbol{W}}_k^{-1} \widetilde{\boldsymbol{H}}_k) (\widetilde{\boldsymbol{H}}_k^T \widetilde{\boldsymbol{W}}_k^{-1} \widetilde{\boldsymbol{H}}_k)^{-1} \end{aligned} \tag{79}$$

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: Under the assumption of Gaussian noise and the absence of outliers, no predicted state and measurement are downweighted. Therefore, $\widehat{\boldsymbol{Q}} = \boldsymbol{I}$ and

$$\widehat{\boldsymbol{x}}_{k|k} = \left(\widetilde{\boldsymbol{H}}_k^T \widetilde{\boldsymbol{W}}_k^{-1} \widetilde{\boldsymbol{H}}_k \right)^{-1} \widetilde{\boldsymbol{H}}_k^T \widetilde{\boldsymbol{W}}_k^{-1} \widetilde{\boldsymbol{z}}_k, \tag{80}$$

$$\boldsymbol{P}_{k|k}^{xx} = \frac{\mathbb{E}_\Phi[\psi^2(r_{S_i})]}{\{\mathbb{E}_\Phi[\psi'(r_{S_i})]\}^2} (\widetilde{\boldsymbol{H}}_k^T \widetilde{\boldsymbol{W}}_k^{-1} \widetilde{\boldsymbol{H}}_k)^{-1}, \tag{81}$$

where α is very close to 1. For example, under Gaussian noise, this value can be calculated as 1.0369. As a result, (80) and (81) are the same of (65) and (68), respectively. Thus, the robust H-infinit UKF reduces to the H-infinit UKF. According to Theorem 1, the uncertainties of the system model and measurement noises are bounded, yielding bounded state estimation errors. This provides the theoretical justification of why the proposed method is robust to uncertainties. On the other hand, according to equations (60) and (62), it can be shown that when γ tends to infinit , the H-infinit criterion reduces to the traditional UKF. Then, the robust H-infinit UKF reduces to the GM-UKF, which is robust to outliers and non-Gaussian noise [20]. In fact, the robust H-infinit UKF has a bounded influenc function, see equation (78). This justifie its robustness to different types of outliers and non-Gaussian noise. From these theoretical analysis, we can conclude that the proposed robust H-infinit UKF leverages the robustness of the GM-estimator to filte out non-Gaussian noise and suppress outliers while relying on the H-infinit criterion to bound system uncertainties.

D. Computational Complexity Analysis

For the traditional UKF algorithm, it has the computational complexity of order $\mathcal{O}(n^3)$ [36]. In particular, there are two operations that lead to the cubic complexity, namely the Cholesky factorization of covariance matrix (see equations (15) and (19)) and the outer product calculations (see equations (20) and (23)). For the traditional linear Kalman filte , the computational complexity is $\mathcal{O}(n^3)$ [37]. In the proposed framework, we have found that the H-infinit UKF can be shown to be a linear Kaman filte using the statistical linearization in the Krein space. Since the Cholesky factorization of covariance matrix and the outer product calculations for the covariance matrix are still required, the computational complexity is $\mathcal{O}(n^3)$ for the state prediction. In the filterin stage, the most time consuming operation is the covariance matrix updating that involves matrix multiplications and inverse. Note that due to the introduction of tuning parameter gamma as well as adding an matrix associated with it, the computational complexity increases to the order $\mathcal{O}((2n)^3)$ in the filterin stage. As a result, the H-infinit UKF has the computational complexity of $\mathcal{O}((2n)^3)$. For the proposed robust H-infinit UKF, the computational complexity is the same as H-infinit UKF at the state prediction stage. However, during the state filterin stage, as the dimensions of the $\tilde{\boldsymbol{H}}$ increase to $3n$, the computational complexity becomes $\mathcal{O}((3n)^3)$. Finally, the proposed robust H-infinit UKF has an approximate computational complexity of order $\mathcal{O}((3n)^3)$. Therefore, it can be concluded that the computational efficiencie of the UKF, the H-infinit UKF and the robust H-infinit UKF are in the similar orders. This will be validated by the simulation results shown in Section IV-D.

E. 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 [17], [18], [25]. 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 in practical power system. This motivates the development of a decentralized DSE that is implemented at each local synchronous generator. The proposed robust H-infinit UKF can be implemented in both ways. However, only the decentralized one is used for demonstration in this paper. 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.

IV. Numerical Results

Extensive simulations are carried out on the IEEE 39-bus system to assess the performance of the proposed robust H-infinit 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 [38]. The transient stability time domain simulations are performed to generate measurements and true state variables using the

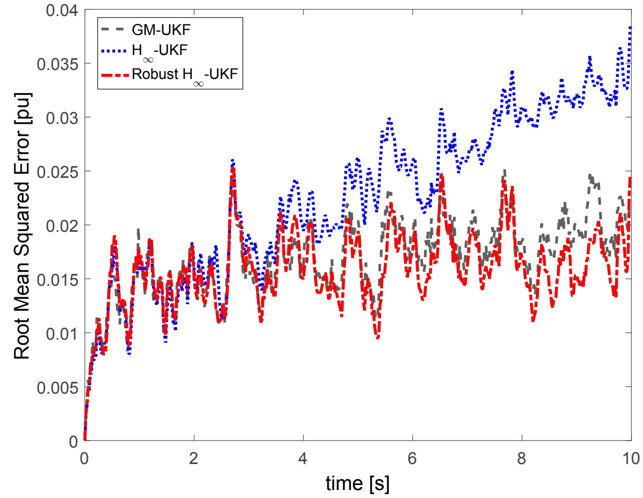

Fig. 1: Root-mean-squared errors of the GM-UKF, the H-infinit UKF and the robust H-infinit UKF in presence of unknown non-Gaussian system process and measurement noise.

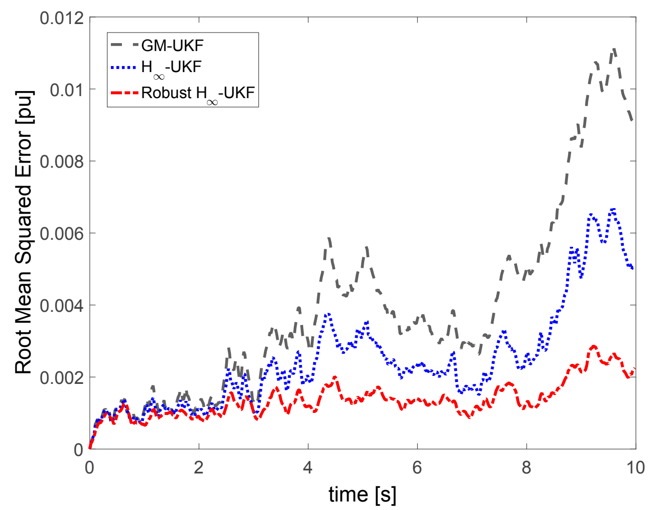

Fig. 2: Root-mean-squared error of the GM-UKF, the H-infinit UKF and the robust H-infinit UKF with model uncertainties. Here, it is assumed that after the disturbance is applied, the transient reactances of Generator 5 deviate from the nominal values by a percentage of 10%.

Matlab-based software PST [39] 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-infinit filte is 10; the root-mean-squared

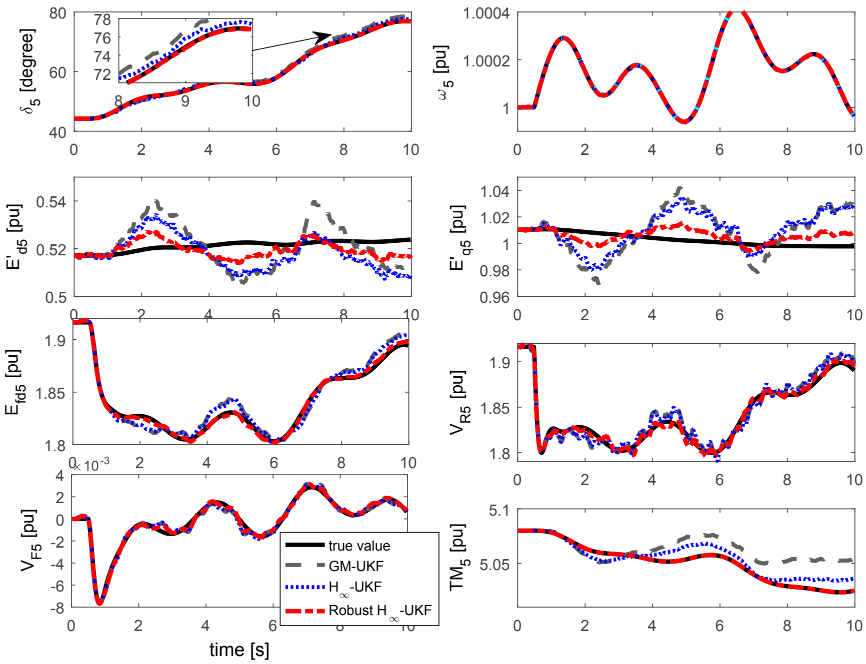

Fig. 3: Estimated state variables by the GM-UKF, the H-infinit UKF and the robust H-infinit UKF with model uncertainties.

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-infinit UKF proposed in this paper and the GM-UKF [19] are used for comparison.

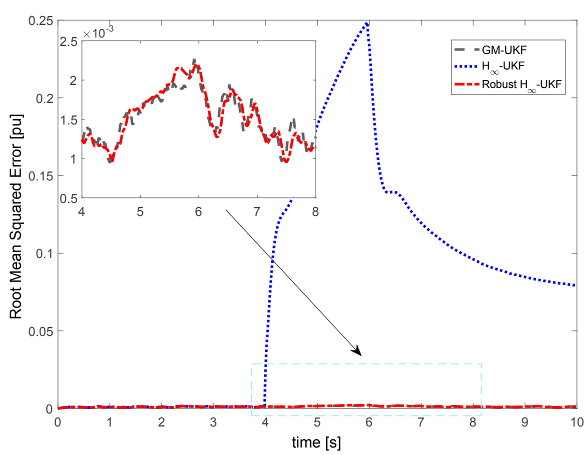

Fig. 4: Root-mean-squared error of the GM-UKF, the H-infinit UKF and the robust H-infinit UKF in presence of observation outliers. Here, it is assumed that the real and reactive power measurements of Generator 5 are corrupted with 20% errors from t=4s to t=6s.

A. 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 approximate any distribution, it is assumed in this paper. Specificall , we assume they follow a Gaussian mixture model, where 10% of the data are drawn with covariance matrices $\boldsymbol{Q} = 5 \times 10^{-5}\boldsymbol{I}_{9\times 9}$ and $\boldsymbol{R} = 5 \times 10^{-5}\boldsymbol{I}_{2\times 2}$ while the true covariance matrices are $\boldsymbol{Q} = 10^{-6}\boldsymbol{I}_{9\times 9}$ and $\boldsymbol{R} = 10^{-6}\boldsymbol{I}_{2\times 2}$. The root-mean-squared errors of each method are displayed in Fig. 1. It can be observed that the H-infinit 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 efficien y. By contrast, both the GM-UKF and the proposed robust H-infinit UKF are able to withstand the non-Gaussian noise thanks to robustness of GM-estimator. However, the proposed robust H-infinit UKF outperforms the GM-UKF slightly. This is because the non-Gaussian noise is actually unknown and the proposed robust H-infinit UKF firs leverages the robustness of GM-estimator to suppress it and then relies on the H-infinit criterion to further bound these uncertainties, yielding improved statistical efficien y.

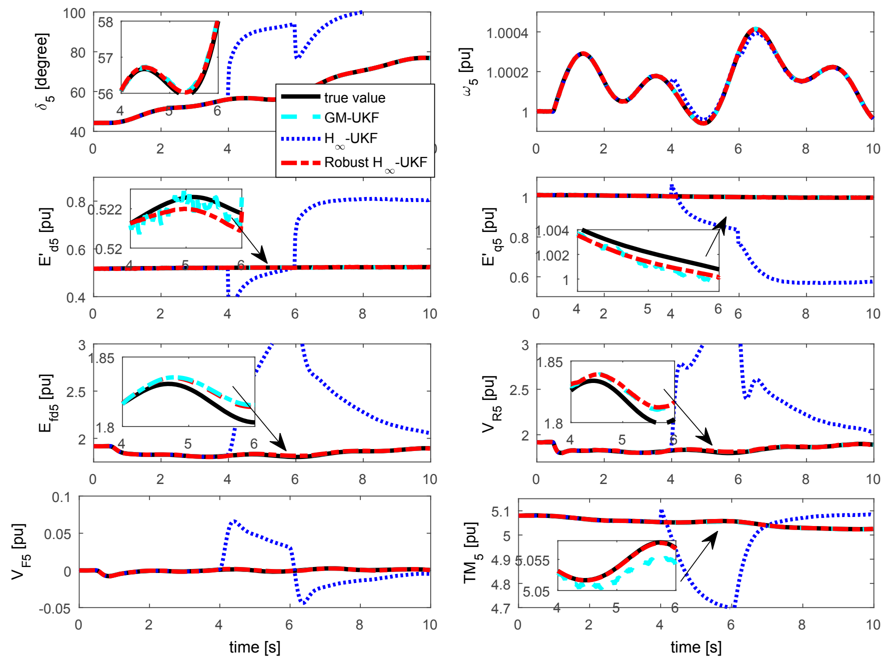

Fig. 5: Estimated state variables of the GM-UKF, the H-infinit UKF and the robust H-infinit UKF in the presence of observation outliers from t=4s to t=6s.

B. 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 significantl . In this paper, 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 [40]–[42], 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 significantl while the changes on other parameters induce negligible difference. By using the trajectory sensitivity analysis approach [40], [41], the transient reactances and the gain of excitor are identifie as these critical parameters. We anticipate that the uncertainties of them would impose huge challenges to the dynamic state estimator.

Figs. 2-3 display the root-mean-squared errors of the GM-UKF, the H-infinit UKF and the robust H-infinit 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-infinit criterion, the H-infinit 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 modifie the estimation error covariance matrix at each iteration to achieve some robustness to system uncertainties, which is similar to the H-infinit UKF. This justifie that why the H-infinit UKF does not outperform significantl the GM-UKF. The proposed robust H-infinit UKF leverages both the robustness of estimation error covariance matrix and the H-infinit 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-infinit filte that can reliably always track the system dynamic states.

C. 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 [15], [17]. 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. 4 and Fig. 5, respectively. It can be found that the Krein space weighted least squares estimator-based H-infinit lacks of robustness to outliers, yielding significantl biased state estimates. By contrast, thanks to the weights provided by projection statistics

TABLE I: Average Computing Times of the GM-UKF, the H-infinit UKF and the Robust H-infinit UKF at Each PMU Scan

Cases	GM-UKF	H-infinit UKF	Robust H-infinit 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

and the GM-estimator, both the GM-UKF and the robust H-infinit 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-infinit UKF achieves the desired performance, that is, leveraging the H-infinit criterion to bound system uncertainties while relying on the robustness of GM-estimator to filte out non-Gaussian noise and suppress outliers.

D. Computational Efficienc

To assess the computational efficien y of each method, the average computing times of the GM-UKF, the H-infinit UKF and the robust H-infinit UKF for Cases 1-3 at each PMU scan are presented in Table. I. 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-infinit UKF is the most computational efficien approach, followed by the GM-UKF. Although the robust H-infinit 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.

V. Conclusions

In this paper, the GM-estimator, the unscented Kalman filte (UKF), and the H-infinit filte are integrated into a unifie framework to yield the general robust H-infinit UKF. The latter is able to handle large system uncertainties as well as suppress outliers while achieving good statistical efficien y under a broad range of non-Gaussian process and observation noise. Specificall , it leverages the H-infinit criterion to bound system uncertainties while relying on the robustness of GM-estimator to filte out non-Gaussian noise and suppress outliers. By contrast, the H-infinit UKF is shown to be based on the Krein space least squares estimator and thus, lacks robustness to outliers and non-Gaussian noise. Comparative results reveal that our proposed robust H-infinit UKF outperforms the H-infinit UKF and the GM-UKF in terms of statistical efficien y and robustness to outliers, non-Gaussian noise, and model uncertainties. As a future work, we will assess the breakdown point of the proposed robust H-infinit UKF. Furthermore, we will investigate other combinations of robust estimators and filter to obtain a good balance between estimation efficien y and robustness to outliers and model uncertainties. In particular, we will improve the statistical efficien y of the filte by resorting to other robust control strategies such as sliding mode control and probabilistic robust control.

Appendix A
Proof of Theorem 1

Proof: The firs step is to convert the suboptimal H-infinit filterin problem (54) to an indefinit form so that the Krein space Kalman filte can be used. Formally, we get

$$
\begin{aligned}
J_\infty &= \left\|\boldsymbol{x}_0-\widehat{\boldsymbol{x}}_{0|0}\right\|^2_{\boldsymbol{P}^{-1}_{0|0}}+\sum_{k=0}^{N-1}\|\boldsymbol{w}_k\|^2_{\boldsymbol{Q}_k^{-1}}+\sum_{k=0}^{N}\|\boldsymbol{v}_k\|^2_{\boldsymbol{R}_k^{-1}}\\
&\quad-\gamma^{-2}\sum_{k=0}^{N}\|\boldsymbol{x}_k-\widehat{\boldsymbol{x}}_{k|k}\|^2_{\boldsymbol{P}^{-1}_{k|k}}\\
&=\|\boldsymbol{x}_0-\widehat{\boldsymbol{x}}_{0|0}\|^2_{\boldsymbol{P}^{-1}_{0|0}}+\sum_{k=0}^{N-1}\|\boldsymbol{w}_k\|^2_{\boldsymbol{Q}_k^{-1}}\\
&\quad+\sum_{k=0}^{N}\begin{bmatrix}\boldsymbol{z}_k-\boldsymbol{h}(\boldsymbol{x}_k)\\ \boldsymbol{x}_k-\widehat{\boldsymbol{x}}_{k|k}\end{bmatrix}^T\begin{bmatrix}\boldsymbol{R}_k & \boldsymbol{0}\\ \boldsymbol{0} & -\gamma^2\boldsymbol{I}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{z}_k-\boldsymbol{h}(\boldsymbol{x}_k)\\ \boldsymbol{x}_k-\widehat{\boldsymbol{x}}_{k|k}\end{bmatrix}
\end{aligned}
\tag{82}
$$

Defin $\boldsymbol{c}_k=\widehat{\boldsymbol{x}}_{k|k-1}-\boldsymbol{F}_k\widehat{\boldsymbol{x}}_{k-1|k-1}+\boldsymbol{e}_k$, $\boldsymbol{m}_k=\widehat{\boldsymbol{z}}_{k|k-1}-\boldsymbol{H}_k\widehat{\boldsymbol{x}}_{k|k-1}+\boldsymbol{\varepsilon}_k$, $\boldsymbol{y}_k=\boldsymbol{z}_k-\boldsymbol{m}_k$, $\widetilde{\boldsymbol{C}}_k^T=[\boldsymbol{H}_k^T\ \boldsymbol{I}]$, $\widetilde{\boldsymbol{m}}_k=[\boldsymbol{y}_k^T\ \widehat{\boldsymbol{x}}_{k|k}^T]^T$ and use the statistical linerization results shown in (25)-(26), (82) can be rewritten as

$$
\begin{aligned}
J_\infty &= \|\boldsymbol{x}_0-\widehat{\boldsymbol{x}}_{0|0}\|^2_{\boldsymbol{P}^{-1}_{0|0}}+\sum_{k=0}^{N-1}\|\boldsymbol{w}_k\|^2_{\boldsymbol{Q}_k^{-1}}\\
&\quad+\sum_{k=0}^{N}\begin{bmatrix}\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{x}_k\\ \boldsymbol{x}_k-\widehat{\boldsymbol{x}}_{k|k}\end{bmatrix}^T\begin{bmatrix}\boldsymbol{R}_k & \boldsymbol{0}\\ \boldsymbol{0} & -\gamma^2\boldsymbol{I}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{y}_k-\boldsymbol{H}_k\boldsymbol{x}_k\\ \boldsymbol{x}_k-\widehat{\boldsymbol{x}}_{k|k}\end{bmatrix}
\end{aligned}
\tag{83}
$$

which resembles the Krein space Kalman filte form (36). Therefore, according to Lemma 1 and Corollary 1, the estimation error covariance matrix is derived as

$$
\begin{aligned}
(\boldsymbol{P}^{xx}_{k|k})^{-1} &= (\boldsymbol{P}^{xx}_{k|k-1})^{-1}+\widetilde{\boldsymbol{C}}_k\boldsymbol{P}^{xx}_{k|k-1}\widetilde{\boldsymbol{C}}_k^T\\
&=(\boldsymbol{P}^{xx}_{k|k-1})^{-1}+\boldsymbol{H}_k^T\boldsymbol{R}_k^{-1}\boldsymbol{H}_k-\gamma^{-2}\boldsymbol{I}.
\end{aligned}
\tag{84}
$$

It is clear that $(\boldsymbol{P}^{xx}_{k|k})^{-1}\succ 0$ must hold for true for the existence of H-infinit UKF. This is because the positive-definitenes of $\boldsymbol{P}^{xx}_{k|k}$ is needed for the generation of sigma points at the next time instant. By applying the matrix inversion Lemma to (84) and using tedious algebraic manipulations, we are able to derive

$$
\begin{aligned}
\boldsymbol{P}^{xx}_{k|k} &= (\boldsymbol{I}-\boldsymbol{P}^{xx}_{k|k-1}[\boldsymbol{H}_k^T\ \boldsymbol{I}]\boldsymbol{R}_{e,k}^{-1}[\boldsymbol{H}_k^T\ \boldsymbol{I}]^T)\boldsymbol{P}^{xx}_{k|k-1}\\
&=\boldsymbol{P}^{xx}_{k|k-1}-[\boldsymbol{P}^{xz}_{k|k-1}\ \boldsymbol{P}^{xx}_{k|k-1}]\boldsymbol{R}_{e,k}^{-1}[\boldsymbol{P}^{xz}_{k|k-1}\ \boldsymbol{P}^{xx}_{k|k-1}]^T,
\end{aligned}
\tag{85}
$$

where the expression of $\boldsymbol{R}_{e,k}$ is the same as (62).

In the meantime, according to the formula of the Krein space Kalman filte , the filtere state vector is derived as

$$\begin{aligned}
\widehat{\boldsymbol{x}}_{k|k} &= \widehat{\boldsymbol{x}}_{k|k-1} \\
&\quad + \boldsymbol{P}^{xx}_{k|k-1} \widetilde{\boldsymbol{C}}_k^T (\widetilde{\boldsymbol{C}}_k \boldsymbol{P}^{xx}_{k|k-1} \widetilde{\boldsymbol{C}}_k^T + \boldsymbol{R}_k)^{-1} (\widetilde{\boldsymbol{m}}_k - \widetilde{\boldsymbol{C}}_k \widehat{\boldsymbol{x}}_{k|k-1}) \\
&= \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{P}^{xx}_{k|k-1} [\boldsymbol{H}_k^T \ \boldsymbol{I}] \begin{bmatrix} \boldsymbol{I} & -\widehat{\boldsymbol{R}}_k \boldsymbol{H}_k \boldsymbol{P}^{xx}_{k|k-1} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \\
&\quad \times \begin{bmatrix} \widehat{\boldsymbol{R}}_k & \boldsymbol{0} \\ \boldsymbol{0} & -\gamma^2 \boldsymbol{I} + ((\boldsymbol{P}^{xx}_{k|k-1})^{-1} + \boldsymbol{H}_k^T \boldsymbol{H}_k)^{-1} \end{bmatrix}^{-1} \\
&\quad \times \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ -\boldsymbol{P}^{xx}_{k|k-1} \boldsymbol{H}_k^T \widehat{\boldsymbol{R}}_k^{-1} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{z}_k - \widehat{\boldsymbol{z}}_{k|k-1} \\ \widehat{\boldsymbol{x}}_{k|k} - \widehat{\boldsymbol{x}}_{k|k-1} \end{bmatrix}^T,
\end{aligned} \tag{86}$$

where $\widehat{\boldsymbol{R}}_k = \boldsymbol{R}_k + \boldsymbol{H}_k \boldsymbol{P}^{xx}_{k|k-1} \boldsymbol{H}_k^T = \boldsymbol{P}^{zz}_{k|k-1}$. Using the matrix inversion Lemma and mathematical manipulations, we can finall get

$$\widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k \left(\boldsymbol{z}_k - \widehat{\boldsymbol{z}}_{k|k-1} \right), \tag{87}$$

$$\boldsymbol{K}_k = \boldsymbol{P}^{xz}_{k|k-1} (\boldsymbol{P}^{zz}_{k|k-1})^{-1}. \tag{88}$$

■

REFERENCES

[1] J. B. Zhao, A. Exposito, M. Netto, L. Mili, A. Abur, V. Terzija, I. Kamwa, B. Pal, A. K. Singh, J. Qi, Z. Huang, A. P. Sakis Meliopoulos, "Power system dynamic state estimation: motivations, definitions methodologies and future work," *IEEE Trans. Power Syst.*, 2019.

[2] H. Modir, Robert A. Schlueter, "A dynamic state estimator for dynamic security assessmeng," *IEEE Trans. Power Syst.*, vol. 100, no. 11, pp. 4644–4652, 1981.

[3] A. K. Singh, B. C Pal, "Decentralized control of oscillatory dynamics in power systems using an extended LQR," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 1715-1728, 2016.

[4] A. P. S. Meliopoulos, G. J. Cokkinides, *et. al*, "Dynamic state estimation-based protection: status and promise," *IEEE Trans. Power Delivery*, vol. 32, no. 1, pp. 320-330, 2017.

[5] J. Chen, R. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems*, Boston, MA: Kluwer Academic Publishers, 2012.

[6] Z. Huang, N. Zhou, R. Diao, S. Wang, S. Elbert, D. Meng, S. Lu, "Capturing real-time power system dynamics: opportunities and challenges," *Proc. IEEE Power Eng. Soc. General Meeting*, July 2015.

[7] S. Wang, J. B. Zhao, Z. Huang, R. Diao, "Assessing Gaussian assumption of PMU measurement error using fiel data," *IEEE Trans. Power Delivery*, vol. 33, no. 6, pp. 3233-3236, 2018.

[8] M. Gandhi, L. Mili, "Robust Kalman filte based on a generalized maximum-likelihood-type estimator, *IEEE Trans. Signal Processing*, vol. 58, no. 5, pp. 2509–2520, 2010.

[9] K. D. Jones, A. Pal, J. S. Thorp, "Methodology for performing synchrophasor data conditioning and validation," *IEEE Trans. Power Syst.*, vol. 30, no. 3, pp. 1121–1130, 2015.

[10] J. B. Zhao, S. Wang, L. Mili, R. Huang, Z. Huang, "A robust state estimation framework considering measurement correlations and imperfect synchronization," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4604-4613, 2018.

[11] D. N. Kosterev, C. W. Taylor, W. A. Mittelstadt, "Model validation for the August 10, 1996 WSCC system outage," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 967–979, 1999.

[12] Z. Huang, P. Du, D. Kosterev, S. Yang, "Generator dynamic model validation and parameter calibration using phasor measurements at the point of connection," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1939–1949, 2013.

[13] J. B. Zhao, "Dynamic state estimation with model uncertainties using H-infinit extended Kalman filte ," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 1099-1100, 2018.

[14] J. B. Zhao, L. Mili, "A decentralized H-infinit unscented Kalman filte for dynamic state estimation against uncertainties," *IEEE Trans. Smart Grid*, 2018.

[15] A. Rouhani, A. Abur, "Linear phasor estimator assisted dynamic state estimation," *IEEE Trans. Smart Grid*, vol. 9, no. 1, pp. 211-219, 2018.

[16] J. B. Zhao, L. Mili, "Robust dynamic state estimator to outliers and cyber attacks," *Proc. IEEE Power Energy General Meeting*, 2017, pp. 1-5.

[17] J. B. Zhao, M. Netto, L. Mili, "A robust iterated extended Kalman filte for power system dynamic state estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 2484-2486, 2017.

[18] J. B. Zhao, L. Mili, "Power system robust decentralized dynamic state estimation based on multiple hypothesis testing," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4553-4562, 2018.

[19] J. B. Zhao, L. Mili, "Robust unscented Kalman filte for power system dynamic state estimation with unknown noise statistics," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1215-1224, 2019.

[20] J. B. Zhao, L. Mili, "A robust generalized-maximum likelihood unscented Kalman filte for power system dynamic state estimation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 4, pp. 578-592, 2018.

[21] W. Li, Y. Jia, "H-infinit filterin for a class of nonlinear discrete-time systems based on unscented transform," *Signal Processing*, vol. 90, no. 12, pp. 3301–3307, 2010.

[22] P. Sauer, M. A. Pai. *Power system dynamics and stability*. Urbana, 1997.

[23] D. Simon, *Optimal State Estimation. Hoboken*, NJ: Wiley, 2006.

[24] K. Zhou, J. C. Doyle, K. Glover, *Robust and Optimal Control*, Prentice-Hall, NewJersey, 1996.

[25] E. Ghahremani, I. Kamwa, "Dynamic state estimation in power system by applying the extended Kalman filte with unknown inputs to phasor measurements," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2556–2566, 2011.

[26] G. Anagnostou, B. C. Pal, "Derivative-free Kalman filterin based approaches to dynamic state estimation for power systems with unknown inputs," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 116-130, 2018.

[27] S. Julier, J. K. Uhlmann, "Unscented filterin and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[28] T. Lefebvre, H. Bruyninckx, and J. De Schutter, "Comment on "a new method for the nonlinear transformation of means and covariances in filter and estimators"," *IEEE Trans. Autom. Control*, vol. 47, no. 8, pp. 1406–1408, Aug. 2002.

[29] B. Hassibi, Ali H. Sayed, T. Kailath, "Linear estimation in Krein spaces-Part I: Theory," *IEEE Trans. Automatic Control*, vol. 41, no. 1, pp. 18-33, 1996.

[30] B. Hassibi, Ali H. Sayed, T. Kailath, "Linear estimation in Krein spaces-Part II: Applications," *IEEE Trans. Automatic Control*, vol. 41, no. 1, pp. 34-49, 1996.

[31] P. J. Huber, *Robust Statistics*. New York: Wiley, 1981.

[32] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust Statistics: The Approach Based on Influenc Functions*. New York: John Wiley & Sons, Inc., 1986.

[33] I. Arasaratnam, S. Haykin, Robert J. Elliott, "Discrete-time nonlinear filterin algorithms using GaussCHermite quadrature," *Proceedings of the IEEE*, vol. 95, no. 5, pp. 953-977, 2007.

[34] I. Arasaratnam, S. Haykin, Thomas R. Hurd, "Cubature Kalman filterin for continuous-discrete systems: theory and simulations," *IEEE Trans. Signal Processing*, vol. 58, no. 10, pp. 4977-4993, 2010.

[35] G. Evensen, "The ensemble Kalman filter theoretical formulation and practical implementation," *Ocean Dyn.*, vol. 53, no. 4, pp. 343C367, 2003.

[36] G. P. Huang, A. I. Mourikis, S. Roumeliotis, "On the complexity and consistency of UKF-based SLAM," *IEEE International Conference on Robotics and Automation*, pp. 4401-4408, 2009.

[37] B. Ait-El-Fquih, I. Hoteit, "Fast Kalman-like filterin for large-dimensional linear and Gaussian state-space models," *IEEE Trans. Signal Processing*, vol. 63, no. 21, pp. 5853-5867, 2015.

[38] IEEE PES TF on Benchmark System for Stability Controls, "Benchmark systems for small-signal stability analysis and control," Aug. 2015.

[39] J. H. Chow, K. W. Cheung, "A toolbox for power system dynamics and control engineering education and research," *IEEE Trans. Power Syst.*, vol. 7, no. 4, pp. 1559-1564, 1992.

[40] S. Benchluch, J. Chow, "A trajectory sensitivity method for the identificatio of nonlinear excitation system models," *IEEE Trans. Energy Conv.*, vol. 8, no. 2, pp. 159-164, 1993.

[41] Z. Huang, P. Du, D. Kosterev, S. Yang, "Generator dynamic model validation and parameter calibration using phasor measurements at the point of connection," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1939–1949, 2013.

[42] R. Huang, R. Diao, Y. Li, et. al, "Calibrating parameters of power system stability models using advanced ensemble Kalman filte ," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 2895-2905, 2018.

Junbo Zhao (M'18) is an assistant professor (research) at Virginia Tech since 2018. He received the Ph.D. degree from Bradley Department of Electrical and Computer Engineering at Virginia Tech in 2018. He did the summer internship at Pacifi Northwest National Laboratory from May-August, 2017. He is now the chair of IEEE Task Force on Power System Dynamic State and Parameter Estimation, the secretary of the IEEE Working Group on State Estimation Algorithms and the IEEE Task Force on Synchrophasor Applications in Power System Operation and Control.

He has written 2 book chapters, published more than 60 peer-reviewed journal and conference papers. His research interests are power system real-time monitoring, operations and cyber security that include state estimation, dynamics and stability, cyber attacks and countermeasures, big data analytics and robust statistics with applications in the smart grid. He serves as the Associate Editor of *International Journal of Electrical Power & Energy Systems* and the subject editor of *IET Generation, Transmission & Distribution*.

Lamine Mili (LF'17) received the Electrical Engineering Diploma from the Swiss Federal Institute of Technology, Lausanne, in 1976, and the Ph.D. degree from the University of Liège, Belgium, in 1987. He is a Professor of Electrical and Computer Engineering, Virginia Tech, Blacksburg. He has f ve years of industrial experience with the Tunisian electric utility, STEG. At STEG, he worked in the planning department from 1976 to 1979 and then at the Test and Meter Laboratory from 1979 till 1981. He was a Visiting Professor with the Swiss Federal Institute of Technology in Lausanne, the Grenoble Institute of Technology, the École Supérieure D'électricité in France and the École Polytechnique de Tunisie in Tunisia, and did consulting work for the French Power Transmission company, RTE.

His research has focused on power system planning for enhanced resiliency and sustainability, risk management of complex systems to catastrophic failures, robust estimation and control, nonlinear dynamics, and bifurcation theory. He is the co-founder and co-editor of the *International Journal of Critical Infrastructure*. He is the chairman of the IEEE Working Group on State Estimation Algorithms. He is a recipient of several awards including the US National Science Foundation (NSF) Research Initiation Award and the NSF Young Investigation Award.