Outlier-Robust Iterative Extended Kalman Filtering

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

***Abstract*—In this letter, we develop OR-IEKF which is a novel outlier-robust iterative extended Kalman filtering (IEKF) framework based on nonlinear regression formulation of update step. A new Kalman-type update step with reweighted prediction covariance and reweighted observation noise covariance are produced under the OR-IEKF framework, which could cut off the large outliers in observations causing by unknown outlier noises. By using various robust cost functions to solve such special nonlinear regression problems, we derive three algorithms. The performances of these new filters are evaluated in a nonlinear system simulation study.**

***Index Terms*—Heavy-tailed noise, iterative extended Kalman filter, Kalman filter, outliers.**

I. Introduction

The Kalman filtering (KF) [1] is a well-known state estimator for linear Gaussian system, which is derived based on minimum mean-square-error (MMSE) criterion [2]. For nonlinear system with Gaussian noises, people develop different Kalman-type filtering algorithms such as extended Kalman filter (EKF) [3], iterative extended Kalman filter (IEKF) [4], unscented Kalman filter (UKF) [5]. They are widely used in a large number of real-world applications of engineering [6], [7]. However, these filters are derived by MMSE criterion making them very sensitive to heavy-tailed observation noises, which are frequently encountered in practical applications. For instance, noises in target tracking, audio communication, and power systems are frequently non-Gaussian heavy-tailed, and are brought on by state and observation outliers from outside interference or defective sensors [8]. Therefore, the above mentioned KF and their extensions performances can be much degraded in such cases.

How to improve state estimate robustness against non-Gaussian heavy-tailed observation noises has been the focus of many recent works. One perspective is to model distribution of heavy-tailed observation noises by pre-selected distribution family. For instance, the authors in [9] used variational Bayes to do inference while assuming t-distributed observation noise. Nevertheless, these filters are not useful for handling multiple dimensions system. The sequential Monte Carlo sampling based methods (or say Particle filter) [10] can also be used to handle the non-Gaussian noises. However, in order to mimic non-Gaussian state and observation noises, their performances largely rely on the pre-selected distribution family. Unfortunately, choosing the appropriate distribution family to characterize the above-mentioned unknown non-Gaussian noises is quite challenging. Using inaccurate or even erroneous noise distributions will drastically reduce the estimation performance. Moreover, high computational complexity of these methods makes them significantly more difficult to implement in real time. More practical approach is regression-based filtering. The resulting algorithm exhibits a similar framework of Kalman-type update step but with a modified prediction covariance and modified observation noise covariance. For example, a classical robust regression technique called M-estimation [11] has been used to develop robust Kalman filer called Huber-based KF [12]. It is derived by Huber function, which is a cost function combining l^1 norm and l^2 norm. Besides Huber's robust statistics, information theoretic quantities (e.g., correntropy [13], Renyi's entropy [14]) can also be used as a robust cost for estimation problems to derive new filters [15], [16], [17]. However, these filters are based on the robust linear regression formulation. Therefore, the natural generalization in nonlinear regression became the main motivation of this letter. More explicitly, we note that, for nonlinear system the update step of Kalman-type filtering can be transformed as a special nonlinear regression problem. This gives rise to an optimization framework to modify its update step. Then the iteratively reweighted least squares (IRLS) [18] can be applied to solve such nonlinear regression problems. With different robust cost function, we get new iterative extended Kalman filtering (IEKF) framework as outlier-robust IEKF (OR-IEKF). We apply Huber function [11], t-likelihood [19] and correntropy and correntropy to derive three new algorithms. It is necessary to note that the robustness of these cost functions is controlled by specific tuning factors. And these derived filters and effects of different tuning factors are evaluated in a nonlinear system with presence of Gaussian mixture noise with large outliers and multi-modal noise.

II. Outlier-Robust Filtering Problems

Throughout this article, 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:

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

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

where $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ and $\mathbf{h} : \mathbb{R}^n \to \mathbb{R}^m$ are nonlinear functions called state function and observation function, respectively. State noise $\mathbf{w}_k$ and observation noise $\mathbf{v}_k$ are uncorrelated multivariate

Manuscript received 20 March 2023; revised 27 May 2023; accepted 29 May 2023. Date of publication 13 June 2023; date of current version 29 June 2023. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 11961141005, in part by Tsinghua University Startup Fund, and in part by Tsinghua University Education Foundation Fund under Grant 042202008. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Donatella Darsena.
(Corresponding author: Stephen Shing-Toung Yau.)

Yangtianze Tao is with the Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China (e-mail: tytz19@mails.tsinghua.edu.cn).

Stephen Shing-Toung Yau is with the Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China, and also with the Yanqi Lake Beijing Institute of Mathematical Sciences and Applications, Beijing 101400, China (e-mail: yau@uic.edu).
Digital Object Identifier 10.1109/LSP.2023.3285118

Gaussian with zero means and nominal covariance matrices $\mathbf{Q}_k \in \mathbb{R}^{n\times n}$ and $\mathbf{R}_k \in \mathbb{R}^{m\times m}$, respectively. In what follows, we assume that the real distribution of observation noise is unknown to us. And the real distribution of observation noise is $(1-\epsilon)\,\mathcal{N}(0,\mathbf{R}_k) + \epsilon\,\mathcal{S}(0,\mathbf{S}_k)$ rather than $\mathcal{N}(0,\mathbf{R}_k)$, where $0<\epsilon\ll 1$ is the unknown probability, and $\mathcal{S}(0,\mathbf{S}_k)$ is an arbitrary unknown distribution with large covariance $\mathbf{S}_k$. Here $\mathbf{R}_k$ is known to us, so it is called the nominal covariance matrix. Let $\mathbf{y}_{1:k}$ denote the σ-algebra generated by noisy observations $\{\mathbf{y}_1,\ldots,\mathbf{y}_k\}$ induced by unknown large outliers. The outlier-robust filtering problem refers to solving the following conditional estimation problem:

$$\hat{\boldsymbol{\varphi}}_k = \arg\min_{\boldsymbol{\varphi}_k} \mathbb{E}\left[\|\boldsymbol{\varphi}_k - \varphi(\mathbf{x}_k)\|^2 \mid \mathbf{y}_{1:k}\right], \tag{2}$$

where $\varphi(\mathbf{x})$ is a function of estimation interest (e.g., $\varphi(\mathbf{x})=\mathbf{x}$ or $\varphi(\mathbf{x})=\mathbf{x}\mathbf{x}^\top$).

III. Nonlinear Regression Form and Robust Optimization Framework

To facilitate the subsequent discussion, $\hat{\mathbf{x}}_{k|k-1}$ and $\hat{\mathbf{x}}_{k|k}$ will be denoted as desired estimation means of prediction at time steps $k-1$ and k respectively. $\mathbf{P}_{k|k-1}$ and $\mathbf{P}_{k|k}$ will be corresponding prediction covariances respectively.

A. Nonlinear Regression Form

Here we shall first illustrate how to view the update step of extended Kalman filtering as a nonlinear regression problem. Assume that we have obtained $\hat{\mathbf{x}}_{k|k-1}$ and observation $\mathbf{y}_k$. Then let us first consider the augmented model, which is given by

$$\begin{bmatrix}\hat{\mathbf{x}}_{k|k-1}\\ \mathbf{y}_k\end{bmatrix} = \begin{bmatrix}\mathbf{x}_k\\ \mathbf{h}(\mathbf{x}_k)\end{bmatrix} + \mathbf{V}_k, \tag{3}$$

where $\mathbf{V}_k$ is given by

$$\mathbf{V}_k = \begin{bmatrix}-(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})\\ \mathbf{v}_k\end{bmatrix}. \tag{4}$$

Then it is easy to see that

$$\mathbb{E}\left[\mathbf{V}_k\mathbf{V}_k^\top\right] = \begin{bmatrix}\mathbf{P}_{k|k-1} & 0\\ 0 & \mathbf{R}_k\end{bmatrix} = \mathbf{B}_k\mathbf{B}_k^\top, \tag{5}$$

with

$$\mathbf{B}_k = \begin{bmatrix}\mathbf{B}^p_{k|k-1} & 0\\ 0 & \mathbf{B}^r_k\end{bmatrix}, \tag{6}$$

where $\mathbf{B}^p_{k|k-1}$ and $\mathbf{B}^r_k$ are Cholesky decompositions of $\mathbf{P}_{k|k-1}$ and $\mathbf{R}_k$, respectively. Left multiplying both sides of (3) by $\mathbf{B}_k^{-1}$, we obtain

$$\mathbf{d}_k = \mathbf{m}_k(\mathbf{x}_k) + \mathbf{e}_k, \tag{7}$$

where $\mathbf{d}_k = \begin{bmatrix}(\mathbf{B}^p_{k|k-1})^{-1}\hat{\mathbf{x}}_{k|k-1}\\ (\mathbf{B}^r_k)^{-1}\mathbf{y}_k\end{bmatrix}$, $\mathbf{m}_k(\mathbf{x}_k) = \begin{bmatrix}(\mathbf{B}^p_{k|k-1})^{-1}\mathbf{x}_k\\ (\mathbf{B}^r_k)^{-1}\mathbf{h}(\mathbf{x}_k)\end{bmatrix}$. Note that $\mathbf{e}_k = \mathbf{B}_k^{-1}\mathbf{V}_k$, which implies $\mathbb{E}[\mathbf{e}_k\mathbf{e}_k^\top] = \mathbf{I}_{n+m}$. Hence the residual error $\mathbf{e}_k$ is white noise, which makes (7) become a nonlinear regression function.

B. Robust Optimization Framework

With the help of regression function (7), we can formulate a optimization-based filtering update step. It is given by

$$\hat{\mathbf{x}}_{k|k} = \arg\min_{\mathbf{x}_k} \mathcal{L}\left(\mathbf{x}_k\right), \tag{8}$$

where the cost function $\mathcal{L}(\cdot)$ is regression-induced in nature. It is given by

$$\mathcal{L}\left(\mathbf{x}_k\right) = \sum_{i=1}^{n+m} \rho\left(e_{k,i}\right), \tag{9}$$

with $e_{k,i}$ is i-th component of the residual vector $\mathbf{e}_k$. ρ is a robust cost function that is used to cut off the outliers. Note that for $\mathcal{L}(\mathbf{x}_k)$, we have

$$\nabla_{\mathbf{x}_k}\mathcal{L}\left(\mathbf{x}_k\right) = \sum_{i=1}^{m+n} \frac{\partial\rho\left(e_{k,i}\right)}{\partial e_{k,i}}\frac{\partial e_{k,i}}{\partial \mathbf{x}_k}. \tag{10}$$

Let we consider the following diagonal matrices,

$$\begin{aligned}\boldsymbol{\Psi}_x\left(\mathbf{e}_k\right) &= \mathrm{diag}\left[\psi(e_{k,1}),\psi(e_{k,2}),\ldots,\psi(e_{k,n})\right]\\ \boldsymbol{\Psi}_y\left(\mathbf{e}_k\right) &= \mathrm{diag}\left[\psi(e_{k,n+1}),\psi(e_{k,2}),\ldots,\psi(e_{k,m+n})\right],\end{aligned} \tag{11}$$

and

$$\boldsymbol{\Psi}\left(\mathbf{e}_k\right) = \begin{bmatrix}\boldsymbol{\Psi}_x\left(\mathbf{e}_k\right) & 0\\ 0 & \boldsymbol{\Psi}_y\left(\mathbf{e}_k\right)\end{bmatrix} \tag{12}$$

with weight function $\psi(e_{k,i}) = \frac{\partial\rho(e_{k,i})}{\partial e_{k,i}}/e_{k,i}$. Notice that $\mathbf{d}_k$ does not depend on $\mathbf{x}_k$. We can denote

$$\begin{aligned}\mathbf{M}(\mathbf{x}_k) = \nabla_{\mathbf{x}_k}\mathbf{e}_k &= \frac{\partial\left(\mathbf{d}_k - \mathbf{m}_k(\mathbf{x}_k)\right)}{\partial\mathbf{x}_k} = -\frac{\partial\mathbf{m}_k(\mathbf{x}_k)}{\partial\mathbf{x}_k}\\ \nabla_{\mathbf{e}_k}\mathcal{L}\left(\mathbf{x}_k\right) &= \boldsymbol{\Psi}\left(\mathbf{e}_k\right)\left(\mathbf{d}_k - \mathbf{m}_k(\mathbf{x}_k)\right).\end{aligned} \tag{13}$$

Then the gradient (10) can be rewritten by

$$\begin{aligned}\nabla_{\mathbf{x}_k}\mathcal{L}\left(\mathbf{x}_k\right) &= \left(\frac{\partial\left(\mathbf{d}_k - \mathbf{m}_k(\mathbf{x}_k)\right)}{\partial\mathbf{x}_k}\right)^\top \boldsymbol{\Psi}\left(\mathbf{e}_k\right)\left(\mathbf{d}_k - \mathbf{m}_k(\mathbf{x}_k)\right)\\ &= \mathbf{M}(\mathbf{x}_k)^\top\nabla_{\mathbf{e}_k}\mathcal{L}\left(\mathbf{x}_k\right).\end{aligned} \tag{14}$$

The solution of $\nabla_{\mathbf{x}_k}\mathcal{L}(\mathbf{x}_k)=0$ can be derived using iteratively reweighted least squares (IRLS) [18], which is of the form

$$\begin{aligned}\mathbf{x}_k^{(j+1)} &= \mathbf{x}_k^{(j)} - \left(\mathbf{M}(\mathbf{x}_k^{(j)})^\top\boldsymbol{\Psi}(\mathbf{e}_k^{(j)})\mathbf{M}(\mathbf{x}_k^{(j)})\right)^{-1}\\ &\quad\times\mathbf{M}(\mathbf{x}_k^{(j)})^\top\boldsymbol{\Psi}(\mathbf{e}_k^{(j)})\left(\mathbf{d}_k - \mathbf{m}_k(\mathbf{x}_k^{(j)})\right),\end{aligned} \tag{15}$$

where the superscript (j) refers to the iteration index.

Remark III.1: Equation (15) is similar to the derivation process of standard IEKF in [4], where the IEKF is equivalent to Gauss-Newton method from an optimization perspective.

IV. Iterative Outlier-Robust Extended Kalman Filtering

In this section, the novel OR-IEKF framework is presented. Its prediction step is the same with the common EKF, i.e.,

$$\begin{aligned}\hat{\mathbf{x}}_{k|k-1} &= \mathbf{f}\left(\hat{\mathbf{x}}_{k-1|k-1}\right)\\ \mathbf{P}_{k|k-1} &= \mathbf{F}_k\mathbf{P}_{k-1|k-1}\mathbf{F}_k^\top + \mathbf{Q}_k,\end{aligned} \tag{16}$$

Algorithm 1: OR-IEKF.

```
Input: f(·), h(·), Q_k, R_k, ρ(·), ε, α
Output: x̂_{k|k} for k = 1, 2, ..., N
Intitialization. Start with initial filtering mean x̂_{0|0}
and filtering covariance P_{0|0}.
for k = 1, 2, ..., N do
   Compute prior mean x̂_{k|k-1} and prior covariance
   P_{k|k-1} via (16).;
   Let x_k^{(0)} = x̂_{k|k-1} and compute x_k^{(1)} via (17).
   while ||x_k^{(j+1)} - x_k^{(j)}|| / ||x_k^{(j)}|| > ε do
      Compute iterative solution x_k^{(j+1)} via (19) with
      corresponding update of K_k^{(j+1)} in (18).
   end while
   Update filtering covariance matrix P_{k|k} using (21).
end for
```

where $\mathbf{F}_k = \frac{\partial \mathbf{f}}{\partial \mathbf{x}_k}|_{\mathbf{x}_k=\hat{\mathbf{x}}_{k-1|k-1}}$. The update step is given in (15). With initial point $\mathbf{x}_k^{(0)} = \hat{\mathbf{x}}_{k|k-1}$, we can further simplify (15) by using the matrix inversion lemma [20]. The result is

$$\mathbf{x}_k^{(j+1)} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k^{(j)} \left[\mathbf{y}_k - \mathbf{h}\left(\mathbf{x}_k^{(j)}\right) - \mathbf{H}\left(\mathbf{x}_k^{(j)}\right)\left(\hat{\mathbf{x}}_{k|k-1} - \mathbf{x}_k^{(j)}\right)\right], \quad (17)$$

where

$$\mathbf{K}_k^{(j)} = \mathbf{P}_{k|k-1}^{(j)} \mathbf{H}\left(\mathbf{x}_k^{(j)}\right)^\top \left(\mathbf{H}\left(\mathbf{x}_k^{(j)}\right) \times \mathbf{P}_{k|k-1}^{(j)} \mathbf{H}\left(\mathbf{x}_k^{(j)}\right)^\top + \mathbf{R}_k^{(j)}\right), \quad (18)$$

with $\mathbf{P}_{k|k-1}^{(j)} = \mathbf{B}_{k|k-1}^p (\boldsymbol{\Psi}_x(\mathbf{e}_k^{(j)}))^{-1} \mathbf{B}_{k|k-1}^p$, $\mathbf{R}_k^{(j)} = \mathbf{B}_k^r (\boldsymbol{\Psi}_y(\mathbf{e}_k^{(j)}))^{-1} \mathbf{B}_k^r$ and $\mathbf{H}(\mathbf{x}_k^{(j)}) = \frac{\partial \mathbf{h}}{\partial \mathbf{x}_k}|_{\mathbf{x}_k = \mathbf{x}_k^{(j)}}$. For numerical stability, when $k \geq 1$, we introduce the step parameter $0 < \alpha \leq 1$ for modification. This means that we first use (17) to compute $\mathbf{x}_k^{(1)}$ without using step parameter α. For $k \geq 1$, we shall use step parameter α with initial guess $\mathbf{x}_k^{(1)}$, which means that (17) can be rewritten as

$$\mathbf{x}_k^{(j+1)} = \hat{\mathbf{x}}_{k|k-1} + \alpha \mathbf{p}_k^{(j)}, \quad (19)$$

where the direction $\mathbf{p}_k^{(j)}$ is given by

$$\mathbf{p}_k^{(j)} = \mathbf{K}_k^{(j)} \left[\mathbf{y}_k - \mathbf{h}\left(\mathbf{x}_k^{(j)}\right) - \mathbf{H}\left(\mathbf{x}_k^{(j)}\right)\left(\hat{\mathbf{x}}_{k|k-1} - \mathbf{x}_k^{(j)}\right)\right]. \quad (20)$$

After iterations, we shall obtain converged solutions $\mathbf{x}_k$ and $\mathbf{e}_k$. The recursion of filtering covariance can be computed by

$$\mathbf{P}_{k|k} = \left(\mathbf{I}_n - \mathbf{K}_k \mathbf{H}\left(\mathbf{x}_k\right)\right) \mathbf{B}_{k|k-1}^p \left(\boldsymbol{\Psi}_x(\mathbf{e}_k)\right)^{-1} \mathbf{B}_{k|k-1}^p. \quad (21)$$

Here we summarise the steps of OR-IEKF in Algorithm 1.

Remark IV.1: Note that, if $\boldsymbol{\Psi}(\mathbf{e}_k) = \mathbf{I}_{n+m}$, the above iterations reduces to the IEKF solution. In addition, when the observation is linear, the solution reduces to the standard linear KF update.

TABLE I
THREE ROBUST COST FUNCTION EXAMPLES

Name	ρ_τ	w_τ
Huber	$\begin{cases} \frac{1}{2}x^2 & \text{if } \lvert x\rvert < \tau, \\ \tau\lvert x\rvert - \frac{1}{2}\tau^2 & \text{if } \lvert x\rvert \geq \tau. \end{cases}$	$\begin{cases} 1 & \text{if } \lvert x\rvert < \tau, \\ \tau \times \frac{\text{sgn}(x)}{x} & \text{if } \lvert x\rvert \geq \tau. \end{cases}$
t-likelihood	$\frac{\tau^2}{2}\log(1 + \frac{x^2}{\tau^2})$	$\frac{1}{1+(\frac{x}{\tau})^2}$
Correntropy-induced cost	$\frac{\tau^2}{2}\left(1 - \exp\left(-\left(\frac{x}{\tau}\right)^2\right)\right)$	$\exp\left(-\left(\frac{x}{\tau}\right)^2\right)$

A. Robust Cost Functions

Note that in (12), the item $\psi(e_{k,i}) = \frac{\partial \rho(e_{k,i})}{\partial e_{k,i}} / e_{k,i}$ plays an important role in cutting of large outliers. We shall call it the weight function. The scalar function $\rho(x)$ is called robust cost function. Let us denote ρ, ψ by ρ_τ, w_τ, where $\tau \in \mathbb{R}$ is a tuning factor. Here we consider three examples, which are given in Table I. They correspond to the three new iterative filtering algorithms; namely, Huber-IEKF, t-likelihood-IEKF and Correntropy-IEKF.

Remark IV.2: While there are several other robust cost functions available, such as Andrew, Biweight, Hampel, Logistic, and more [21], [22], this letter focuses on developing an algorithmic framework. Our future work will concentrate on developing IEKF that integrates these robust cost functions.

V. SIMULATION

In this section, we shall evaluate our proposed filters in a nonlinear system benchmark. We abbreviate our three new filters as Huber-IEKF, Tlike-IEKF and Corr-IEKF. We shall compare them with EKF, IEKF, MCC-EKF [15] and MEE-EKF [23], where MCC-EKF and MEE-EKF follow the nonlinear extensions proposed in [24]. The iterative convergence threshold is $\epsilon = 0.001$ and step parameter is $\alpha = 0.5$. The time steps and Monte Carlo runs are set by $N = 1000$ and $M = 100$, respectively. The comparison metrics are RMSE and ARMSE, which are defined as

$$\text{RMSE}(k) = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(\|\mathbf{x}_k^{(i)} - \hat{\mathbf{x}}_k^{(i)}\|\right)^2}$$

$$\text{ARMSE} = \frac{1}{N}\sum_{k=1}^{N}\text{RMSE}(k). \quad (22)$$

Here $\text{RMSE}(k)$ is the RMSE at time step k, and $\mathbf{x}_k^{(i)}$, $\hat{\mathbf{x}}_k^{(i)}$ are the true and estimated state in i-th experiment at time step k.

A. Nonlinear Model

We consider a nonlinear system, which is given as follows:

$$\mathbf{x}_k = \left(\mathbf{I}_2 + \kappa_1 \begin{bmatrix} -1 & 0.1 \\ 0.1 & -1 \end{bmatrix}\right) \mathbf{x}_{k-1} + \kappa_2 \cos(\mathbf{x}_{k-1}) + \mathbf{w}_k,$$
$$\mathbf{y}_k = \mathbf{x}_k + \sin(\mathbf{x}_k) + \mathbf{v}_k, \quad (23)$$

where $\mathbf{x}_0 \sim \mathcal{N}(0, 25\mathbf{I}_2)$, $\kappa_1 = \kappa_2 = 0.1$ are constants controlling the state dynamics. Let $\mathbf{Q} = \mathbf{R} = \mathbf{I}_2$ be nominal state covariance and nominal observation covariance. We shall look at two different non-Gaussian noises.

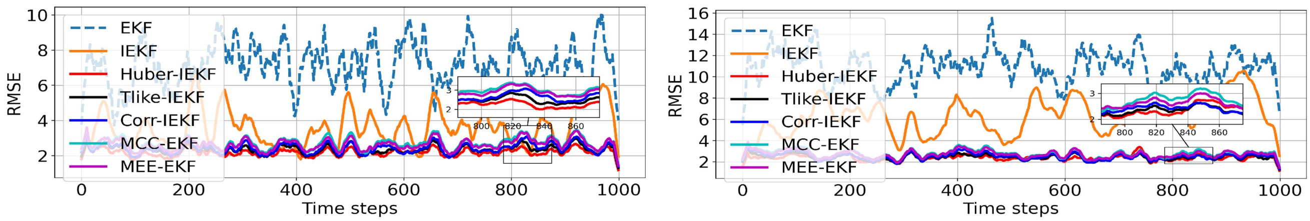

Fig. 1. RMSE comparisons over time steps for different noise cases. The left and right in the figure above correspond to case A, case B, respectively.

TABLE II
PERFORMANCE COMPARISONS WITH $\mathbf{R}_k$ CHOICES FOR NOISE CASE A

Method	ARMSE	CPU Time	ARMSE	CPU Time	ARMSE	CPU Time
	$\mathbf{R}_k$		$\hat{\mathbf{R}}_k$		$\mathbf{S}_k$	
EKF	7.1902	0.0292	4.5997	0.0289	**3.2794**	0.0285
IEKF	3.6231	0.3119	2.6333	0.2259	3.3248	0.1098
MCC-EKF	2.7141	0.8294	2.7043	0.5179	3.3563	0.3231
MEE-EKF	2.6681	1.3791	2.6653	1.0358	3.2839	0.6356
Huber-IEKF	**2.1948**	0.7834	**2.4681**	0.5927	3.3473	0.3210
Tlike-IEKF	2.3530	0.6377	2.5988	0.5344	3.3478	0.3325
Corr-IEKF	2.4332	0.6864	2.4882	0.5640	3.3473	0.3231

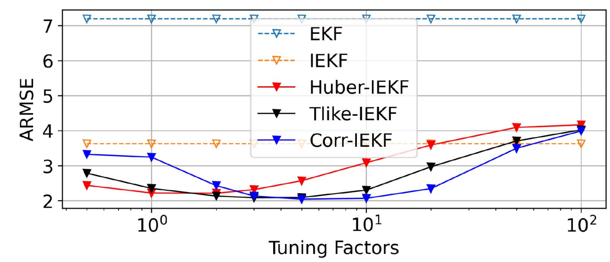

Fig. 2. ARMSE comparisons with different tuning factors for noise case A.

- *Case A:* The real state noise is Gaussian and real observation noise is a mixture of Gaussian, i.e.,

$$\mathbf{w}_k \sim \mathcal{N}(0, \mathbf{Q}),\ \mathbf{v}_k \sim 0.9\,\mathcal{N}(0, \mathbf{R}) + 0.1\,\mathcal{N}(0, 1000\mathbf{R}), \tag{24}$$

- *Case B:* The real state noise is Gaussian, and the real observation noise has multi-modal distribution, i.e.,

$$\mathbf{w}_k \sim \mathcal{N}(0, \mathbf{Q}),\ \mathbf{v}_k \sim 0.8\,\mathcal{N}(0, \mathbf{R}) + 0.1\,\mathcal{N}(-10, 1000\mathbf{R}) + 0.1\,\mathcal{N}(10, 1000\mathbf{R}), \tag{25}$$

For the above two noise cases, we present RMSEs of EKF, IEKF, MCC-EKF, MEE-EKF, Huber-IEKF, Tlike-IEKF and Corr-IEKF over time steps for different noise cases in Fig. 1. To better exhibit the RMSEs, the RMSEs are smoothed using a moving average method with span of 20 steps. Here we set tuning factor $\tau = 1.345, 1, 2$ for these robust functions, respectively. Note that for both cases, our proposed OR-IEKF algorithms outperform the EKF, IEKF, MCC-EKF and MEE-EKF. These results demonstrate their effectiveness in the presence of various non-Gaussian observation noises. Moreover, as the OR-IEKF algorithm has shown superior performance compared to the MCC-EKF and MEE-EKF, the effectiveness of the robust algorithm using the IEKF framework instead of the conventional EKF framework has been highlighted. We found that these three filters maintain very robust and stable estimates that change over time, unlike EKF and IEKF, which are greatly affected by large outliers, thus making their estimation performance degraded severely.

In [25], the selection of $\mathbf{R}_k$ is highlighted as a key factor in the KF performance. To evaluate our algorithm's performance with different $\mathbf{R}_k$ choices, we compare results using three choices: $\hat{\mathbf{R}}_k$, the true covariance of observation noise; $\mathbf{R}_k$; and $\mathbf{S}_k$. While we may not have access to $\hat{\mathbf{R}}_k$ and $\mathbf{S}_k$, comparing results using these values can help identify an optimal choice for $\mathbf{R}_k$. These results are listed in Table II. It should be noted that when using $\mathbf{R}_k$ and $\hat{\mathbf{R}}_k$, the OR-IEKF algorithms outperform other algorithms and have similar robust performance. This implies that the robust estimate of OR-IEKF algorithms can be achieved simply by using $\mathbf{R}_k$. Moreover,all algorithms exhibit similar performance when using $\mathbf{S}_k$ and their performance are worse than those of the former two cases. These results demonstrate that $\mathbf{R}_k$ is a good choice for implementing robust filtering algorithms in the presence of unknown large outliers. Moreover, we found that their CPU time is only roughly twice that of IEKF, demonstrating that good robustness may be attained with only a little increase in computational cost. Additionally, since they are all Kalman-type, real-time performance can be guaranteed. Besides, to see the effect of different tuning factor τ. We conduct experiments for case A with tuning factors ranging in $[0.5, 1, 2, 3, 5, 10, 20, 50, 100]$. These results are presented in Fig. 2. Notice that as the tuning factor increases, their robustness first decreases, then increases, and finally degenerates to the scale close to the robustness of IEKF. Furthermore, observe that, for Table I, Huber-IEKF performs better than Tlike-IEKF and Corr-IEKF, but as the tuning factor increases, Tlike-IEKF and Corr-IEKF perform better than Huber-IEKF. This demonstrates that by selecting appropriate tuning factors, higher robustness can be attained in practical situations.

VI. CONCLUSION

In this letter, we introduce OR-IEKF, a novel outlier-robust IEKF framework built on a nonlinear regression formulation. In order to solve such unique nonlinear regression problems with various robust cost functions, we introduce a new Kalman-type update step with reweighted prediction covariance and reweighted observation noise covariance under the OR-IEKF framework. This step is controlled by a predetermined tuning factor related to the given robust cost function. For three distinct robust cost functions, we derive three different algorithms. The efficiency of the proposed algorithms has been validated through the simulation study. Moreover, the effects of different tuning factors are also evaluated.

REFERENCES

[1] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Trans. ASME–J. Basic Eng.*, vol. 82, pp. 35–45, 1960.

[2] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. Chelmsford, MA, USA: Courier Corporation, 2007.

[3] D. Simon, *Optimal State Estimation: Kalman, H. Infinity, and Nonlinear Approaches*. Hoboken, NJ, USA: Wiley, 2006.

[4] B. M. Bell and F. W. Cathey, “The iterated Kalman filter update as a gauss-newton method,” *IEEE Trans. Autom. Control*, vol. 38, no. 2, pp. 294–297, Feb. 1993.

[5] S. J. Julier and J. K. Uhlmann, “New extension of the Kalman filter to nonlinear systems,” in *Proc. SPIE*, vol. 3068, 1997, pp. 182–193.

[6] S. Chen, “Kalman filter for robot vision: A survey,” *IEEE Trans. Ind. Electron.*, vol. 59, no. 11, pp. 4409–4420, Nov. 2012.

[7] K. Law, A. Stuart, and K. Zygalakis, *Data Assimilation*, vol. 214. Cham, Switzerland: Springer, 2015.

[8] Y. Huang, Y. Zhang, N. Li, and J. Chambers, “A robust Gaussian approximate fixed-interval smoother for nonlinear systems with heavy-tailed process and measurement noises,” *IEEE Signal Process. Lett.*, vol. 23, no. 4, pp. 468–472, Apr. 2016.

[9] Y. Huang, Y. Zhang, N. Li, Z. Wu, and J. A. Chambers, “A novel robust student’s T-based Kalman filter,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 3, pp. 1545–1554, Jun. 2017.

[10] N. J. Gordon and D. J. Salmond, “Novel approach to nonlinear/non-Gaussian Bayesian state estimation,” *IEE Proc. Part F*, vol. 140, no. 2, pp. 107–113, 1993.

[11] P. J. Huber, “Robust estimation of a location parameter,” in *Breakthroughs in Statistics*. Berlin, Germany: Springer, 1992, pp. 492–518.

[12] P. Petrus, “Robust Huber adaptive filter,” *IEEE Trans. Signal Process.*, vol. 47, no. 4, pp. 1129–1133, Apr. 1999.

[13] W. Liu, P. P. Pokharel, and J. C. Principe, “Correntropy: Properties and applications in non-Gaussian signal processing,” *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5286–5298, Nov. 2007.

[14] J. C. Principe, *Information Theoretic Learning: Renyi’s Entropy and Kernel Perspectives*. Berlin, Germany: Springer, 2010.

[15] X. Liu, H. Qu, J. Zhao, and B. Chen, “Extended Kalman filter under maximum correntropy criterion,” in *Proc. IEEE Int. Joint Conf. Neural Netw.*, 2016, pp. 1733–1737.

[16] B. Chen, X. Liu, H. Zhao, and J. C. Principe, “Maximum correntropy Kalman filter,” *Automatica*, vol. 76, pp. 70–77, 2017.

[17] B. Chen, L. Dang, Y. Gu, N. Zheng, and J. C. Príncipe, “Minimum error entropy Kalman filter,” *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 51, no. 9, pp. 5819–5829, Sep. 2021.

[18] R. Wolke and H. Schwetlick, “Iteratively reweighted least squares: Algorithms, convergence analysis, and numerical comparisons,” *SIAM J. Sci. Stat. Comput.*, vol. 9, no. 5, pp. 907–921, 1988.

[19] M. J. Hinich and P. P. Talwar, “A simple method for robust regression,” *J. Amer. Stat. Assoc.*, vol. 70, no. 349, pp. 113–119, 1975.

[20] W. W. Hager, “Updating the inverse of a matrix,” *SIAM Rev.*, vol. 31, no. 2, pp. 221–239, 1989.

[21] R. V. Hogg, “Statistical robustness: One view of its use in applications today,” *Amer. Statistician*, vol. 33, no. 3, pp. 108–115, 1979.

[22] C. Liu, G. Wang, X. Guan, and C. Huang, “Robust M-estimation-based maximum correntropy Kalman filter,” *ISA Trans.*, vol. 136, pp. 198–209, 2022.

[23] G. Wang, B. Chen, X. Yang, B. Peng, and Z. Feng, “Numerically stable minimum error entropy Kalman filter,” *Signal Process.*, vol. 181, 2021, Art. no. 107914.

[24] L. Dang, B. Chen, Y. Xia, J. Lan, and M. Liu, “Dual extended Kalman filter under minimum error entropy with fiducial points,” *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 52, no. 12, pp. 7588–7599, Dec. 2022.

[25] X. Fan, G. Wang, J. Han, and Y. Wang, “A background-impulse Kalman filter with non-Gaussian measurement noises,” *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 53, no. 4, pp. 2434–2443, Apr. 2023.