The Stochastic Stability Analysis for Outlier Robustness of Kalman-type Filtering Framework Based on Correntropy-Induced Cost

Yangtianze Tao, Jiayi Kang and Stephen Shing-Toung Yau, *Life Fellow, IEEE*

***Abstract*—This note introduces the Modified Extended Kalman Filter (MEKF), reformulating the EKF update step within a nonlinear regression framework. We propose a novel outlier-robust scheme, MCIC-MEKF, utilizing the minimum correntropy-induced cost criterion. We provide a theoretical analysis of its outlier robustness through stochastic stability, proving exponentially bounded mean square posterior estimation error under natural conditions. Additionally, we present a technical approximation for the adaptive Kalman gain, enhancing efficiency without compromising performance. Simulation results confirm MCIC-MEKF's robustness against various non-Gaussian noises with large outliers, outperforming several filtering benchmarks.**

***Index Terms*—Stochastic Stability, Maximum Correntropy Criterion, Extended Kalman Filter, Outliers.**

I. Introduction

State estimation in state-space models is crucial for automatic control and signal processing, utilizing fundamental filtering techniques applied in numerous engineering applications [1]. The Kalman Filter (KF) is the optimal solution for linear systems with Gaussian noises [2], while nonlinear systems often employ the Extended Kalman Filter (EKF) [3], the Unscented Kalman Filter (UKF) [4], or the Cubature Kalman Filter (CKF) [5], all effective under Gaussian noise assumptions. However, non-Gaussian heavy-tailed noises, common in target tracking, audio communication, and power systems, necessitate alternatives like the particle filter (PF) [6], [7], which estimates posterior distributions through sequential Monte Carlo sampling [8] or control method modifications [9].

Non-Gaussian noise estimation challenges have spurred research into robust estimators for heavy-tailed noises. Key approaches include M-estimation-based Huber KF [10], [11], information theoretic methods [12], [13], and correntropy-based MCC filters [14]–[16]. While MCC filters show promise in handling outliers, their theoretical robustness quantification remains an open research area.

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 11961141005, in part by the Tsinghua University Startup Fund, and in part by the Tsinghua University Education Foundation Fund under Grant 042202008. *Corresponding author: Stephen Shing-Toung Yau.*

Yangtianze Tao is with the Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China (E-mail: calibertytz@163.com). Jiayi Kang is with Beijing Institute of Mathematical Sciences and Applications, Huairou district, Beijing 101400, China. (E-mail: kangjiayi@bimsa.cn). Yangtianze Tao and Jiayi Kang contribute equally.

Stephen Shing-Toung Yau is with the Beijing Institute of Mathematical Sciences and Applications, Huairou district, Beijing 101400, China and with the Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China, and (E-mail: yau@uic.edu).

In filtering problems, robustness is primarily considered in two contexts: noise uncertainty and model uncertainty. The former, as highlighted, involves resilience against inaccuracies in noise characterization, while the latter deals with discrepancies between the actual and assumed models, often termed as "model mismatch" [17], [18]. The model mismatch may result from discretization errors in continuous systems or an inadequate understanding of the noise processes. Detailed exploration of model uncertainty is provided in [19] and [20]. However, our discussion will mainly concentrate on addressing the robustness of noise uncertainty, especially in managing outliers within the noise.

This note presents a novel Modified EKF (MEKF), reformulating the EKF update step as a nonlinear regression optimization problem. We introduce the MCIC-MEKF, an outlier-robust scheme utilizing the minimum correntropy-induced cost (MCIC) [21]. Based on technical error representations [22], this approach aligns with the maximum correntropy criterion (MCC) but is framed as a regression cost minimization. The Kalman-type update features an adaptive Kalman gain computed via adjustable prior and observation noise covariance matrices. Our work analyzes the robustness of a novel Kalman-type update, examining its stochastic stability [23]–[25]. We quantify MCC-based filters' robustness, establishing conditions for consistent prior error estimation and exponentially bounded posterior error. Our approach optimizes adaptive Kalman gain without iterative solutions [15], [26], [27]. Simulations demonstrate MCIC-MEKF's robustness under non-Gaussian noise, contrasting with [19] by using a robust cost function instead of a probabilistic method.

In this note, vectors are denoted in boldface lowercase, and matrices in boldface uppercase. Transposition and expectation are symbolized by $\{\cdot\}^\top$ and $\mathbb{E}[\cdot]$, respectively. The Gaussian distribution is represented as $\mathcal{N}(\mu, \mathbf{\Sigma})$. The spaces $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$ indicate n-dimensional Euclidean space and the set of all $m \times n$ real matrices. The identity and zero matrices of dimension n are $\mathbf{I}_n$ and $\mathbf{0}_n$. The 2-norm, spectral norm, and Frobenius norm of a vector $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ and matrix $\mathbf{X} = [x_{ij}]_{n\times n} \in \mathbb{R}^{n\times n}$ are $\|\mathbf{x}\|$, $\|\mathbf{X}\|$, and $\|\mathbf{X}\|_F$, respectively. The weighted l^2 norm with respect to matrix $\mathbf{A}$ is $\|\mathbf{x}\|^2_{\mathbf{A}} = \mathbf{x}^\top\mathbf{A}^{-1}\mathbf{x}$. $\mathrm{diag}(x_1, \dots, x_p)$ constructs a diagonal matrix, and $\mathrm{Tr}(\cdot)$ is the trace. Matrix inequality $\mathbf{A} \geq \mathbf{B}$ ($\mathbf{A} > \mathbf{B}$) indicates $\mathbf{A} - \mathbf{B}$ is non-negative (positive) definite.

II. Preliminaries

In this section, we shall first introduce the formulation of robust filtering problems with outliers and some basic settings. Then we shall review the concept of the correntropy-induced cost.

A. Robust Filtering Problems with Outliers

Throughout this note, we consider the nonlinear autonomous system with state $\mathbf{x}_k \in \mathbb{R}^n$ and observation $\mathbf{y}_k \in \mathbb{R}^m$. It is given by the following state and observation equations:

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k) + \mathbf{W}_k \mathbf{w}_k \quad \text{(state equation)}, \tag{1a}$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{V}_k \mathbf{v}_k \quad \text{(observation equation)}, \tag{1b}$$

where $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ and $\mathbf{h} : \mathbb{R}^n \to \mathbb{R}^m$ are nonlinear functions that are assumed to be second-order continuously differentiable, called state function and observation function, respectively. State noise $\mathbf{w}_k$ and observation noise $\mathbf{v}_k$ are uncorrelated multivariate Gaussian with zero means and covariance matrices $\mathbf{Q}_k \in \mathbb{R}^{n\times n}$ and $\mathbf{R}_k \in \mathbb{R}^{m\times m}$, respectively.

B. Correntropy-Induced Cost

We now introduce the correntropy-induced cost [21].

Definition II.1 (Correntropy-Induced Cost)**.** *The correntropy-induced cost $l_\sigma(x) : \mathbb{R} \to [0, \infty)$ is defined as $l_\sigma(x) = \sigma^2 (1 - \mathcal{G}_\sigma(x))$, where $\mathcal{G}_\sigma(\cdot)$ is defined in Eq.(2) with $\sigma > 0$ being a scale parameter, where $\mathcal{G}_\sigma$ is the Gaussian Kernel given by*

$$\mathcal{G}_\sigma(e) = \exp\left(-\frac{e^2}{2\sigma^2}\right), \tag{2}$$

and $\sigma > 0$ stands for the kernel bandwidth.

They are denoted by minimum correntropy-induced cost (MCIC). Suppose our goal is to learn a parameter θ for a given estimator $\mathbf{x}(\theta)$, and let $\mathbf{y}$ denote the desired output. Then MCIC-based estimation can be formulated as solving the following optimzation problems:

$$\hat{\boldsymbol{\theta}}_{MCIC} = \arg\min_{\boldsymbol{\theta}\in\Omega} \ \mathbb{E}\left[l_\sigma\left(\|\mathbf{x}(\theta) - \mathbf{y}\|\right)\right], \tag{3}$$

III. Modified Extended Kalman Filter Framework

In this section, we present the MEKF Framework. Let $\hat{\mathbf{x}}_{k|k-1}$ and $\hat{\mathbf{x}}_{k|k}$ denote its estimated prior mean and estimated posterior mean, respectively. Then its prior estimation error $\tilde{\mathbf{e}}_{k|k-1}$ and posterior estimation error $\tilde{\mathbf{e}}_{k|k}$ can be written by

$$\tilde{\mathbf{e}}_{k|k-1} = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}, \tilde{\mathbf{e}}_{k|k} = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k}. \tag{4}$$

Let

$$\mathbf{F}_k = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}_k}\right|_{\mathbf{x}_k = \hat{\mathbf{x}}_{k-1|k-1}}, \quad \mathbf{H}_k = \left.\frac{\partial \mathbf{h}}{\partial \mathbf{x}_k}\right|_{\mathbf{x}_k = \hat{\mathbf{x}}_{k|k-1}}. \tag{5}$$

To facilitate the subsequent deduction, it is essential to establish a representation for the linearization error, akin to the approach presented in [22]. This approach elucidates that the discrepancy between the function evaluated at the true state and its estimate can be expressed as follows:

$$\begin{aligned}\mathbf{f}(\mathbf{x}_k) - \mathbf{f}(\hat{\mathbf{x}}_{k|k}) &= \mathbf{F}_{k+1}\tilde{\mathbf{e}}_{k|k} + \mathbf{A}_k \boldsymbol{\alpha}_k \tilde{\mathbf{e}}_{k|k} \\ &= \hat{\mathbf{F}}_{k+1}\tilde{\mathbf{e}}_{k|k}.\end{aligned} \tag{6}$$

Similarly there exist a problem-dependent scaling matrix $\mathbf{B}_k \in \mathbb{R}^{m\times n}$ and a unknown time-varying matrix $\boldsymbol{\beta}_k \in \mathbb{R}^{m\times n}$ with $\|\boldsymbol{\beta}_k\| \le 1$ making $\mathbf{h}$ can be rewritten by

$$\begin{aligned}\mathbf{h}(\mathbf{x}_k) - \mathbf{h}(\hat{\mathbf{x}}_{k|k-1}) &= \mathbf{H}_k\tilde{\mathbf{e}}_{k|k-1} + \mathbf{B}_k \boldsymbol{\beta}_k \tilde{\mathbf{e}}_{k|k-1} \\ &= \hat{\mathbf{H}}_k\tilde{\mathbf{e}}_{k|k-1}.\end{aligned} \tag{7}$$

Here $\hat{\mathbf{F}}_k$ and $\hat{\mathbf{H}}_k$ are given by

$$\hat{\mathbf{F}}_k = \mathbf{F}_k + \mathbf{A}_{k-1}\boldsymbol{\alpha}_{k-1}, \quad \hat{\mathbf{H}}_k = \mathbf{H}_k + \mathbf{B}_k\boldsymbol{\beta}_k. \tag{8}$$

Remark III.1. *We use matrices $\boldsymbol{\alpha}_k, \boldsymbol{\beta}_k, \mathbf{A}_k, \mathbf{B}_k$ to address linearization errors, ensuring exact equalities. This approach, similar to methods in [28]–[30], supports Kalman-type filter stability. Our stability analysis is independent of $\mathbf{A}_k, \mathbf{B}_k, \boldsymbol{\alpha}_k, \boldsymbol{\beta}_k$ magnitudes, enabling further theoretical exploration.*

A. Nonlinear Regression Form and Optimization-based Update Step

Here we shall illustrate how to transform the robust filtering as a nonlinear regression problem. Let $\mathbf{P}_{k|k-1}$ and $\mathbf{P}_{k|k}$ be prior covariance and posterior covariance, respectively. Then we consider the augmented model, which is given by

$$\begin{bmatrix} \hat{\mathbf{x}}_{k|k-1} \\ \mathbf{y}_k \end{bmatrix} = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{h}(\mathbf{x}_k) \end{bmatrix} + \tilde{\mathbf{V}}_k, \tag{9}$$

where $\tilde{\mathbf{V}}_k$ is given by $\tilde{\mathbf{V}}_k = \begin{bmatrix} -\tilde{\mathbf{e}}_{k|k-1} \\ \mathbf{V}_k\mathbf{v}_k \end{bmatrix}$. Here we consider Cholesky decompositions of $\mathbf{P}_{k|k-1}$ and $\mathbf{R}_k$, which are given by

$$\mathbf{P}_{k|k-1} = \mathbf{B}^p_{k|k-1}\left(\mathbf{B}^p_{k|k-1}\right)^\top, \mathbf{R}_k = \mathbf{B}^r_k\left(\mathbf{B}^r_k\right)^\top. \tag{10}$$

Then it is easy to see that

$$\mathbb{E}\left[\tilde{\mathbf{V}}_k\tilde{\mathbf{V}}_k^\top\right] = \begin{bmatrix} \mathbf{P}_{k|k-1} & 0 \\ 0 & \mathbf{V}_k\mathbf{R}_k\mathbf{V}_k^\top \end{bmatrix} = \mathbf{B}_k\mathbf{B}_k^\top, \tag{11}$$

with $\mathbf{B}_k = \begin{bmatrix} \mathbf{B}^p_{k|k-1} & 0 \\ 0 & \mathbf{V}_k\mathbf{B}^r_k \end{bmatrix}$. Left multiplication on both sides of (9) by $\mathbf{B}_k^{-1}$, we obtain

$$\mathbf{d}_k = \mathbf{m}_k(\mathbf{x}_k) + \mathbf{e}_k, \tag{12}$$

where $\mathbf{d}_k = \begin{bmatrix} \left(\mathbf{B}^p_{k|k-1}\right)^{-1}\hat{\mathbf{x}}_{k|k-1} \\ \left(\mathbf{V}_k\mathbf{B}^r_k\right)^{-1}\mathbf{y}_k \end{bmatrix}$ and $\mathbf{m}_k(\mathbf{x}_k) = \begin{bmatrix} \left(\mathbf{B}^p_{k|k-1}\right)^{-1}\mathbf{x}_k \\ \left(\mathbf{V}_k\mathbf{B}^r_k\right)^{-1}\mathbf{h}(\mathbf{x}_k) \end{bmatrix}$. Here we need to assume that $\mathbb{E}\left[\tilde{\mathbf{e}}_{k|k-1}\tilde{\mathbf{e}}_{k-l|k-l-1}^\top\right] = 0$ for $l \neq 0$. Note tha $\mathbf{e}_k = \mathbf{B}_k^{-1}\tilde{\mathbf{V}}_k$, which implies $\mathbb{E}\left[\mathbf{e}_k\mathbf{e}_k^\top\right] = \mathbf{I}_{n+m}$, hence the residual error $\mathbf{e}_k$ is white noise, which makes Eq.(12) become a nonlinear regression function. With the help of regression function (12), we can formulate a optimization-based filtering update step. It will be given by mean of a cost function $\mathcal{J}$,

$$\hat{\mathbf{x}}_{k|k} = \arg\min_{\mathbf{x}_k} \mathcal{J}(\mathbf{x}_k). \tag{13}$$

Let us denote

$$\begin{aligned} \mathbf{e}_k^x &= \left(\mathbf{B}_{k|k-1}^p\right)^{-1} \left(\hat{\mathbf{x}}_{k|k-1} - \mathbf{x}_k\right) \\ \mathbf{e}_k^y &= \left(\mathbf{V}_k \mathbf{B}_k^r\right)^{-1} \left(\mathbf{y}_k - \mathbf{h}(\mathbf{x}_k)\right). \end{aligned} \tag{14}$$

Then $\mathcal{J}$ can take the following two forms $\mathcal{J}_1$ and $\mathcal{J}_2$,

$$\mathcal{J}_1(\mathbf{x}_k) = \rho\left(\|\mathbf{e}_k^x\|\right) + \rho\left(\|\mathbf{e}_k^y\|\right), \mathcal{J}_2(\mathbf{x}_k) = \sum_{i=1}^{n+m} \rho\left(e_{k,i}\right), \tag{15}$$

where $e_{k,i}$ is the i-th component of the residual vector $\mathbf{e}_k$. ρ is a robust cost function (e.g., correntropy-induce cost) which is used to cut off the outliers in residual vector $\mathbf{e}_k$.

B. Outlier-Robust MEKF Schemes

Based on the optimization framework (13), we derive a novel outlier-robust EKF scheme using MCIC, termed MCIC-MEKF with two variants, MCIC-MEKF-1 and MCIC-MEKF-2, using cost functions $\mathcal{J}_1$ and $\mathcal{J}_2$ respectively, with $\rho(x) = l_\sigma(x)$. These variants differ in adaptive adjustment scale but share the same adjustment method. These two schemes share the similar prediction step with the common EKF, which is given by

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{f}(\hat{\mathbf{x}}_{k|k-1}), \tag{16}$$

and

$$\mathbf{P}_{k|k-1} = \hat{\mathbf{F}}_k \mathbf{P}_{k-1|k-1} \hat{\mathbf{F}}_k^\top + \mathbf{W}_k \mathbf{Q}_k \mathbf{W}_k^\top. \tag{17}$$

Nextly, we shall deduce their update steps.

1) MCIC-MEKF-1: Let us take $\rho(x) = l_\sigma(x)$ in $\mathcal{J}_1$ defined in Eq.(15) and denote it by $\mathcal{L}_1$, which is given by

$$\begin{aligned} \mathcal{L}_1(\mathbf{x}_k) = &\, l_\sigma\left(\|\mathbf{y}_k - \mathbf{h}(\mathbf{x}_k)\|_{\mathbf{V}_k \mathbf{R}_k \mathbf{V}_k^\top}\right) \\ &+ l_\sigma\left(\|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}\|_{\mathbf{P}_{k|k-1}}\right). \end{aligned} \tag{18}$$

Then based on $\mathcal{L}_1(\mathbf{x}_k)$, the update step for MCIC-MEKF-1 can be obtained by solving the following optimization problem:

$$\hat{\mathbf{x}}_{k|k} = \arg\min_{\mathbf{x}_k} \ \mathcal{L}_1(\mathbf{x}_k). \tag{19}$$

Solving optimzation problem Eq.(19), we can obtain the following new Kalman-type update step.

$$\mathbf{x}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k^\star\left(\mathbf{y}_k - \mathbf{h}\left(\hat{\mathbf{x}}_{k|k-1}\right)\right), \tag{20}$$

and its Kalman gain $\mathbf{K}_k^\star$ is given by

$$\mathbf{K}_k^\star = \mathbf{P}_{k|k-1}^\star \hat{\mathbf{H}}_k^\top \left(\hat{\mathbf{H}}_k \mathbf{P}_{k|k-1}^\star \hat{\mathbf{H}}_k^\top + \mathbf{R}_k^\star\right)^{-1}, \tag{21}$$

where $\mathbf{P}_{k|k-1}^\star$ and $\mathbf{R}_k^\star$ are given by

$$\mathbf{P}_{k|k-1}^\star = \frac{\mathbf{P}_{k|k-1}}{l_{\mathbf{P}_{k|k-1}}}, \quad \mathbf{R}_k^\star = \frac{\mathbf{V}_k \mathbf{R}_k \mathbf{V}_k^\top}{l_{\mathbf{R}_k}}, \tag{22}$$

with

$$\begin{aligned} l_{\mathbf{R}_k} &= \mathcal{G}_\sigma\left(\|\mathbf{y}_k - \mathbf{h}\left(\mathbf{x}_k\right)\|_{\mathbf{V}_k \mathbf{R}_k \mathbf{V}_k^\top}\right) \\ l_{\mathbf{P}_{k|k-1}} &= \mathcal{G}_\sigma\left(\|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}\|_{\mathbf{P}_{k|k-1}}\right). \end{aligned} \tag{23}$$

The corresponding estimation error covariance can be calculated as

$$\mathbf{P}_{k|k} = \left(\mathbf{I}_n - \mathbf{K}_k^\star \hat{\mathbf{H}}_k\right) \mathbf{P}_{k|k-1} \left(\mathbf{I}_n - \mathbf{K}_k^\star \hat{\mathbf{H}}_k\right)^\top + \mathbf{K}_k^\star \mathbf{V}_k \mathbf{R}_k \mathbf{V}_k^\top \left(\mathbf{K}_k^\star\right)^\top. \tag{24}$$

2) MCIC-MEKF-2: Let us still take $\rho(x) = l_\sigma(x)$ in $\mathcal{J}_2$ defined in Eq.(15) and denote it by $\mathcal{L}_2$, which is given by

$$\mathcal{L}_2(\mathbf{x}_k) = \sum_{i=1}^{n} l_\sigma\left(e_{k,i}^x\right) + \sum_{j=1}^{m} l_\sigma\left(e_{k,j}^y\right), \tag{25}$$

where $e_{k,i}^x$ and $e_{k,j}^y$ are the i-th and the j-th components of $\mathbf{e}_k^x$ and $\mathbf{e}_k^x$, respectively. Then, under cost function $\mathcal{L}_2$, the update step for MCIC-MEKF-2 can be obtained by solving the following optimization problem:

$$\hat{\mathbf{x}}_{k|k} = \arg\min_{\mathbf{x}_k} \ \mathcal{L}_2(\mathbf{x}_k). \tag{26}$$

Solving optimzation problem Eq.(26), we can obtain the following another new Kalman-type update step.

$$\mathbf{x}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k^*(\mathbf{y}_k - \mathbf{h}\left(\hat{\mathbf{x}}_{k|k-1}\right)), \tag{27}$$

where the Kalman gain is given by

$$\mathbf{K}_k^* = \mathbf{P}_{k|k-1}^* \hat{\mathbf{H}}_k^\top \left(\hat{\mathbf{H}}_k \mathbf{P}_{k|k-1}^* \hat{\mathbf{H}}_k^\top + \mathbf{R}_k^*\right)^{-1}, \tag{28}$$

with

$$\begin{aligned} \mathbf{P}_{k|k-1}^* &= \mathbf{B}_{k|k-1}^p \left(\mathbf{C}_k^x\right)^{-1} \left(\mathbf{B}_{k|k-1}^p\right)^\top \\ \mathbf{R}_k^* &= \mathbf{V}_k \mathbf{B}_k^r \left(\mathbf{C}_k^y\right)^{-1} \left(\mathbf{B}_k^r\right)^\top \mathbf{V}_k^\top. \end{aligned} \tag{29}$$

Here $\mathbf{C}_k^x$ and $\mathbf{C}_k^y$ are given by

$$\begin{aligned} \mathbf{C}_k^x &= \text{diag}(\mathcal{G}_\sigma(e_{k,1}^x), \mathcal{G}_\sigma(e_{k,2}^x), \ldots, \mathcal{G}_\sigma(e_{k,n}^x)) \\ \mathbf{C}_k^y &= \text{diag}(\mathcal{G}_\sigma(e_{k,1}^y), \mathcal{G}_\sigma(e_{k,2}^y), \ldots, \mathcal{G}_\sigma(e_{k,m}^y)). \end{aligned} \tag{30}$$

The corresponding estimation error covariance can be calculated as

$$\begin{aligned} \mathbf{P}_{k|k} = &\left(\mathbf{I}_n - \mathbf{K}_k^* \hat{\mathbf{H}}_k\right) \mathbf{P}_{k|k-1} \left(\mathbf{I}_n - \mathbf{K}_k^* \hat{\mathbf{H}}_k\right)^\top \\ &+ \mathbf{K}_k^* \mathbf{V}_k \mathbf{R}_k \mathbf{V}_k^\top \left(\mathbf{K}_k^*\right)^\top. \end{aligned} \tag{31}$$

IV. Stochastic Stability Analysis and Practical Algorithm

In this section, we shall understand the robustness of our proposed MCIC-MEKF-1 in a quantitative way.

A. Preparations

We shall first deduce the recursion of prior estimation error $\tilde{\mathbf{e}}_{k|k-1}$ and posterior estimation error $\tilde{\mathbf{e}}_{k|k}$. In view of Eq.(7) and Eq.(1b), one can conclude that

$$\tilde{\mathbf{e}}_{k|k} = \left(\mathbf{I}_n - \mathbf{K}_k^\star \hat{\mathbf{H}}_k\right) \tilde{\mathbf{e}}_{k|k-1} - \mathbf{K}_k^\star \mathbf{V}_k \mathbf{v}_k. \tag{32}$$

Here we recall the following definition.

Definition IV.1. *The stochastic process $\mathbf{z}_k$ is said to be exponentially bounded in mean square, if $\mathbb{E}\left[\|\mathbf{z}_0\|^2\right] < \infty$ and there are real numbers $\beta, v > 0$ and $0 < \alpha < 1$ such that*

$$\mathbb{E}\left[\|\mathbf{z}_k\|^2\right] \leq \beta \mathbb{E}\left[\|\mathbf{z}_0\|^2\right] \alpha^k + v$$

holds for every integer $k \geq 0$.

Nextly, we shall give the following assumptions.

Assumption IV.1. *There are positive real numbers* $\underline{f}, \bar{f}, \underline{h}, \bar{h}, \underline{p}, \underline{q}, \bar{q}, \bar{p}, \underline{r}, \bar{r}, \kappa_{\mathbf{w}}, \kappa_{\mathbf{v}}, \bar{w}, \bar{v} > 0$ *such that for above mentioned matrices the following bounds are satisfied for every integer* $k \geq 0$*:*

$$\underline{f} \leq \|\hat{\mathbf{F}}_k\| \leq \bar{f}, \quad \underline{h} \leq \|\hat{\mathbf{H}}_k\| \leq \bar{h}, \tag{33}$$

$$\underline{p}\mathbf{I}_n \leq \mathbf{P}_{k+1|k} \leq \bar{p}\mathbf{I}_n, \tag{34}$$

$$\underline{q}\mathbf{I}_n \leq \mathbf{Q}_k \leq \bar{q}\mathbf{I}_n, \quad \underline{r}\mathbf{I}_m \leq \mathbf{R}_k \leq \bar{r}\mathbf{I}_m. \tag{35}$$

$$\mathbb{E}\left[\mathbf{w}_k^\top \mathbf{w}_k\right] \leq \kappa_{\mathbf{w}}, \mathbb{E}\left[\mathbf{v}_k^\top \mathbf{v}_k\right] \leq \kappa_{\mathbf{v}}. \tag{36}$$

$$1 \leq \|\mathbf{W}_k\| \leq \bar{w}, 1 \leq \|\mathbf{V}_k\| \leq \bar{v}. \tag{37}$$

Assumption IV.2. *For every integer* $k \geq 0$, $\hat{\mathbf{F}}_k$ *is non-singular.*

Remark IV.1. *In fact, Eq.*(36) *can be deduced by Eq.*(35)*, details can be found in Appendix A-B. We put it in the Assumption IV.1 for convenience.*

Remark IV.2. *These boundedness conditions* (33)*,* (35) *in Assumption IV.1 are relatively natural for the practical system. In fact, bounded condition* (34) *is related to the uniform detectability [31] for the linear time-varying system. It is very similar to the uniform observability [32] of linear systems. The upper bounds in* (37) *portrays the upper bounds of the influence of outliers.*

B. Main results

Before we proceed, we need some technical lemmas. Our first lemma regarding the boundedness of stochastic processes is crucial to our theoretical analysis.

Lemma IV.1. *Assume there is a stochastic Lyapunov function* $\mathcal{V}_k(\mathbf{x}_k)$ *and real numbers* $\underline{v}, \bar{v}$ *such that for every integer* $k \geq 0$,

$$\underline{v}\|\mathbf{x}_k\|^2 \leq \mathcal{V}_k(\mathbf{x}_k) \leq \bar{v}\|\mathbf{x}_k\|^2, \tag{38}$$

and there are corresponding $\mu_k > 0$ *and* $0 < \alpha_k < 1$ *making*

$$\mathbb{E}\left[\mathcal{V}_{k+1}(\mathbf{x}_{k+1}) \mid \mathbf{x}_k\right] \leq (1-\alpha_k)\mathcal{V}_k(\mathbf{x}_k) + \mu_k \tag{39}$$

Then for every integer $k \geq 0$,

$$\mathbb{E}\left[\|\mathbf{x}_k\|^2\right] \leq \frac{\bar{v}}{\underline{v}}\mathbb{E}\left[\|\mathbf{x}_0\|^2\right]\prod_{i=0}^{k-1}(1-\alpha_i) + \frac{1}{\underline{v}}\sum_{i=0}^{k-1}\mu_i(1-\alpha_i)^i. \tag{40}$$

Moreover, the stochastic process $\mathbf{x}_k$ *is exponentially bounded in mean square, i.e.,*

$$\mathbb{E}\left[\|\mathbf{x}_k\|^2\right] \leq \frac{\bar{v}}{\underline{v}}\mathbb{E}\left[\|\mathbf{x}_0\|^2\right](1-\alpha)^k + \frac{\mu}{\underline{v}\alpha}, \tag{41}$$

where $\alpha = \min_{i=0,1,\ldots,k-1}\{\alpha_i\}$ *and* $\mu = \max_{i=0,1,\ldots,k-1}\{\mu_i\}$.

Proof. See Appendix IV.1. □

Lemma IV.2 (Bound of Kalman Gain)**.** *If the assumption IV.1 holds, the lower bound and the uppper bound of Kalman gain* $\mathbf{K}_k^\star$ *in* (21) *are given by*

$$\iota_{\mathbf{K}_k^\star} \leq \|\mathbf{K}_k^\star\| \leq \bar{\iota}_{\mathbf{K}_k^\star}, \tag{42}$$

where $\underline{\iota}_{\mathbf{K}_k^\star} = \frac{\underline{p}\underline{h}}{\bar{h}^2\bar{p}+\lambda_k\bar{v}\bar{r}}$*, and* $\bar{\iota}_{\mathbf{K}_k^\star} = \frac{\bar{p}\bar{h}}{\underline{h}^2\underline{p}+\lambda_k\underline{r}}$*. Here adaptive weight ratio* λ_k *is given* $\lambda_k = \frac{l_{\mathbf{P}_{k|k-1}}}{l_{\mathbf{R}_k}}$.

Proof. See Appendix A-C. □

Lemma IV.3. *If the assumption IV.1 and assumption IV.2 hold. Then for every integer* $k \geq 0$*, there exist a real number* $0 < \alpha_k < 1$ *such that* $\mathbf{P}_{k+1|k}^{-1}$ *satisfies:*

$$\begin{aligned}&\left(\hat{\mathbf{F}}_{k+1} - \hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\hat{\mathbf{H}}_k\right)^\top \mathbf{P}_{k+1|k}^{-1}\left(\hat{\mathbf{F}}_{k+1} - \hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\hat{\mathbf{H}}_k\right)\\&\leq (1-\alpha_k)\mathbf{P}_{k|k-1}^{-1},\end{aligned} \tag{43}$$

where $\alpha_k = \frac{\underline{q}}{\bar{p}\left(\bar{f}+\bar{f}\bar{\iota}_{\mathbf{K}_k^\star}\bar{h}\right)^2+\underline{q}}$.

Proof. See Appendix A-D. □

Lemma IV.4. *If the assumption IV.1 holds. Then for every integer* $k \geq 0$,

$$\begin{aligned}&\mathbb{E}\left[\left(\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\mathbf{V}_k\mathbf{v}_k\right)^\top \mathbf{P}_{k+1|k}^{-1}\left(\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\mathbf{V}_k\mathbf{v}_k\right) \mid \tilde{\mathbf{e}}_{k|k-1}\right]\\&\leq \frac{\left(\bar{f}\bar{\iota}_{\mathbf{K}_k^\star}\|\mathbf{V}_k\|\right)^2\kappa_{\mathbf{v}}}{\underline{p}}.\end{aligned} \tag{44}$$

Proof. See Appendix A-E. □

Now we are ready to present our main result.

Theorem IV.1. *Consider the nonlinear system given by* (1) *and the MCIC-MEKF-1 as stated above. If Assumption IV.1 and Assumption IV.2 hold and initial prior estimation error satisfies* $\mathbb{E}\left[\|\tilde{\mathbf{e}}_{1|0}\|^2\right] < \infty$*, then it holds, 1. The one-step prior estimation error is given by*

$$\mathbb{E}\left[\|\tilde{\mathbf{e}}_{k+1|k}\|^2\right] \leq \frac{\bar{p}}{\underline{p}}(1-\alpha_k)\mathbb{E}\left[\|\tilde{\mathbf{e}}_{k|k-1}\|^2\right] + \bar{p}\mu_k, \tag{45}$$

where and μ_k *is given by*

$$\mu_k = \frac{\left(\bar{f}\bar{\iota}_{\mathbf{K}_k^\star}\|\mathbf{V}_k\|\right)^2\kappa_{\mathbf{v}} + \kappa_{\mathbf{w}}\|\mathbf{W}_k\|^2}{\underline{p}}. \tag{46}$$

2. For every integer $k > 0$*, the posterior estimation error is given by*

$$\begin{aligned}\mathbb{E}\left[\|\tilde{\mathbf{e}}_{k|k}\|^2\right] &\leq 2\left(\underline{f}\right)^{-2}\frac{\bar{p}}{\underline{p}}\mathbb{E}\left[\|\tilde{\mathbf{e}}_{1|0}\|^2\right]\prod_{i=0}^{k-1}(1-\alpha_i)^i\\&+ 2\left(\underline{f}\right)^{-2}\left(\bar{p}\sum_{i=0}^{k-1}\mu_i(1-\alpha_i)^i + \|\mathbf{W}_k\|^2\kappa_{\mathbf{w}}\right).\end{aligned} \tag{47}$$

Moreover, if adaptive weight ratio λ_k *is bounded, i.e.,* $0 < \underline{\lambda} \leq \lambda_k \leq \bar{\lambda}$ *for some* $0 < \underline{\lambda} < \bar{\lambda}$*. Then the posterior estimation* *error $\tilde{\mathbf{e}}_{k|k}$ is exponentially bounded in mean square, i.e., for every integer $k > 0$,*

$$\begin{aligned}\mathbb{E}\left[\left\|\tilde{\mathbf{e}}_{k|k}\right\|^2\right] \leq 2\left(\underline{f}\right)^{-2}\frac{\bar{p}}{\underline{p}}\mathbb{E}\left[\left\|\tilde{\mathbf{e}}_{1|0}\right\|^2\right](1-\alpha)^k \\ + 2\left(\underline{f}\right)^{-2}\left(\frac{\bar{p}\mu}{\alpha} + \bar{w}^2\kappa_{\mathbf{w}}\right),\end{aligned} \tag{48}$$

where $\mu = \frac{\left(\frac{\bar{f}\bar{p}\bar{h}\bar{v}}{\underline{h}^2\underline{p}+\bar{\lambda}\underline{r}}\right)^2\kappa_{\mathbf{v}} + \bar{w}^2\kappa_{\mathbf{w}}}{\underline{p}}$ *and* $\alpha = \frac{\underline{q}}{\bar{p}\left(\bar{f}+\frac{\bar{f}\bar{p}\bar{h}^2}{\underline{h}^2\underline{p}+\bar{\lambda}\underline{r}}\right)^2+\underline{q}}$.

Proof. We consider a stochastic Lyapunov function $\mathcal{V}_{k+1}\left(\tilde{\mathbf{e}}_{k+1|k}\right)$, which is given by

$$\mathcal{V}_{k+1}\left(\tilde{\mathbf{e}}_{k+1|k}\right) = \tilde{\mathbf{e}}_{k+1|k}^\top\mathbf{P}_{k+1|k}^{-1}\tilde{\mathbf{e}}_{k+1|k}. \tag{49}$$

It is easy to see that

$$\frac{1}{\bar{p}}\|(\tilde{\mathbf{e}}_{k+1|k}\|^2 \leq \mathcal{V}_{k+1}\left(\tilde{\mathbf{e}}_{k+1|k}\right) \leq \frac{1}{\underline{p}}\|(\tilde{\mathbf{e}}_{k+1|k}\|^2. \tag{50}$$

We consider

$$\begin{aligned}&\mathbb{E}\left[\mathcal{V}_{k+1}\left(\tilde{\mathbf{e}}_{k+1|k}\right) \mid \tilde{\mathbf{e}}_{k|k-1}\right] \\ &= \tilde{\mathbf{e}}_{k|k-1}^\top\left(\hat{\mathbf{F}}_{k+1} - \hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\hat{\mathbf{H}}_k\right)^\top\mathbf{P}_{k+1|k}^{-1} \\ &\times\left(\hat{\mathbf{F}}_{k+1} - \hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\hat{\mathbf{H}}_k\right)\tilde{\mathbf{e}}_{k|k-1} \\ &+ \mathbb{E}\left[\left(\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\mathbf{V}_k\mathbf{v}_k\right)^\top\mathbf{P}_{k+1|k}^{-1}\left(\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\mathbf{V}_k\mathbf{v}_k\right) \mid \tilde{\mathbf{e}}_{k|k-1}\right] \\ &+ \mathbb{E}\left[\mathbf{w}_k^\top\mathbf{W}_k^\top\mathbf{P}_{k+1|k}^{-1}\mathbf{W}_k\mathbf{w}_k \mid \tilde{\mathbf{e}}_{k|k-1}\right].\end{aligned} \tag{51}$$

Note that in Eq.(51), the cross terms all vanish after taking the conditional expectation since $\mathbf{w}_k$ and $\mathbf{v}_k$ are uncorrelated noises with zero means. Then in view of lemma IV.3, we have

$$\begin{aligned}&\mathbb{E}\left[\mathcal{V}_{k+1}\left(\tilde{\mathbf{e}}_{k+1|k}\right) \mid \tilde{\mathbf{e}}_{k|k-1}\right] \\ &\leq (1-\alpha)\mathcal{V}_k\left(\tilde{\mathbf{e}}_{k|k-1}\right) \\ &+ \mathbb{E}\left[\left(\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\mathbf{V}_k\mathbf{v}_k\right)^\top\mathbf{P}_{k+1|k}^{-1}\left(\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\mathbf{V}_k\mathbf{v}_k\right) \mid \tilde{\mathbf{e}}_{k|k-1}\right] \\ &+ \mathbb{E}\left[\mathbf{w}_k^\top\mathbf{W}_k^\top\mathbf{P}_{k+1|k}^{-1}\mathbf{W}_k\mathbf{w}_k \mid \tilde{\mathbf{e}}_{k|k-1}\right].\end{aligned} \tag{52}$$

Note that

$$\mathbb{E}\left[\mathbf{w}_k^\top\mathbf{W}_k^\top\mathbf{P}_{k+1|k}^{-1}\mathbf{W}_k\mathbf{w}_k \mid \tilde{\mathbf{e}}_{k|k-1}\right] \leq \frac{\kappa_{\mathbf{w}}\|\mathbf{W}_k\|^2}{\underline{p}}. \tag{53}$$

Then in view of lemma IV.4 and Eq.(46), we have

$$\mathbb{E}\left[\mathcal{V}_{k+1}\left(\tilde{\mathbf{e}}_{k+1|k}\right) \mid \tilde{\mathbf{e}}_{k|k-1}\right] \leq (1-\alpha_k)\mathcal{V}_k\left(\tilde{\mathbf{e}}_{k|k-1}\right) + \mu_k. \tag{54}$$

Here we shall first prove Eq.(45). In view of Eq.(50), we can deduce that

$$\begin{aligned}&\frac{1}{\bar{p}}\mathbb{E}\left[\|\tilde{\mathbf{e}}_{k+1|k}\|^2\right] \leq \mathbb{E}\left[\mathcal{V}_{k+1}\left(\tilde{\mathbf{e}}_{k+1|k}\right) \mid \tilde{\mathbf{e}}_{k|k-1}\right] \\ &(1-\alpha_k)\mathcal{V}_k\left(\tilde{\mathbf{e}}_{k|k-1}\right) \leq \frac{1}{\underline{p}}(1-\alpha_k)\|\tilde{\mathbf{e}}_{k|k-1}\|^2.\end{aligned} \tag{55}$$

Then in view of Eq.(54), we have

$$\frac{1}{\bar{p}}\mathbb{E}\left[\|\tilde{\mathbf{e}}_{k+1|k}\|^2\right] \leq \frac{1}{\underline{p}}(1-\alpha_k)\mathbb{E}\left[\|\tilde{\mathbf{e}}_{k|k-1}\|^2\right] + \mu_k \tag{56}$$

It follows that Eq.(45) holds. Now we shall prove Eq.(47). Let $\underline{v} = \frac{1}{\bar{p}}$ and $\bar{v} = \frac{1}{\underline{p}}$. By Lemma IV.1, we have

$$\mathbb{E}\left[\|\tilde{\mathbf{e}}_{k+1|k}\|^2\right] \leq \frac{\bar{p}}{\underline{p}}\mathbb{E}\left[\left\|\tilde{\mathbf{e}}_{1|0}\right\|^2\right]\prod_{i=0}^{k}(1-\alpha_i)^i + \bar{p}\sum_{i=0}^{k}\mu_i(1-\alpha_i)^i. \tag{57}$$

We can conclude that

$$\begin{aligned}\|\tilde{\mathbf{e}}_{k+1|k}\| &= \|\hat{\mathbf{F}}_{k+1}\tilde{\mathbf{e}}_{k|k} + \mathbf{W}_k\mathbf{w}_k\| \\ &\geq \|\hat{\mathbf{F}}_{k+1}\tilde{\mathbf{e}}_{k|k}\| - \|\mathbf{W}_k\mathbf{w}_k\|,\end{aligned} \tag{58}$$

which implies

$$\|\hat{\mathbf{F}}_{k+1}\tilde{\mathbf{e}}_{k|k}\|^2 \leq 2\left(\|\tilde{\mathbf{e}}_{k+1|k}\|^2 + \|\mathbf{W}_k\mathbf{w}_k\|^2\right). \tag{59}$$

It follows that

$$\begin{aligned}\mathbb{E}\left[\|\tilde{\mathbf{e}}_{k|k}\|^2\right] &\leq 2\left(\underline{f}\right)^{-2}\left(\mathbb{E}\left[\|\tilde{\mathbf{e}}_{k+1|k}\|^2\right] + \mathbb{E}\left[\|\mathbf{W}_k\mathbf{w}_k\|^2\right]\right) \\ &\leq 2\left(\underline{f}\right)^{-2}\left(\mathbb{E}\left[\|\tilde{\mathbf{e}}_{k+1|k}\|^2\right] + \|\mathbf{W}_k\|^2\kappa_{\mathbf{w}}\right).\end{aligned} \tag{60}$$

Substituting Eq.(57) into Eq.(60), we can deduce that

$$\begin{aligned}\mathbb{E}\left[\|\tilde{\mathbf{e}}_{k|k}\|^2\right] &\leq 2\left(\underline{f}\right)^{-2}\frac{\bar{p}}{\underline{p}}\mathbb{E}\left[\left\|\tilde{\mathbf{e}}_{1|0}\right\|^2\right]\prod_{i=0}^{k}(1-\alpha_i)^i \\ &+ 2\left(\underline{f}\right)^{-2}\left(\bar{p}\sum_{i=0}^{k}\mu_i(1-\alpha_i)^i + \|\mathbf{W}_k\|^2\kappa_{\mathbf{w}}\right),\end{aligned} \tag{61}$$

which is our desired Eq.(47). At last, we shall derive Eq.(48). Note that

$$\frac{\bar{p}\bar{h}}{\underline{h}^2\underline{p}+\bar{\lambda}\underline{r}} \leq \bar{\iota}_{\mathbf{K}_k^\star} \leq \frac{\bar{p}\bar{h}}{\underline{h}^2\underline{p}+\underline{\lambda r}}. \tag{62}$$

It follows that $\alpha_k \leq \alpha$ and $\mu_k \leq \mu$. Then Eq.(48) can be easily derived following Lemma IV.1. □

Parameters α_k and μ_k depend on $\mathbf{A}_k, \mathbf{B}_k, \boldsymbol{\alpha}_k, \boldsymbol{\beta}_k$ magnitudes but remain bounded when $\mathbf{Q}_k$ and $\mathbf{R}_k$ are positive definite. MCIC-MEKF-1's robustness, with exponentially bounded posterior error, is ensured by bounded λ_k. This adaptive weight ratio influences error variance upper bound via $\bar{\iota}_{\mathbf{K}_k^\star}$, enhancing robustness to observation outliers by adjusting $\|\mathbf{V}_k\|$ impact, with indirect effect on state outliers. The essential formula encapsulating the effect of adaptive weighting, μ_k, elegantly decomposes into components reflecting the influence of observation noise and prior state uncertainty, moderated by λ_k.

C. Practical Algorithm

At last, we shall consider the practical algorithm forms for MCIC-MEKF-1 and MCIC-MEKF-2 proposed previously. In fact, recall $l_{\mathbf{R}_k}$ and $l_{\hat{\mathbf{C}}_k}$ defined in Eq.(23). They all contain the item $\mathbf{x}_k$, hence the optimal solution Eq.(13) is in fact the zero point of the (20). A typical way of dealing with Eq.(20) is to view it as a fixed-point equation of $\mathbf{x}_k$, whose solution can be obtained by a fixed-point iterative algorithm. such a perspective has been investigated in [15].

Notice that in Lemma IV.2 and Theorem IV.1, the effect of two weights $l_{\mathbf{R}_k}$ and $l_{\hat{\mathbf{C}}_k}$ on the Kalman gain and error estimate depends on the ratio λ_k. In addition to this approach, we propose a alternative approach to avoid the problem of solving a fixed point iteration. Recall the definitions of $l_{\mathbf{R}_k}$ and $l_{\mathbf{P}_{k|k-1}}$ in (23). Notice that $l_{\mathbf{R}_k}$ and $l_{\mathbf{P}_{k|k-1}}$ all contains variable $\mathbf{x}_k$. To avoid computing a fixed point solution, we hope to use approximation for $l_{\mathbf{R}_k}$ and $l_{\mathbf{P}_{k|k-1}}$. More specifically, we use $\hat{\mathbf{x}}_{k|k-1}$ to approximate $\mathbf{x}_k$ contained in $\mathcal{G}_\sigma\left(\|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}\|_{\mathbf{P}_{k|k-1}}\right)$ and $\mathcal{G}_\sigma\left(\|\mathbf{y}_k - \mathbf{h}(\mathbf{x}_k)\|_{\mathbf{V}_k\mathbf{R}_k\mathbf{V}_k^\top}\right)$. They are listed by

$$\begin{aligned}\hat{l}_{\mathbf{P}_{k|k-1}} &= \mathcal{G}_\sigma\left(\left\|\hat{\mathbf{x}}_{k|k-1} - \hat{\mathbf{x}}_{k|k-1}\right\|_{\mathbf{P}_{k|k-1}}\right) = 1 \approx l_{\mathbf{P}_{k|k-1}} \\ \hat{l}_{\mathbf{R}_k} &= \mathcal{G}_\sigma\left(\left\|\mathbf{y}_k - \mathbf{h}\left(\hat{\mathbf{x}}_{k|k-1}\right)\right\|_{\mathbf{R}_k}\right) \approx l_{\mathbf{R}_k}.\end{aligned} \tag{63}$$

Similarly, for MCIC-MEKF-2, we still use $\hat{\mathbf{x}}_{k|k-1}$ to approximate $\mathbf{x}_k$ contained in $\mathbf{C}_k^x$ and $\mathbf{C}_k^y$, i.e,

$$\begin{aligned}\hat{\mathbf{e}}_k^x &= \left(\mathbf{B}_{k|k-1}^p\right)^{-1}\left(\hat{\mathbf{x}}_{k|k-1} - \hat{\mathbf{x}}_{k|k-1}\right) \approx \mathbf{e}_k^x \\ \hat{\mathbf{e}}_k^y &= \left(\mathbf{B}_k^r\right)^{-1}\left(\mathbf{y}_k - \mathbf{h}(\hat{\mathbf{x}}_{k|k-1})\right) \approx \mathbf{e}_k^y.\end{aligned} \tag{64}$$

Besides, we choose $\mathbf{A}_k\boldsymbol{\alpha}_k = \mathbf{0}_n$ and $\mathbf{B}_k\boldsymbol{\beta}_k = \mathbf{0}_m$ for algorithm implementation. We shall study how to estimate $\mathbf{A}_k, \boldsymbol{\alpha}_k, \mathbf{B}_k, \boldsymbol{\beta}_k$ in an online manner to improve algorithm performance in future work. Here we summarise the steps of MCIC-MEKF-1 and MCIC-MEKF-2 in Algorithm 1.

Algorithm 1 MCIC-MEKF

1: **Intitialization.** Start with initial posterior mean $\hat{\mathbf{x}}_{0|0}$ and posterior covariance $\mathbf{P}_{0|0}$.
2: **for** $k = 1, 2, \ldots, T$ **do**
3: Compute prior mean $\hat{\mathbf{x}}_{k|k-1}$ and prior covariance $\mathbf{P}_{k|k-1}$ via Eq.(16) and Eq.(17), respectively.
4: The following update steps correspond to MCIC-MEKF-1 and MCIC-MEKF-2, respectively.
- (MCIC-MEKF-1) Compute posterior mean $\hat{\mathbf{x}}_{k|k}$ and posterior covariance $\mathbf{P}_{k|k}$ via Eq.(20) and Eq.(24), respectively, where $\mathbf{K}_k^\star$ defined in Eq.(21) is computed using approximations (63).
- (MCIC-MEKF-2) Compute posterior mean $\hat{\mathbf{x}}_{k|k}$ and posterior covariance $\mathbf{P}_{k|k}$ via Eq.(27) and Eq.(31), respectively, where $\mathbf{K}_k^*$ defined in Eq.(28) is computed using approximations (64).

5: **end for**

TABLE I
COMPUTATIONAL COMPLEXITY COMPARISON

Algorithm	Complexity
EKF	$8n^3 + 10n^2m - n^2 + 6nm^2 - n + O(m^3) + O(n^2)$
MC-EKF	$S_{\text{EKF}} + (2T+8)n^3 + (4T+6)n^2m + \cdots + TO(n^3) + 2TO(m^3)$
MCIC-MEKF-1	$S_{\text{EKF}} + 2(m^2+n^2) + m + n$
MCIC-MEKF-2	$S_{\text{EKF}} + 4n^3 + O(n^3) + O(m^3)$

Note: S_{EKF} represents the complexity of the standard EKF algorithm. The notation $O(\cdot)$ denotes the big-O computational complexity notation.

V. EXPERIMENTS

We verify MCIC-MEKF-1 and MCIC-MEKF-2's robustness against EKF, RS-EKF [17], MCC-EKF [26], and robust EKF (REKF) [18] on a non-linear model with non-Gaussian noises. The RS-EKF was configured with a risk-sensitive parameter μ set to 0.1, while the REKF utilized an initial scaling parameter θ of 0.01. For the REKF, an approximate method was employed to iteratively solve the γ equation at each step. Performance is assessed by estimation errors and computation speeds. We employ the Root Mean Square Error (RMSE) and its Average (ARMSE) across 100 Monte Carlo simulations to quantify the accuracy of filters. We consider a 2D constant acceleration model for tracking a target. The state, denoted as $\mathbf{x}_k$, includes the position and velocity in the x and y directions ($\mathbf{x}_k = [x, y, v_x, v_y]$). The radar is positioned at $[p_x, p_y] = [-100, -100]$, and the time sampling interval is $T_0 = 1$.

$$\begin{aligned}\mathbf{x}_k &= \begin{bmatrix} 1 & 0 & T_0 & 0 \\ 0 & 1 & 0 & T_0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_{k-1} + \mathbf{w}_k \\ \mathbf{y}_k &= \begin{bmatrix} \sqrt{(x-p_x)^2 + (y-p_y)^2} \\ \arctan\frac{y-p_y}{x-p_x} \end{bmatrix} + \mathbf{v}_k,\end{aligned} \tag{65}$$

where initial state is $\mathbf{x}_0 = [-40, 10, 3, 1]$ and its covariance is $\mathbf{P}_0 = \text{diag}(4, 4, 0.01, 0.01)$. The nominal state noise covariance and observation noise are given as follows.

$$\mathbf{Q} = 0.04 \begin{bmatrix} \frac{T_0^3}{3} & 0 & \frac{T_0^2}{2} & 0 \\ 0 & \frac{T_0^3}{3} & 0 & \frac{T_0^2}{2} \\ \frac{T_0^2}{2} & 0 & T_0 & 0 \\ 0 & \frac{T_0^2}{2} & 0 & T_0 \end{bmatrix}, \mathbf{R} = \text{diag}(0.2^2, 0.015^2) \tag{66}$$

We consider the radar with Gaussian mixture observation noise, which is given by

$$\mathbf{v}_k \sim 0.9\,\mathcal{N}(0, \text{diag}(0.2^2, 0.015^2)) + 0.1\,\mathcal{N}(0, \text{diag}(5^2, 0.75^2)). \tag{67}$$

We analyze 100 time steps using RMSE and ARMSE metrics: RMSE$(pos)_k$, ARMSE(pos), RMSE$(vel)_k$, and ARMSE(vel). Results are shown in Figs.1 and Table II. EKF performs poorly under non-Gaussian noise, while REKF and SR-EKF show improvement but underperforms compared to robust filters. MCC-EKF, MCIC-MEKF-1, and MCIC-MEKF-2 effectively mitigate non-Gaussian noise effects, with MCIC-MEKF variants achieving the best accuracy. The shaded regions in Figs.1 indicate that MCIC-MEKF-1 and MCIC-MEKF-2 have smaller estimation STD, suggesting consistent estimation.

TABLE II
THE PERFORMANCE COMPARISONS BETWEEN EKF, REKF, SR-EKF, MCC-EKF, MCIC-MEKF-1 AND MCIC-MEKF-2 WITH $\sigma = 2$ ON TARGET TRACKING EXAMPLE.

Methods	ARMSE(pos)	ARMSE(vel)	Inference time [sec]
EKF	29.0929	5.0089	0.0026
REKF	15.9128	3.8178	0.0273
SR-EKF	7.3046	1.0210	0.0030
MCC-EKF ($\sigma = 2$)	4.1012	0.5930	0.0286
MCIC-MEKF-1 ($\sigma = 2$)	3.5836	**0.5382**	0.0057
MCIC-MEKF-2 ($\sigma = 2$)	**3.4808**	0.5562	0.0056

VI. CONCLUSION

This note introduces MCIC-MEKF, a robust modified Kalman-type filter based on MCIC. We provide theoreti-

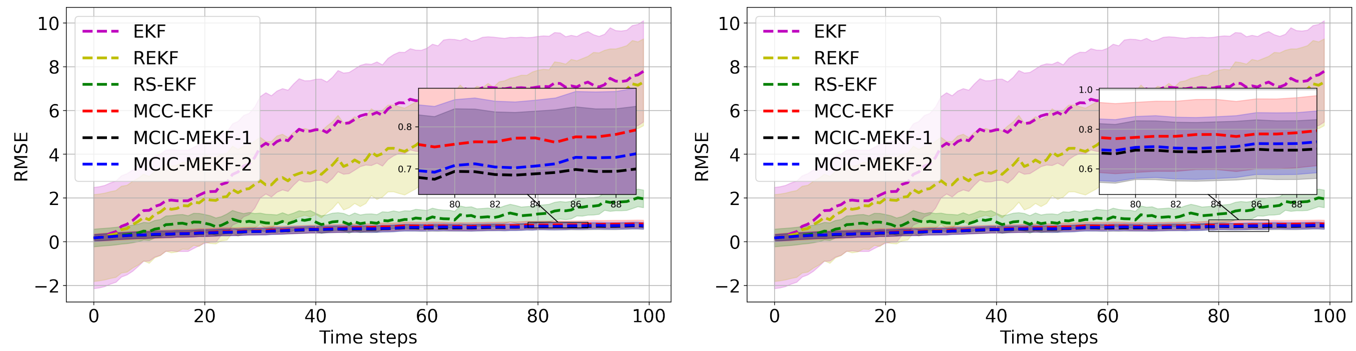

Fig. 1. RMSE comparisons over time steps for a target tracking example (the estimates of position and velocity on the left and right, respectively)

cal analysis of its stochastic stability, proving exponentially bounded posterior estimation errors under certain conditions. Our approach includes an efficient approximation for adaptive Kalman gain. Simulations demonstrate the method's robustness against non-Gaussian noises with large outliers, outperforming various benchmark filters in target tracking state-space model.

Appendix A
Proof of Lemmas

A. Proof of Lemma IV.1

Taking mathematical expectation for both sides of Eq.(39), we have

$$\mathbb{E}\left[\mathcal{V}_{k+1}\left(\mathbf{x}_{k+1}\right)\right] \leq \left(1-\alpha_k\right)\mathbb{E}\left[\mathcal{V}_k\left(\mathbf{x}_k\right)\right]+\mu_k. \tag{68}$$

Then iterate the above equation, we have

$$\mathbb{E}\left[\mathcal{V}_{k+1}\left(\mathbf{x}_{k+1}\right)\right] \leq \prod_{i=0}^{k}\left(1-\alpha_i\right)\mathbb{E}\left[\mathbf{V}_0\left(\mathbf{x}_0\right)\right]+\sum_{i=0}^{k}\mu_i\left(1-\alpha_i\right)^i. \tag{69}$$

Using Eq.(38), one can conclude that

$$\underline{v}\mathbb{E}\left[\|\mathbf{x}_{k+1}\|^2\right] \leq \mathbb{E}\left[\mathcal{V}_{k+1}\left(\mathbf{x}_{k+1}\right)\right], \bar{v}\mathbb{E}\left[\|\mathbf{x}_0\|^2\right] \geq \mathbb{E}\left[\mathbf{V}_0\left(\mathbf{x}_0\right)\right] \tag{70}$$

It follows that

$$\underline{v}\mathbb{E}\left[\|\mathbf{x}_{k+1}\|^2\right] \leq \prod_{i=0}^{k}\left(1-\alpha_i\right)\bar{v}\mathbb{E}\left[\|\mathbf{x}_0\|^2\right]+\sum_{i=0}^{k}\mu_i\left(1-\alpha_i\right)^i. \tag{71}$$

Hence we can conclude Eq.(40). Note that $1-\alpha_i \leq 1-\alpha$, it follows that

$$\sum_{i=0}^{k-1}\mu_i(1-\alpha_i)^i \leq \mu\sum_{i=0}^{k-1}(1-\alpha_i)^i \leq \mu\sum_{i=0}^{\infty}(1-\alpha_i)^i = \frac{\mu}{\alpha}, \tag{72}$$

which implies Eq.(41).

B. Proof of Eq.(36)

Based the property of trace operator and in view of Eq.(35), we have

$$\mathbb{E}\left[\mathbf{w}_k^\top\mathbf{w}_k\right] = \mathbb{E}\left[\mathrm{Tr}\left(\mathbf{w}_k^\top\mathbf{w}_k\right)\right] = \mathrm{Tr}\left(\mathbb{E}\left[\mathbf{w}_k\mathbf{w}_k^\top\right]\right) \leq n\bar{q}. \tag{73}$$

Let $\kappa_{\mathbf{w}} = n\bar{q}$. Similarly, we can choose $\kappa_{\mathbf{v}} = m\bar{r}$.

C. Proof of Lemma IV.2

In terms of the definition of Kalman gain $\mathbf{K}_k^\star$ in (21) and considering our assumption, we have

$$\|\mathbf{K}_k^\star\| \geq \frac{\underline{p}l_{\mathbf{P}_{k|k-1}}\underline{h}}{\bar{h}^2\frac{\bar{p}}{l_{\mathbf{P}_{k|k-1}}}+\frac{\|\mathbf{V}_k\|\bar{r}}{l_{\mathbf{R}_k}}} = \frac{\underline{p}\underline{h}}{\bar{h}^2\bar{p}+\lambda_k\|\mathbf{V}_k\|\bar{r}}. \tag{74}$$

Similarly, we have

$$\|\mathbf{K}_k^\star\| \leq \frac{\bar{p}l_{\mathbf{P}_{k|k-1}}\bar{h}}{\underline{h}^2\frac{\underline{p}}{l_{\mathbf{P}_{k|k-1}}}+\frac{\|\mathbf{V}_k\|\underline{r}}{l_{\mathbf{R}_k}}} = \frac{\bar{p}\bar{h}}{\underline{h}^2\underline{p}+\lambda_k\|\mathbf{V}_k\|\underline{r}}. \tag{75}$$

D. Proof of Lemma IV.3

Note that

$$\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\mathbf{V}_k\mathbf{R}_k\mathbf{V}_k^\top\left(\mathbf{K}_k^\star\right)^\top\hat{\mathbf{F}}_{k+1}^\top \geq 0. \tag{76}$$

Then we have

$$\begin{aligned}\mathbf{P}_{k+1|k} \geq &\left(\hat{\mathbf{F}}_{k+1}-\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\hat{\mathbf{H}}_k\right)\mathbf{P}_{k|k-1}\\ &\times\left(\hat{\mathbf{F}}_{k+1}-\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\hat{\mathbf{H}}_k\right)^\top+\mathbf{W}_k\mathbf{Q}_k\mathbf{W}_k^\top.\end{aligned} \tag{77}$$

In view of assumption IV.2, $\hat{\mathbf{F}}_{k+1}$ is non-singular. Then use Matrix inversion Lemma and (21), we have

$$\begin{aligned}&\left(\hat{\mathbf{F}}_{k+1}\right)^{-1}\left(\hat{\mathbf{F}}_{k+1}-\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\hat{\mathbf{H}}_k\right)\mathbf{P}_{k|k-1}\\ &=\mathbf{I}_n-\mathbf{P}_{k|k-1}^\star\hat{\mathbf{H}}_k^\top\left(\hat{\mathbf{H}}_k\mathbf{P}_{k|k-1}^\star\hat{\mathbf{H}}_k^\top+\mathbf{R}_k^\star\right)^{-1}\hat{\mathbf{H}}_k\mathbf{P}_{k|k-1}\\ &=\left(\mathbf{P}_{k|k-1}^{-1}+\hat{\mathbf{H}}_k\left(\frac{\mathbf{R}_{k|k-1}^\star}{l_{\mathbf{P}_{k|k-1}}}\right)^{-1}\hat{\mathbf{H}}_k^\top\right)^{-1}.\end{aligned} \tag{78}$$

It follows that $\left(\hat{\mathbf{F}}_{k+1}-\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\hat{\mathbf{H}}_k\right)^{-1}$ exists. Then we have

$$\begin{aligned}\mathbf{P}_{k+1|k} \geq &\left(\hat{\mathbf{F}}_{k+1}-\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\hat{\mathbf{H}}_k\right)\left(\mathbf{P}_{k|k-1}\right.\\ &+\left(\hat{\mathbf{F}}_{k+1}-\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\hat{\mathbf{H}}_k\right)^{-1}\mathbf{W}_k\mathbf{Q}_k\mathbf{W}_k^\top\\ &\times\left.\left(\hat{\mathbf{F}}_{k+1}-\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\hat{\mathbf{H}}_k\right)^{-\top}\right)\left(\hat{\mathbf{F}}_{k+1}-\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\hat{\mathbf{H}}_k\right)^\top.\end{aligned} \tag{79}$$

In view of lemma IV.2, we have $\|\mathbf{K}_k^\star\| \leq \bar{\iota}_{\mathbf{K}_k^\star}$. Then bounds of $\hat{\mathbf{F}}_{k+1}$, $\hat{\mathbf{H}}_k$, $\mathbf{W}_k$, $\mathbf{Q}_k$, Eq.(79) imply that

$$\mathbf{P}_{k+1|k} \geq \left(\hat{\mathbf{F}}_{k+1} - \hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\hat{\mathbf{H}}_k\right)\left(\mathbf{P}_{k|k-1} + \frac{\underline{q}}{\left(\bar{f} + \bar{f}\bar{\iota}_{\mathbf{K}_k^\star}\bar{h}\right)^2}\mathbf{I}_n\right) \times \left(\hat{\mathbf{F}}_{k+1} - \hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\hat{\mathbf{H}}_k\right)^\top. \tag{80}$$

Note that $\mathbf{P}_{k|k-1} \geq \underline{p}\mathbf{I}_n$ and $\left(\hat{\mathbf{F}}_{k+1} - \hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\mathbf{H}_k\right)$ is non-singular, we can take the inverse of both sides for (80). Then by multiplying $\left(\hat{\mathbf{F}}_{k+1} - \hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\mathbf{H}_k\right)^\top$ and $\left(\hat{\mathbf{F}}_{k+1} - \hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\mathbf{H}_k\right)$ from left and right respectively, one can conclude that

$$\left(\hat{\mathbf{F}}_{k+1} - \hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\mathbf{H}_k\right)^\top \mathbf{P}_{k+1|k}^{-1}\left(\hat{\mathbf{F}}_{k+1} - \hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\mathbf{H}_k\right) \leq \left(1 + \frac{\underline{q}}{\bar{p}\left(\bar{f} + \bar{f}\bar{\iota}_{\mathbf{K}_k^\star}\bar{h}\right)^2}\right)^{-1}\mathbf{P}_{k|k-1}^{-1}, \tag{81}$$

which is our desired inequality (43) with $1 - \alpha_k = \frac{1}{1+\frac{\underline{q}}{\bar{p}\left(\bar{f}+\bar{f}\bar{\iota}_{\mathbf{K}_k^\star}\bar{h}\right)^2}}$.

E. Proof of Lemma IV.4

Note that

$$\mathbb{E}\left[\left(\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\mathbf{V}_k\mathbf{v}_k\right)^\top \mathbf{P}_{k+1|k}^{-1}\left(\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\mathbf{V}_k\mathbf{v}_k\right) \mid \tilde{\mathbf{e}}_{k|k-1}\right] = \mathbb{E}\left[\mathbf{v}_k^\top\left(\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\mathbf{V}_k\right)^\top \mathbf{P}_{k+1|k}^{-1}\left(\hat{\mathbf{F}}_{k+1}\mathbf{K}_k^\star\mathbf{V}_k\right)\mathbf{v}_k\right] \leq \frac{\left(\bar{f}\bar{\iota}_{\mathbf{K}_k^\star}\|\mathbf{V}_k\|\right)^2\kappa_{\mathbf{v}}}{\underline{p}}. \tag{82}$$

References

[1] S. Särkkä, *Bayesian filtering and smoothing*. Cambridge university press, 2013, no. 3.

[2] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.

[3] D. Simon, *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.

[4] S. J. Julier and J. K. Uhlmann, "New extension of the kalman filter to nonlinear systems," in *Signal processing, sensor fusion, and target recognition VI*, vol. 3068. International Society for Optics and Photonics, 1997, pp. 182–193.

[5] I. Arasaratnam and S. Haykin, "Cubature kalman filters," *IEEE Transactions on automatic control*, vol. 54, no. 6, pp. 1254–1269, 2009.

[6] N. J. Gordon and D. J. Salmond, "Novel approach to nonlinear/non-gaussian bayesian state estimation," *IEE Proceedings. Part F*, vol. 140, no. 2, pp. P.107–113, 1993.

[7] J. Kang, X. Chen, Y. Tao, and S. S.-T. Yau, "Optimal transportation particle filter for linear filtering systems with correlated noises," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 6, pp. 5190–5203, 2022.

[8] A. Doucet, N. De Freitas, and N. Gordon, "An introduction to sequential monte carlo methods," in *Sequential Monte Carlo methods in practice*. Springer, 2001, pp. 3–14.

[9] A. Taghvaei and P. G. Mehta, "A survey of feedback particle filter and related controlled interacting particle systems (cips)," *arXiv preprint arXiv:2301.00935*, 2023.

[10] P. J. Huber, "Robust estimation of a location parameter," in *Breakthroughs in statistics*. Springer, 1992, pp. 492–518.

[11] C. D. Karlgaard, "Nonlinear regression huber–kalman filtering and fixed-interval smoothing," *Journal of guidance, control, and dynamics*, vol. 38, no. 2, pp. 322–330, 2015.

[12] J. C. Principe, *Information theoretic learning: Renyi's entropy and kernel perspectives*. Springer Science & Business Media, 2010.

[13] Y. Huang, Y. Zhang, N. Li, Z. Wu, and J. A. Chambers, "A novel robust student's t-based kalman filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 3, pp. 1545–1554, 2017.

[14] W. Liu, P. P. Pokharel, and J. C. Principe, "Correntropy: Properties and applications in non-gaussian signal processing," *IEEE Transactions on signal processing*, vol. 55, no. 11, pp. 5286–5298, 2007.

[15] B. Chen, X. Liu, H. Zhao, and J. C. Principe, "Maximum correntropy kalman filter," *Automatica*, vol. 76, pp. 70–77, 2017.

[16] G. Wang, Y. Zhang, and X. Wang, "Iterated maximum correntropy unscented kalman filters for non-gaussian systems," *Signal Processing*, vol. 163, pp. 87–94, 2019.

[17] J. L. Speyer, C.-H. Fan, and R. N. Banavar, "Optimal stochastic estimation with exponential cost criteria," in *[1992] Proceedings of the 31st IEEE Conference on Decision and Control*. IEEE, 1992, pp. 2293–2299.

[18] P. Chauchat, J. Vilà-Valls, and E. Chaumette, "Robust information filtering under model mismatch for large-scale dynamic systems," *IEEE Control Systems Letters*, vol. 6, pp. 818–823, 2021.

[19] A. Longhini, M. Perbellini, S. Gottardi, S. Yi, H. Liu, and M. Zorzi, "Learning the tuned liquid damper dynamics by means of a robust ekf," in *2021 American Control Conference (ACC)*. IEEE, 2021, pp. 60–65.

[20] R. K. Tiwari and S. Bhaumik, "Risk sensitive filtering with randomly delayed measurements," *Automatica*, vol. 142, p. 110409, 2022.

[21] Y. Feng, X. Huang, L. Shi, Y. Yang, J. A. Suykens *et al.*, "Learning with the maximum correntropy criterion induced losses for regression." *J. Mach. Learn. Res.*, vol. 16, no. 30, pp. 993–1034, 2015.

[22] G. Calafiore, "Reliable localization using set-valued nonlinear filters," *IEEE Transactions on systems, man, and cybernetics-part A: systems and humans*, vol. 35, no. 2, pp. 189–197, 2005.

[23] K. Reif, S. Gunther, E. Yaz, and R. Unbehauen, "Stochastic stability of the discrete-time extended kalman filter," *IEEE Transactions on Automatic control*, vol. 44, no. 4, pp. 714–728, 1999.

[24] S. Kluge, K. Reif, and M. Brokate, "Stochastic stability of the extended kalman filter with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 55, no. 2, pp. 514–518, 2010.

[25] L. Li and Y. Xia, "Stochastic stability of the unscented kalman filter with intermittent observations," *Automatica*, vol. 48, no. 5, pp. 978–981, 2012.

[26] X. Liu, H. Qu, J. Zhao, and B. Chen, "Extended kalman filter under maximum correntropy criterion," in *2016 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2016, pp. 1733–1737.

[27] Y. Huang, Y. Zhang, Y. Zhao, P. Shi, and J. A. Chambers, "A novel outlier-robust kalman filtering framework based on statistical similarity measure," *IEEE Transactions on Automatic Control*, vol. 66, no. 6, pp. 2677–2692, 2020.

[28] K. Xiong, H. Zhang, and C. Chan, "Performance evaluation of ukf-based nonlinear filtering," *Automatica*, vol. 42, no. 2, pp. 261–270, 2006.

[29] X. Kai, C. Wei, and L. Liu, "Robust extended kalman filtering for nonlinear systems with stochastic uncertainties," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 40, no. 2, pp. 399–405, 2009.

[30] J. Hu, Z. Wang, H. Gao, and L. K. Stergioulas, "Extended kalman filtering with stochastic nonlinearities and multiple missing measurements," *Automatica*, vol. 48, no. 9, pp. 2007–2015, 2012.

[31] B. D. Anderson and J. B. Moore, "Detectability and stabilizability of time-varying discrete-time linear systems," *SIAM Journal on Control and Optimization*, vol. 19, no. 1, pp. 20–32, 1981.

[32] A. H. Jazwinski, *Stochastic processes and filtering theory*. Courier Corporation, 2007.