An Outlier Robust Finite Impulse Response Filter With Maximum Correntropy

YANDA GUO, XUYOU LI[ID], AND QINGWEN MENG[ID]
College of Automation, Harbin Engineering University, Harbin 150001, China

Corresponding author: Xuyou Li (lixuyou@hrbeu.edu.cn)

ABSTRACT Non-Gaussian noise is common in industrial applications, and it is a severe challenge to existing state estimators. In this paper, a novel robust maximum correntropy finite impulse response (MCFIR) filter is proposed to deal with the state estimation problem in the linear state-space system corrupted by outliers. The filter operates as a finite memory form, and thus it obtains superior immunity to noise statistics and process uncertainties than existing Kalman-like robust filters. Gaussian correntropy is adopted to generate a new cost function, which improves the filter robustness to outlier interference. We derive an unbiased MCFIR filter that ignores noise statistics and propose an improvement bias-constrained MCFIR filter to achieve better estimate accuracy. To improve the filtering performance degradation caused by improper kernel size, an adaptive kernel size algorithm is further proposed, which adjusts the bandwidth within a specific range adaptively and achieves significant improvement in the MCFIR filter. An illustrative example based on moving target tracking is presented to evaluate the performance of the proposed filter, and simulation results confirmed that the MCFIR filter obtained superior immunity to outliers than the existing robust filters.

INDEX TERMS Kalman filter, finite impulse response, maximum correntropy criterion, state estimation.

I. INTRODUCTION

The estimation problem has been one of the essential subjects from industrial applications to research areas, including signal processing, navigation, target tracking, etc. A Kalman filter (KF) is one of the most popular filters for the linear Gaussian state-space model due to its simple structure and excellent performance. It is optimal for the minimum mean square error (MMSE) under the ideal conditions [1]. To solve the filtering problem in a nonlinear system, researchers have made many improvements based on the KF's structure: Extended KF (EKF) based on Taylor transform [2], unscented KF (UKF), and Cubature KF (CKF) based on sigma point transform [3], [4], particle filters (PF) based on a Monte Carlo sample, and other variants of Kalman filters [5] have been applied into nonlinear systems and achieve excellent performances.

The preconditions for the reliable operation of a KF are accurately defined noise covariance matrix and initial values, and a Gaussian distribution assumption is necessary to keep

the KF optimal. Once the dynamic processes are disturbed by non-Gaussian noise like outliers, the KF cannot maintain optimal estimation. A Student-t distribution-based KF (STF) is an effective cope method for non-Gaussian noise or impulsive noise [6], [7]. However, an STF suffers from high-order statistics loss due to moment matching and improper statistical parameters [8], [9]. In addition to the STF, robust filters based on m-estimate theory have been researched extensively over the years [10]. One of the most widely used m-estimators based on the Huber loss function combines the minimum $\ell 1$ and $\ell 2$ -norm. The Huber loss function has been applied in many applications and has high reliability [11]. Combining the covariance matrix rescaled operation, the Huber KF (HKF) derived in prior work [12], [13] can be taken as a generalized maximum likelihood estimator, eliminating the outliers effectively after the measurement update. Similarly, this method can also be applied in nonlinear systems after deformation [14], [15]. Recently, correntropy based filters have introduced the concept of information entropy as a local similarity measure, which is robust to non-Gaussian noise. The correntropy contains the high-order statistic characteristics, which can obtain significant performance

improvement than other m-estimators. For deterministic systems, correntropy has been applied in state-space recursive least squares [16]. Regarding state-space systems with process noise, the maximum correntropy KF (MCKF) was proposed [17] and obtains better performance than many existing robust KFs. Because the analytical solutions of correntropy filters are difficult to find, the optimal solution needs to be obtained by an iterative method like fixed-point iteration [17], Gauss-Newton method in [18] and half-quadratic minimization method in [19], the convergence condition of the fixed-point iteration was further analyzed [20]. Several variants of MCKF have also been developed for nonlinear systems as [18], [19], [21] and practical applications like INS/GNSS navigation systems in [22]. Because the estimation accuracy of the MCKF is primarily related to the kernel parameters, an improper kernel size may even cause filtering divergence. Therefore, an adaptive filter with variable kernel size in [23], and MCKF with adaptive kernel size in [24] were proposed to solve the problem. The concept of mixture correntropy was recently proposed in [25], and an outlier-robust KF with mixture correntropy derived in [26] further improve the kernel size adaptive problem. Although MCKF has achieved significant improvement on robustness to non-Gaussian noise, the filter cannot avoid being limited by the Kalman filter structure, incorrect initial state or model uncertainties, and other unfavorable factors are likely to corrupt the stability of MCKF.

From the view of Bayesian estimation, KF can be taken as an infinite impulse filter (IIR) filter, which utilizes all measurements on the infinite interval to obtain the current estimation [27]. In addition to the IIR filters, finite impulse response (FIR) filters for state-space systems have been developed extensively over the years. The FIR filters have several inherent properties, such as bounded input/output (BIBO) stability and excellent robustness to temporary process uncertainties and noise statistics [28]. In practical applications, such as research working in statistical signal processing and control, FIR filters might own better practical performance than KF. The FIR filters can be divided into various forms according to their characteristics. Among these, the unbiased FIR (UFIR) filter that ignores statistics and initial conditions [29], the optimal FIR filter with an unbiased constraint [30], and the MLFIR filter based on maximum likelihood estimate with Gaussian noise assumption are several representative FIR filters that have been proposed [31]. Like STF, a robust student-t distribution FIR filter was also proposed to deal with the heavy-tailed noise filtering problem [32]. For FIR filters, the horizon length determines the filtering performance, which needs to be appropriately preset or adjusted adaptively [33], [34]. Compared with typical IIR filters like Kalman filter, FIR filters avoid the error accumulation and own better robust inherent properties to process uncertainties and noise statistics. However, the implementation of FIR filters depends on reliable measurement vectors, for measurement contaminated by non-Gaussian noise,

a convenient robust FIR filter remains rarely studied, which motivates the present work.

This work's main contributions can be summarized as follows: 1) We propose a novel robust maximum correntropy FIR filter, called the MCFIR filter. The filter operates as a finite memory form and takes the correntropy as the local similarity measure to attenuate outliers' interference. 2) This paper further modifies the unbiased MCFIR filter into a bias-constrained form, which improves the unbiased MCFIR filter's performance by compensating process bias error accumulation. 3) By combining the unique processing mode of the MCFIR filter, we propose an adaptive kernel bandwidth algorithm to prevent filtering performance degradation caused by improper kernel bandwidth parameters.

The rest of the paper is organized as follows: In Section 2, we briefly introduce the concept of correntropy and an unbiased FIR filter. In Section 3, we derive the batch form and iterative form of the unbiased MCFIR filter. Moreover, an adaptive kernel bandwidth method is proposed. In Section 4, we modify the unbiased MCFIR filter into a bias-constrained MCFIR filter. Simulation results in Section 5 demonstrate the performance of the MCFIR filter. Finally, conclusions are drawn in Section 6.

II. PRELIMINARIES AND PROBLEM STATEMENT
A. CONCEPT OF CORRENTROPY
The correntropy is a useful local similarity measure tool for state estimate in heavy-tailed noise conditions. Given two random variables $X, Y \in R$ with joint distribution function $F_{XY}(x, y)$, correntropy can be defined by

$$V(X, Y) = \mathrm{E}[\kappa(x, y)] = \int \kappa(x, y) dF_{XY}(x, y), \quad (1)$$

where $\kappa(x, y)$ is a shift-invariant Mercer Kernel. In this work, the general Gaussian kernel function is given by

$$\kappa(x, y) = G_\sigma(e) = \exp(-\frac{e^2}{2\sigma^2}), \quad (2)$$

where $e = x - y$, and $\sigma > 0$ denotes the kernel bandwidth, the correntropy function with Gaussian kernel is insensitive to large errors and hence can be used as a robust cost function in estimation problems.

Taking the Taylor series expansion of the Gaussian kernel yields

$$V(X, Y) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n \sigma^{2n} n!} E\left[(X - Y)^{2n}\right], \quad (3)$$

the correntropy is a reweighted sum of all even order moments of the error variable $X - Y$, the kernel bandwidth appears as a parameter to weight the second-order and higher-order moments. With a large kernel bandwidth (compared to the dynamic range of the data), the correntropy (3) dominated by the second-order moment.

For a sequence of error data $\mathbf{e}$, we can estimate the correntropy using a sample mean estimator as

$$\hat{V}(X, Y) = \frac{1}{N} \sum_{i=1}^{N} G_\sigma(\mathbf{e}(i)), \qquad (4)$$

where $\mathbf{e}(i) = \mathbf{x}(i) - \mathbf{y}(i)$, $\mathbf{x}$ and $\mathbf{y}$ represent two vectors with N samples. The robust correntropy cost function could be established based on (4), which could effectively eliminate the adverse effects of noise outliers during filtering.

B. EXTENDED STATE SPACE MODEL AND UNBIASED FIR FILTER

Consider the standard linear state-space system equations as

$$\mathbf{x}_k = \mathbf{A}_k \mathbf{x}_{k-1} + \mathbf{w}_k, \qquad (5)$$
$$\mathbf{y}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k, \qquad (6)$$

where k is a discrete time index, $\mathbf{x}_k \in \mathbb{R}^n$ is the system state vector at discrete-time index k, $\mathbf{A}_k \in \mathbb{R}^{n \times n}$ is the state transition matrix, $\mathbf{y}_k \in \mathbb{R}^m$ is the measurement vector, and $\mathbf{C}_k \in \mathbb{R}^{m \times n}$ is the measurement matrix. The process noise vectors $\mathbf{w}_k \in \mathbb{R}^n$ and the measurement noise vectors $\mathbf{v}_k \in \mathbb{R}^m$ are zero mean with covariance matrix $\mathbf{Q}_k$ and $\mathbf{R}_k$. Assume the noise vectors $\mathbf{w}_k$ and $\mathbf{v}_k$ are uncorrelated at each step.

The FIR filters take N neighbor measurements to update the state estimation $\hat{\mathbf{x}}_k$, therefore, an extended state-space model implemented on the horizon ranging $[s, k]$, which includes N discrete steps from $s = k - N + 1$ to k, can be formulated as

$$\mathbf{X}_{k,s} = \mathbf{A}_{k,s} \mathbf{x}_s + \mathbf{B}_{k,s} \mathbf{W}_{k,s}, \qquad (7)$$
$$\mathbf{Y}_{k,s} = \mathbf{C}_{k,s} \mathbf{x}_s + \mathbf{H}_{k,s} \mathbf{W}_{k,s} + \mathbf{V}_{k,s}, \qquad (8)$$

where $\mathbf{X}_{k,s} = [\mathbf{x}_k^T \mathbf{x}_{k-1}^T \ldots \mathbf{x}_s^T]^T$, $\mathbf{Y}_{k,s} = [\mathbf{y}_k^T \mathbf{y}_{k-1}^T \ldots \mathbf{y}_s^T]^T$, $\mathbf{W}_{k,s} = [\mathbf{w}_k^T \mathbf{w}_{k-1}^T \ldots \mathbf{w}_s^T]^T$, $\mathbf{V}_{k,s} = [\mathbf{v}_k^T \mathbf{v}_{k-1}^T \ldots \mathbf{v}_s^T]^T$, the extended state and measurement vectors $\mathbf{X}_{k,s} \in \mathbb{R}^{nN \times 1}$, $\mathbf{Y}_{k,s} \in \mathbb{R}^{mN \times 1}$, and $\mathbf{W}_{k,s} \in \mathbb{R}^{nN \times 1}$, $\mathbf{V}_{k,s} \in \mathbb{R}^{mN \times 1}$, and the extended state-space matrix follows as

$$\mathbf{A}_{k,s} = [\mathcal{A}_{k,s}^T \mathcal{A}_{k-1,s}^T \cdots \mathbf{A}_{s+1}^T \mathbf{I}]^T, \qquad (9)$$

$$\mathbf{B}_{k,s} = \begin{bmatrix} \mathbf{I} & \mathbf{A}_k & \cdots & \mathcal{A}_{k,s+1} & \mathcal{A}_{k,s} \\ \mathbf{0} & \mathbf{I} & \cdots & \mathcal{A}_{k-1,s+1} & \mathcal{A}_{k-1,s} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I} & \mathbf{A}_{s+1} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{I} \end{bmatrix}, \qquad (10)$$

$$\mathbf{C}_{k,s} = \bar{\mathbf{C}}_{k,s} \mathbf{A}_{k,s}, \qquad (11)$$

then $\bar{\mathbf{C}}_{k,s} = \mathrm{diag}(\mathbf{C}_k \mathbf{C}_{k-1} \cdots \mathbf{C}_s)$, $\mathbf{H}_{k,s} = \bar{\mathbf{C}}_{k,s} \mathbf{B}_{k,s}$, and $\mathcal{A}_{j,i} = \mathbf{A}_j \mathbf{A}_{j-1} \ldots \mathbf{A}_{i+1}$, where $\mathbf{A}_{k,s} \in \mathbb{R}^{nN \times n}$, $\mathbf{B}_{k,s} \in \mathbb{R}^{nN \times nN}$, $\mathbf{C}_{k,s} \in \mathbb{R}^{mN \times n}$, $\bar{\mathbf{C}}_{k,s} \in \mathbb{R}^{mN \times nN}$.

On horizon $[s, k]$, the FIR filter using N past neighboring measurement points can be described as

$$\hat{\mathbf{x}}_k = \mathbf{K}_{k,s} \mathbf{Y}_{k,s}, \qquad (12)$$

where $\hat{\mathbf{x}}_k$ is the estimation of $\mathbf{x}_k$ determined by measurements from s to k, and $\mathbf{K}_{k,s}$ is the FIR filter gain given by the estimate principle.

According to [29], the estimation of the unbiased FIR filter (12) retains the unbiased constraint as

$$\mathrm{E}[\hat{\mathbf{x}}_k] = \mathrm{E}[\mathbf{x}_k], \qquad (13)$$

where $\mathrm{E}[\mathbf{x}_k]$ represent the expectation of $\mathbf{x}_k$. By substituting (12) into (13), then

$$\mathrm{E}[\mathbf{K}_{k,s} \mathbf{Y}_{k,s}] = \mathrm{E}[\mathbf{x}_k], \qquad (14)$$

and the unbiased constrained $\mathbf{K}_{k,s} \tilde{\mathbf{C}}_{k,s} = \mathbf{I}$ needs to be satisfied to ensure unbiasedness characteristic of $\hat{\mathbf{x}}_k$, which defines a extended measurement matrix $\tilde{\mathbf{C}}_{k,s}$ as

$$\tilde{\mathbf{C}}_{k,s} = \bar{\mathbf{C}}_{k,s} [\mathbf{I} \mathbf{A}_k^{-T} \cdots \mathcal{A}_{k,s+1}^{-T} \mathcal{A}_{k,s}^{-T}]^T, \qquad (15)$$

By multiplying the identity matrix $(\tilde{\mathbf{C}}_{k,s}^T \tilde{\mathbf{C}}_{k,s})^{-1} \tilde{\mathbf{C}}_{k,s}^T \tilde{\mathbf{C}}_{k,s}$ on both sides of the unbiased constraint (14), we can derive an unbiased gain as

$$\mathbf{K}_{k,s} = (\tilde{\mathbf{C}}_{k,s}^T \tilde{\mathbf{C}}_{k,s})^{-1} \tilde{\mathbf{C}}_{k,s}^T, \qquad (16)$$

where $\mathbf{G}_{k,s} = (\tilde{\mathbf{C}}_{k,s}^T \tilde{\mathbf{C}}_{k,s})^{-1}$ represents the generalized noise power gain (NPG), and $\mathbf{G}_{k,s}$ represents the ratio of the filter output noise variance to the input noise variance. Therefore, the unbiased estimation of $\mathbf{x}_k$ can be written in batch form as

$$\hat{\mathbf{x}}_k = (\tilde{\mathbf{C}}_{k,s}^T \tilde{\mathbf{C}}_{k,s})^{-1} \tilde{\mathbf{C}}_{k,s}^T \mathbf{Y}_{k,s}. \qquad (17)$$

The unbiased FIR filter is a type of one parameter filter that only horizon length N is used to adjust the filter properties. Different from the recursive ordinary least squares method derived for deterministic systems, the horizon length of the unbiased FIR filters cannot be infinitely increased due to process bias accumulation, which needs to be limited within an appropriate range to ensure reliable estimation of $\mathbf{x}_k$. The optimal horizon size can be determined by minimizing the trace of the mean square error matrix as summarized in [30].

The unbiased FIR filter ignores the statistics of noise and only obtains the unbiased estimate. It has been proved in prior work [27]–[29] that unbiased FIR filter owns robust properties to model uncertainties and noise statistics, due to its FIR structure avoid error accumulation. However, as the operation of the UFIR filter is similar to the ordinary least squares (OLS), without considering the effect of process noise, we can approximate that finite measurement samples assigned with equal weights, and unbiased FIR filter is susceptible to abnormal measurement noise interference especially outliers.

Therefore, the problems can be formulated as follow. We want to provide a feasible method to introduce the maximum correntropy method into unbiased FIR filter to further improve the filtering robustness to abnormal noises. Meanwhile, consider the bias accumulation problem in the unbiased FIR filter, we then solve the problem by modifying the unbiased MCFIR filter into a bias constrained form. Finally, just as the prior work [17] summarized, the kernel size is a key parameter in the correntropy function, which remains an unsolved problem to be set properly. We further work to explore an adaptive kernel size method to solve the problem.

III. UNBIASED MAXIMUM CORRENTROPY FIR FILTER
A. BATCH SOLUTION OF THE UNBIASED MCFIR FILTER

Considering the extended state-space model (7), $\mathbf{x}_k$ can be expressed by $\mathbf{x}_s$ as

$$\mathbf{x}_k = \mathcal{A}_{k,s}\mathbf{x}_s + \bar{\mathbf{B}}_{k,s}\mathbf{W}_{k,s}. \qquad (18)$$

where $\bar{\mathbf{B}}_{k,s}$ is the first vector row in $\mathbf{B}_{k,s}$, the unbiased FIR filter ignores the noise vectors and assumes that $\mathbf{x}_k \approx \mathcal{A}_{k,s}\mathbf{x}_s$ within the estimate window; therefore, the unbiased solution obtained by (17) can be regarded as a typical solution of a linear regression problem derived on the quadratic loss function

$$\hat{\mathbf{x}}_k = \arg\min_k (\left\| \mathbf{Y}_{k,s} - \tilde{\mathbf{C}}_{k,s}\mathbf{x}_k \right\|). \qquad (19)$$

Therefore, the unbiased FIR filter is susceptible to measurement outliers, and, to improve the robustness of unbiased FIR filter to outliers, a new cost function with respect to $\mathbf{x}_k$ based on the Gaussian correntropy function can be reformulated as

$$J_{\text{MCC}}(\mathbf{x}_k) = \sum_{i=s}^{k} G_\sigma\left(\mathbf{e}_i^{(1)}\right) + \cdots + G_\sigma\left(\mathbf{e}_i^{(m)}\right), \quad (20)$$

where $\mathbf{e}_i = \mathbf{R}_i^{-1/2}\left(\mathbf{y}_i - \tilde{\mathbf{C}}_{k,i}\mathbf{x}_k\right)$, $\mathbf{e}_i^{(m)}$ is the mth element of $\mathbf{e}_i$, nominal measurement noise covariance matrix $\mathbf{R}_i$ is used to normalized the residual error sequence.

The optimal solution of $\hat{x}_k$ is obtained by maximize the object function $J_{\text{MCC}}(\mathbf{x}_k)$ as

$$\partial J_{\text{MCC}}(\mathbf{x}_k)/\partial \mathbf{x}_k = \sum_{i=s}^{k}\sum_{j=1}^{m}\left[G_\sigma(\mathbf{e}_i^{(j)})\tilde{\mathbf{C}}_{k,i}^{(j)}\mathbf{e}_i^{(j)} \right] = 0, \qquad (21)$$

it follows easily that a robust solution of (19) is

$$\hat{\mathbf{x}}_k = (\tilde{\mathbf{C}}_{k,s}^T\mathbf{\Omega}_{k,s}\tilde{\mathbf{C}}_{k,s})^{-1}\tilde{\mathbf{C}}_{k,s}^T\mathbf{\Omega}_{k,s}\mathbf{Y}_{k,s}. \qquad (22)$$

Here, $\mathbf{\Omega}_{k,s} = \mathbf{\Theta}_{k,s}\text{diag}[\mathbf{I}_m\theta\mathbf{I}_m\cdots\theta^{N-2}\mathbf{I}_m\theta^{N-1}\mathbf{I}_m]$, where $\mathbf{\Theta}_{k,s}$ denote the correntropy coefficients correspond to $[s, k]$ measurement vectors as

$$\mathbf{\Theta}_{k,s} = \text{diag}[G_\sigma(\mathbf{e}_k); G_\sigma(\mathbf{e}_{k-1}); \cdots; G_\sigma(\mathbf{e}_s)]. \qquad (23)$$

Defining the forgetting factor as $\theta \leq 1$, for the non-deterministic systems, the bias error accumulation degrades the accuracy of the correntropy coefficients and the unbiased FIR filter's filtering performance. Adding an appropriate forgetting factor may improve filtering performance. Therefore, the robust unbiased gain of (22) can be rewritten as

$$\tilde{\mathbf{K}}_{k,s} = (\tilde{\mathbf{C}}_{k,s}^T\mathbf{\Omega}_{k,s}\tilde{\mathbf{C}}_{k,s})^{-1}\tilde{\mathbf{C}}_{k,s}^T\mathbf{\Omega}_{k,s}. \qquad (24)$$

Because (22) operates as a batch processing form, it may bring an excessive calculation burden. According to [29], appropriate deformations can be made to obtain a faster iterative Kalman-like unbiased MCFIR filter.

B. ITERATIVE SOLUTION OF THE UNBIASED MCFIR FILTER

Consider the batch form unbiased estimate gain, in which $\mathbf{\Omega}_{k,s}$ is a diagonal matrix represents the product of the forget factor matrix and the correntropy coefficient matrix.

To recursively obtain the correntropy function factor at each step, a backward recursion can be adopted. By setting an intermediate variable $\boldsymbol{\xi}$, we have $\boldsymbol{\xi}_k = \hat{x}_{k|k-1}$ and $\boldsymbol{\xi}_{i-1} = \mathbf{A_i}^{-1}\boldsymbol{\xi}_i$ with $i \in [s, k]$, and we can obtain

$$\mathbf{\Omega}_i = \theta^{k-i}\text{diag}[G_{\sigma_k}(\mathbf{e}_i)], \qquad (25)$$

where intermediate variable $\boldsymbol{\xi}_i$ is used to update $\mathbf{\Omega}_i$ by recursion.

According to [29], the iterative solution of the unbiased MCFIR filter can be formulated as a Kalman-like filter as follows:

$$\tilde{\mathbf{x}}_i = \tilde{\mathbf{x}}_i^- + \mathbf{K_i}(\mathbf{y}_i - \mathbf{C}_i\tilde{\mathbf{x}}_i^-), \qquad (26)$$

where the filter gain $\mathbf{K_i}$ is defined as

$$\mathbf{K_i} = \tilde{G}_i\mathbf{C_i^T}\mathbf{\Omega_i}. \qquad (27)$$

The noise power gain factor $\tilde{G}_i$ can be calculated iteratively as

$$\tilde{G}_i = (\mathbf{C}_i^T\mathbf{\Omega}_i\mathbf{C}_i + (\mathbf{A}_i\tilde{G}_{i-1}\mathbf{A}_i^T)^{-1})^{-1}. \qquad (28)$$

The detailed derivation can be referred to in Appendix A.

By taking $\hat{\mathbf{x}}_k = \tilde{\mathbf{x}}_i$ as the output if $i = k$, the initial iteration noise power gain $\tilde{G}_m$ is

$$\tilde{G}_m = (\tilde{\mathbf{C}}_{m,s}^T\mathbf{\Omega}_{m,s}\tilde{\mathbf{C}}_{m,s})^{-1}, \qquad (29)$$

and the initial state $\tilde{\mathbf{x}}_m$ is calculated by

$$\tilde{\mathbf{x}}_m = \tilde{\mathbf{K}}_{m,s}\mathbf{Y}_{m,s}, \qquad (30)$$

which can be calculated by small-batch form.

It should be noted that only $\tilde{\mathbf{x}}_k$ is taken as the output, and the intermediate estimation $\tilde{\mathbf{x}}_i$ is ignored. The estimate results of the iterative form and batch form are equivalent. The iterative form unbiased MCFIR filter operates as a flexible Kalman-like form, making it much easier to make various extensions than the batch solution. The iterative solution may lower the computational burden than the batch form because it avoids high-dimensional matrix calculation that requires more computational memory to obtain the result.

C. ADAPTIVE KERNEL BANDWIDTH ALGORITHM

The kernel bandwidth σ is an essential parameter for the proposed filter, which determines the correntropy function's performance to the outliers. It is difficult to preset an optimal kernel bandwidth for each step [23]. Nevertheless, for the unbiased MCFIR filter proposed in this paper, an adaptive kernel size within the appropriate range is available.

In this work, the problem is further analyzed, and a feasible adjustment method is proposed. Considering the filter's unique FIR processing method, we design an adaptive kernel bandwidth algorithm that realizes the adaptive adjustment of the kernel size during filtering.

In the estimate window $[k-N+1,k]$, the normalized residual sequence $\mathbf{e}_{k,s}$ within the estimate window is analyzed.

$$|\mathbf{e}_{k,s}| = [|\mathbf{e}_k|^T\, |\mathbf{e}_{k-1}|^T \cdots |\mathbf{e}_{k-N+1}|^T]^T. \qquad (31)$$

A new parameter γ is defined as

$$\gamma = |\mathbf{e}_{\mathrm{med}} - \mathbf{e}_{\mathrm{min}}|/|\mathbf{e}_k - \mathbf{e}_{\mathrm{min}}|, \qquad (32)$$

where $|\cdot|$ represents the absolute value operation and $\gamma \in (0, +\infty)$, $|\mathbf{e}_{\mathrm{min}}|$ and $|\mathbf{e}_{\mathrm{med}}|$ respectively represent taking the minimum and median selection operation in the estimate window, and $\mathbf{e}_k$ is the residual vector at step k and $|\mathbf{e}_k| > |\mathbf{e}_{\mathrm{min}}|$.

Given the upper limit of the bandwidth σ_{max} and the adjustment coefficient k_σ, once $\gamma \geq \sigma_{\mathrm{max}}/k_\sigma$, then

$$\sigma_k = \sigma_{\mathrm{max}}. \qquad (33)$$

If there are existing distinct outliers in the measurement vector at step k, then $\gamma < \sigma_{\mathrm{max}}/k_\sigma$, and σ_k can be defined as

$$\sigma_k = k_\sigma \gamma, \qquad (34)$$

where k_σ represents the adaptive adjustment coefficient of σ_k.

Due to the operation of taking the minimum and median remove the impulsive noise interference, $|\mathbf{e}_{\mathrm{min}}|$ and $|\mathbf{e}_{\mathrm{med}}|$ always maintained within a stable range, and thus γ reduces significantly only at the $|\mathbf{e}_k|$ interference by the outliers. Therefore, in most cases, an appropriate k_σ can be determined to realize the adaptation of the kernel size to ensure the MCFIR filter's accuracy. As too small kernel bandwidth might cause filtering divergence, a minimum kernel bandwidth σ_{min} might need to be set to ensure filtering stability.

With the above derivations, we summarize the proposed MCFIR filter as follows:

IV. BIAS CONSTRAINED MCFIR FILTER

The unbiased MCFIR filter ignores the noise statistics, limiting the filtering accuracy due to process bias error accumulation. Appropriate forgetting factors may improve the unbiased MCFIR filter's performance in certain conditions, but the forgetting factor is set manually and cannot be guaranteed to be optimal.

In this section, we modify the unbiased MCFIR into a bias-constrained form, which considers the noise covariance $\mathbf{Q}_k$ and $\mathbf{R}_k$ to compensate bias error accumulation within estimate window, and further improve the performance of the unbiased MCFIR filter.

A. BATCH SOLUTION OF THE BIAS CONSTRAINED MCFIR FILTER

Consider the extended state space model (7) and equation (18), $\mathbf{x}_s$ can be expressed by $\mathbf{x}_k$ as

$$\mathbf{x}_s = \mathbf{A}_{k,s}^{-1}(\mathbf{x}_k - \bar{\mathbf{B}}_{k,s}\mathbf{W}_{k,s}), \qquad (35)$$

where both process and measurement noise interference are considered. Then,

$$\mathbf{Y}_{k,s} - \tilde{\mathbf{C}}_{k,s}\mathbf{x}_k = \mathbf{N}_{k,s}, \qquad (36)$$

Algorithm 1 Robust Unbiased MCFIR Filter

Input σ_{max}, k_σ and θ, N, and $\mathbf{Y}_{k,s}$
$\hat{x}_{k|k-1} = \mathbf{A}_k\hat{x}_{k-1}$
A. Batch form solution
$\quad \tilde{\mathbf{C}}_{k,s} = \bar{\mathbf{C}}_{k,s}[\mathbf{IA}_k^{-T}\cdots\mathbf{A}_{k,s+1}^{-T}\mathbf{A}_{k,s}^{-T}]^T$
$\quad \mathbf{e}_k = \mathbf{R}_{k,s}^{-1/2}(\mathbf{Y}_{k,s} - \tilde{\mathbf{C}}_{k,s}\hat{x}_{k|k-1})$
$\quad$ Calculate σ_k using (31)-(34)
$\quad \mathbf{\Theta}_{k,s} = \mathbf{diag}[G_\sigma(\mathbf{e}_k); G_\sigma(\mathbf{e}_{k-1}); \cdots; G_\sigma(\mathbf{e}_s)]$
$\quad \mathbf{\Omega}_{k,s} = \mathbf{\Theta}_{k,s}\mathrm{diag}[\mathbf{I}_m; \theta\mathbf{I}_m; \cdots; \theta^{N-2}\mathbf{I}_m; \theta^{N-1}\mathbf{I}_m]$
$\quad \hat{\mathbf{x}}_k = (\tilde{\mathbf{C}}_{k,s}^T\mathbf{\Omega}_{k,s}\tilde{\mathbf{C}}_{k,s})^{-1}\tilde{\mathbf{C}}_{k,s}^T\mathbf{\Omega}_{k,s}\mathbf{Y}_{k,s}$
B. Iterative form solution
$\quad \boldsymbol{\xi}_k = \hat{x}_{k|k-1}$
$\quad$ Backward for $i:k \to s$
$\quad \mathbf{e}_i = \mathbf{R}_i^{-1/2}(\mathbf{y}_i - \mathbf{C}_i\boldsymbol{\xi}_i)$
$\quad \boldsymbol{\xi}_{i-1} = \mathbf{A}_i^{-1}\boldsymbol{\xi}_i$
$\quad$ end for
$\quad$ Calculate $\tilde{G}_m$, $\tilde{\mathbf{x}}_m$ using (29)-(30)

$\quad$ **for** $i:m+1 \to k$
$\quad\quad \tilde{\mathbf{x}}_i^- = \mathbf{A}_i\tilde{\mathbf{x}}_{i-1}$
$\quad\quad$ Calculate σ_k using (31)-(34)
$\quad\quad \mathbf{\Omega}_i = \theta^{k-i}\mathrm{diag}[G_\sigma(\mathbf{e}_i)]$
$\quad\quad \tilde{G}_i = (\mathbf{C}_i^T\mathbf{\Omega}_i\mathbf{C}_i + (\mathbf{A}_i\tilde{G}_{i-1}\mathbf{A}_i^T)^{-1})^{-1}$
$\quad\quad \tilde{\mathbf{x}}_i = \tilde{\mathbf{x}}_i^- + \tilde{G}_i\mathbf{C}_i^T\mathbf{\Omega}_i(\mathbf{y}_i - \mathbf{C}_i\tilde{\mathbf{x}}_i^-)$
$\quad$ **end for**
Output $\hat{\mathbf{x}}_k = \tilde{\mathbf{x}}_k$

---

which represents the measurement vector $\mathbf{Y}_{k,s}$ derived by $\mathbf{x}_k$. According to (35), (36), and (8), within the estimate window, the nominal measurement noise $\mathbf{N}_{k,s}$ can be expressed as

$$\mathbf{N}_{k,s} = (\mathbf{H}_{k,s} - \tilde{\mathbf{C}}_{k,s}\bar{\mathbf{B}}_{k,s})\mathbf{W}_{k,s} + \mathbf{V}_{k,s}. \qquad (37)$$

Here, (37) contains the bias accumulation and measurement noise vectors.

Taking $\sum_{k,s}$ represent the expectation $\mathrm{E}[\mathbf{N}_{k,s}\mathbf{N}_{k,s}^T]$ as follows

$$\sum\nolimits_{k,s} = \mathrm{E}[\mathbf{N}_{k,s}\mathbf{N}_{k,s}^T]$$
$$= (\mathbf{H}_{k,s} - \tilde{\mathbf{C}}_{k,s}\bar{\mathbf{B}}_{k,s})\mathbf{Q}_{k,s}(\mathbf{H}_{k,s} - \tilde{\mathbf{C}}_{k,s}\bar{\mathbf{B}}_{k,s})^T + \mathbf{R}_{k,s}. \qquad (38)$$

with $\mathbf{Q}_{k,s} = \mathrm{diag}[\mathbf{Q}_k\mathbf{Q}_{k-1}\ldots\mathbf{Q}_s], \mathbf{R}_{k,s} = \mathrm{diag}[\mathbf{R}_k\mathbf{R}_{k-1}\ldots\mathbf{R}_s]$.

Then, a new correntropy based object function weighted by the residual covariance matrix $\sum_{k,s}$ is derived as follows:

$$J_{\mathrm{MCC}}(\mathbf{x}_k) = \sum_{i=s}^{k} G_\sigma\left(\left\|\mathbf{y_i} - \tilde{\mathbf{C}}_{k,i}\mathbf{x}_k\right\|_{\sum_i^{-1}}\right). \qquad (39)$$

Here, $\|\cdot\|_{\sum_i^{-1}}$ denotes the $\sum_i^{-1}$-weighted two-norm of a vector.

By taking the derivation of (39) with respect to $\mathbf{x}_k$, and equating it to zero, the robust solution $\hat{\mathbf{x}}_k$ is obtained as

$$\hat{\mathbf{x}}_k = (\tilde{\mathbf{C}}_{k,s}^T\sum\nolimits_{k,s}^{-1}\tilde{\mathbf{C}}_{k,s})^{-1}\tilde{\mathbf{C}}_{k,s}^T\sum\nolimits_{k,s}^{-1}\mathbf{Y}_{k,s}. \qquad (40)$$

Because the finite memory filters are robust to process uncertainties, it can be approximated that only $\mathbf{R}_k$ needs to be rescaled to attenuate the measurement outliers, therefore, the residual covariance matrix with rescaled measurement noise covariance $\tilde{\mathbf{R}}_{k,s} = \mathbf{R}_{k,s}^{1/2}\boldsymbol{\Theta}_{\mathbf{k,s}}^{-1}\mathbf{R}_{k,s}^{T/2}$, and the reweighted matrix $\tilde{\sum}_{k,s}$ can be formulated as

$$\tilde{\sum}_{k,s} = (\mathbf{H}_{k,s} - \tilde{\mathbf{C}}_{k,s}\bar{\mathbf{B}}_{k,s})\mathbf{Q}_{k,s}(\mathbf{H}_{k,s} - \tilde{\mathbf{C}}_{k,s}\bar{\mathbf{B}}_{k,s})^T + \tilde{\mathbf{R}}_{k,s}.$$

(41)

The robust solution (40) is the batch form solution of the bias-constrained MCFIR filter, which is formulated as a weighted least-square form. The unbiased MCFIR filter can be taken as a special case of (40) with $\sum_{k,s}^{-1} = \boldsymbol{\Omega}_{\mathbf{k,s}}$, where the process noise accumulation is ignored within the estimate window and ensures only an unbiased estimate result.

B. ITERATIVE SOLUTION OF THE BIAS CONSTRAINED MCFIR FILTER

The batch form bias-constrained MCFIR filter's operation brings an additional calculation burden than original unbiased MCFIR filter due to complex matrix operations, which might make implementation difficult in practical applications.

The nominal measurement residual covariance matrix $\boldsymbol{\Sigma}_k$ needs to be obtained in advance to implement the iterative form solution. Taking the process and measurement noise into consideration, the nominal measurement residual covariance of the $[s, k]$ interval can be expressed as

$$\boldsymbol{\Sigma}_{k,s} = \bar{\mathbf{C}}_{k,s}\mathbf{D}_{k,s}\mathbf{Q}_{k,s}\mathbf{D}_{k,s}^T\bar{\mathbf{C}}_{k,s}^T + \mathbf{R}_{k,s},$$

(42)

where

$$\mathbf{D}_{k,s} = \mathbf{B}_{k,s} - [\mathbf{I}\mathbf{A}_k^{-T}\cdots\mathcal{A}_{k,s+1}^{-T}\mathcal{A}_{k,s}^{-T}]^T\bar{\mathbf{B}}_{k,s}.$$

(43)

Taking $\boldsymbol{\Phi}_{k,s} = \mathbf{D}_{k,s}\mathbf{Q}_{k,s}\mathbf{D}_{k,s}^T$ to represent the error covariance of predict state vector in $[s, k]$. An approximate covariance matrix $\boldsymbol{\Phi}_i = \mathbf{D_i}\mathbf{Q_i}\mathbf{D_i}^T$ can be adopted for simplicity, and the corresponding nominal measurement noise covariance can be obtained by $\boldsymbol{\Sigma}_i = \mathbf{C}_i\boldsymbol{\Phi}_i\mathbf{C}_i^T + \mathbf{R}_i$, which can be determined by a backward recursion, and then

$$G_\sigma(\mathbf{e}_i) = G_\sigma(\mathbf{S}_i^{-1}(\mathbf{y}_i - \mathbf{C}_i\boldsymbol{\xi}_i)),$$

(44)

where $\boldsymbol{\Sigma}_i = \mathbf{S}_i\mathbf{S}_i^T$. The rescaled measurement noise covariance calculation is completed by $\tilde{\mathbf{R}}_i = \mathbf{R}_i^{1/2}\boldsymbol{\Theta}_i^{-1}\mathbf{R}_i^{T/2}$, where $\boldsymbol{\Theta_i} = \text{diag}[\mathbf{G}_\sigma(\mathbf{e_i})]$. After simplification, the proposed filter only rescaled the measured noise covariance matrix, while the rest of the formulas remained the original structure. As the iterative solution of the bias-constrained MCFIR filter is not unique, one of the representative Kalman-like methods is formulated as (43)-(45) proposed in prior work [30].

The remaining computation can be operated as following iterative Kalman-like formula as

$$\tilde{\mathbf{x}}_i = \mathbf{A}_i\tilde{\mathbf{x}}_{i-1} + \mathbf{L}_i(\mathbf{y}_i - \mathbf{C}_i\mathbf{A}_i\tilde{\mathbf{x}}_{i-1}),$$

(45)

where filter gain $\mathbf{L}_i = \boldsymbol{\Delta}_i + \bar{\boldsymbol{\Delta}}_i$ can be decomposed by two parts as

$$\boldsymbol{\Delta}_i = \mathbf{P}_i\mathbf{C}_i^T(\mathbf{C}_i\mathbf{P}_i\mathbf{C}_i^T + \tilde{\mathbf{R}}_i)^{-1},$$

(46)

$$\bar{\boldsymbol{\Delta}}_i = (\mathbf{I} - \boldsymbol{\Delta}_i\mathbf{C}_i)\mathbf{G}_i\mathbf{Z}_i\mathbf{J}_i,$$

(47)

where $\mathbf{J}_i = \mathbf{G}_i^T\mathbf{C}_i^T(\mathbf{C}_i\mathbf{P}_i\mathbf{C}_i^T + \tilde{\mathbf{R}}_i)^{-1}$ is defined for simplicity.

The iteration parameters $\mathbf{P}_i$, $\mathbf{G}_i$, $\mathbf{Z}_i$ are formulated as

$$\mathbf{P}_i = \mathbf{A}_i\mathbf{P}_{i-1}\mathbf{A}_i^T + \mathbf{Q}_i - \mathbf{A}_i\boldsymbol{\Delta}_{i-1}\mathbf{C}_{i-1}\mathbf{P}_{i-1}\mathbf{A}_i^T,$$

(48)

$$\mathbf{G}_i = \mathbf{A}_i(\mathbf{I} - \boldsymbol{\Delta}_{i-1}\mathbf{C}_{i-1})\mathbf{G}_{i-1},$$

(49)

$$\mathbf{Z}_i = (\mathbf{Z}_{i-1}^{-1} + \mathbf{G}_i\mathbf{C}_i\mathbf{J}_i)^{-1}.$$

(50)

Assuming that the iteration starts at m, the initial values of $\mathbf{P}_i$, $\mathbf{G}_i$, $\mathbf{Z}_i$ are given by

$$\mathbf{P}_m = \bar{\mathbf{B}}_{m,s}\mathbf{Q}_{m,s}\bar{\mathbf{B}}_{m,s}^T - \mathbf{A}_m\boldsymbol{\gamma}_{m-1}\prod_{m-1,s}^{-1}\boldsymbol{\gamma}_{m-1}^T\mathbf{A}_m^T,$$

(51)

$$\mathbf{G}_m = \mathcal{A}_{m,s} - \mathbf{A}_m\boldsymbol{\gamma}_{m-1,s}\prod_{m-1,s}^{-1}\mathbf{C}_{m-1,s},$$

(52)

$$\mathbf{Z}_m = (\mathbf{C}_{m,s}^T\prod_{m,s}^{-1}\mathbf{C}_{m,s})^{-1}.$$

(53)

To simplify the computational complexity, covariance matrix $\prod_{m,s}$, $\boldsymbol{\gamma}_{m-1}$ can be formulated as $\prod_{m,s} = \mathbf{H}_{m,s}\mathbf{Q}_{m,s}\mathbf{H}_{m,s}^T + \tilde{R}_{m,s}$, $\boldsymbol{\gamma}_{m-1} = \bar{\mathbf{B}}_{m-1,s}\mathbf{Q}_{m-1,s}\mathbf{C}_{m-1,s}^T$, which can be calculated offline in time invariant systems.

The increase of N makes the FIR filters converge to the optimal IIR filter and lose the characteristics of the finite memory filter. If the selected horizon length N is small, the modified MCFIR filter is dominated by unbiasedness, which is almost equivalent to an unbiased MCFIR filter. The iterative solution of the bias-constrained MCFIR filter requires more iterations than batch form methods. As for systems that do not require updating the filter parameters at each step, offline calculation saves much of the calculation amount.

V. SIMULATION RESULTS

In this section, we illustrate the performance of the proposed algorithms (the unbiased MCFIR filter is denoted as MCFIR-1, and the biased-constrained MCFIR filter is denoted as MCFIR-2) after applying them to a moving target tracking system. In addition to the classical KF, UFIR filter of [29], several latest robust filters, HKF [14] based on Huber m-estimator, a novel robust student-t Kalman filter (RSTKF) [9] derived for the heavy-tailed process, and MCKF [17] based on maximum correntropy criterion, and optimal Kalman filter (OKF) with exact instantaneous noise covariance simulated as benchmarks for comparisons. All filtering algorithms are coded with MATLAB and executed on a laptop with Intel Core i5-7300HQ CPU @2.5 GHz.

We considered a constant turn-rate motion target tracking system specified by (1) and (2) with

$$\mathbf{A} = \begin{bmatrix} 1 & \sin(\alpha T)/\alpha & 0 & -(1-\cos(\alpha T))/\alpha \\ 0 & \cos(\alpha T) & 0 & -\sin(\alpha T) \\ 0 & -(1-\cos(\alpha T))/\alpha & 1 & \sin(\alpha T)/\alpha \\ 0 & \sin(\alpha T) & 0 & \cos(\alpha T) \end{bmatrix},$$

(54)

Algorithm 2 Robust Bias Constrained MCFIR Filter

Input $\sigma_{\max}, k_\sigma, N$, and $\mathbf{Y}_{k,s}$
$\hat{x}_{k|k-1} = \mathbf{A}_k \hat{x}_{k-1}$
A. Batch form solution
   Calculate $\tilde{\mathbf{C}}_{k,s}, \sum_{k,s}$ using (15) and (38)
   $\mathbf{e}_k = \mathbf{S}_{k,s}^{-1}(\mathbf{Y}_{k,s} - \tilde{\mathbf{C}}_{k,s}\hat{x}_{k|k-1})$
   Calculate σ_k using (31)-(34)
   $\mathbf{\Theta}_{k,s} = \text{diag}[G_\sigma(\mathbf{e}_k); G_\sigma(\mathbf{e}_{k-1}); \cdots; G_\sigma(\mathbf{e}_s)]$
   Calculate $\tilde{\sum}_{k,s}$ using (41)
   $\hat{\mathbf{x}}_k = (\tilde{\mathbf{C}}_{k,s}^T \tilde{\sum}_{k,s}^{-1} \tilde{\mathbf{C}}_{k,s})^{-1} \tilde{\mathbf{C}}_{k,s}^T \tilde{\sum}_{k,s}^{-1} \mathbf{Y}_{k,s}$
B. Iterative form solution
   Backward for $i:k \to s$
     $\mathbf{\Phi}_i = \mathbf{D}_i \mathbf{Q}_i \mathbf{D}_i^T$
     $\mathbf{\Sigma}_i = \mathbf{C}_i \mathbf{\Phi}_i \mathbf{C}_i^T + \mathbf{R}_i$
     $\mathbf{e}_i = \mathbf{S}_i^{-1}(\mathbf{y}_i - \mathbf{C}_i \boldsymbol{\xi}_i)$
     $\boldsymbol{\xi}_{i-1} = \mathbf{A_i}^{-1} \boldsymbol{\xi}_i$
   end for
   Calculate $\tilde{\sum}_{m,s}$ and $\tilde{\mathbf{x}}_m$ using (41), (40)
  for $i:m+1 \to k$
   │ $\tilde{\mathbf{x}}_i^- = \mathbf{A}_i \tilde{\mathbf{x}}_{i-1}$
   │ Calculate σ_k using (31)-(34)
   │ $\tilde{R}_i = \mathbf{R}_i^{1/2} \mathbf{\Theta}_i^{-1} \mathbf{R}_i^{T/2}$
   │ Calculate $\mathbf{P}_i, \mathbf{G}_i$ and $\mathbf{Z}_i$ using (48)-(50) and obtain
   │ $\mathbf{\Delta}_i$ and $\bar{\Delta}_i$
   │ $\mathbf{L}_i = \mathbf{\Delta}_i + \bar{\Delta}_i$
   │ $\tilde{\mathbf{x}}_i = \tilde{\mathbf{x}}_i^- + \mathbf{L}_i(\mathbf{y}_i - \mathbf{C}_i \tilde{\mathbf{x}}_i^-)$
  end for
Output $\hat{\mathbf{x}}_k = \tilde{\mathbf{x}}_k$

$$\mathbf{\Gamma} = \begin{bmatrix} 0.5T^2 & 0 \\ T & 0 \\ 0 & 0.5T^2 \\ T & 0 \end{bmatrix}, \tag{55}$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{56}$$

Here, $T = 0.2$ and $\alpha = 0.1$, $\mathbf{\Gamma}$ represents the process noise transfer matrix, and set $\mathbf{Q} = \mathbf{diag}[0.05\ 0.1]$ and $\mathbf{R} = \mathbf{diag}[10\ 10]$. The state vector $\mathbf{x}_k = [x\dot{x}y\dot{y}]^T$, in which $x, y, \dot{x}, \dot{y}$, denotes the cartesian coordinates and corresponding velocities, and the initial state $\mathbf{x}_0 = [1111]^T$. It is assumed that $\hat{x}_0 = \mathbf{0}_{4\times1}$ as the filtering initial state.

The process and measurement noise vectors were generated from $\mathbf{v}_k \sim 0.95N(0, \mathbf{R}) + 0.05N(0, 100\mathbf{R})$, $\mathbf{w}_k \sim 0.95N(0, \mathbf{Q}) + 0.05N(0, 100\mathbf{Q})$ respectively.

The horizon length of the MCFIR filters $N = 35$, r = 1.345 for the Huber loss function, and $\sigma = 5$ for the kernel bandwidth of correntropy function. For MCFIR-1, the forget factor $\theta = 0.99$. For RSTKF, the degrees of freedom parameter, tuning parameter, were set as $w = v = 5, \tau = 5$. For robust Kalman filters, the iteration times $n = 10$. The $\sigma_{\max} = 9$, and $k_\sigma = 15$ as the kernel size adaptive tuning factor for the proposed filters.

To evaluate the estimate performance, taking

$$\text{RMSE}_{\text{pos}} \triangleq \sqrt{\frac{1}{M}\sum_{i=1}^{M}((\mathbf{x}_k^i - \hat{\mathbf{x}}_k^i)^2 + (\mathbf{y}_k^i - \hat{y}_k^i)^2)}, \tag{57}$$

$$\text{ARMSE}_{\text{pos}} \triangleq \sqrt{\frac{1}{MK}\sum_{k=1}^{K}\sum_{i=1}^{M}((\mathbf{x}_k^i - \hat{\mathbf{x}}_k^i)^2 + (\mathbf{y}_k^i - \hat{y}_k^i)^2)}, \tag{58}$$

represent the estimate errors of different filters, where $(\mathbf{x}_k^i, \mathbf{y}_k^i)$ and $(\hat{\mathbf{x}}_k^i, \hat{y}_k^i)$ were the true and estimated positions at k th step of the ith Monte Carlo run, $M = 500$, and $K = 500$. Like the root mean square error (RMSE) and average root mean square error (ARMSE) of the position, we define the RMSE and ARMSE of velocity by the corresponding estimated states. The RMSE of the position and velocity of existing robust filter algorithms and the proposed filters are shown in Fig. 5. The MCFIR filters proposed in this paper had no output until the Nth step, and we evaluated the performance of the filters after the Nth step.

1) In Fig. 1, we made a preliminary comparison of the trajectories generated by different robust filters within a typical moving trajectory. The proposed filters achieved better tracking performance than the existing robust filters, which could eliminate the filter failure caused by measurement outliers. Moreover, as the target's motion state changed, the proposed MCFIR filters showed strong tracking performance, and the estimation trajectories were closer to the real trajectory.
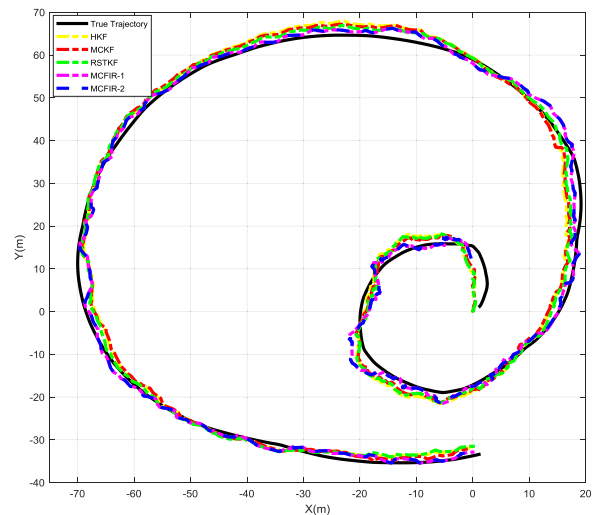

FIGURE 1. Tracking trajectories of different algorithms.

In Table 1 and Table 2, we explore the performance of the MCFIR filters with a distinct kernel bandwidth. An improper σ causes the filter accuracy to degrade. The kernel bandwidth adaptive adjustment method proposed in this work ensures that the MCFIR filters operate at a relatively optimal status.

As shown in Table 3 and Fig. 2, the MCFIR filters obtained lower errors than the existing robust filter, while the process and measurement noise were corrupted by a certain percentage of outliers. Moreover, MCFIR-1 and MCFIR-2 retained

TABLE 1. ARMSE of MCFIR-1 with distinct kernel size.

Filters	$ARMSE_p$ (m)	$ARMSE_v$ (m/s)
MCFIR-1($\sigma = 2$)	1.715	0.779
MCFIR-1($\sigma = 3$)	1.704	0.776
MCFIR-1($\sigma = 5$)	1.805	0.790
MCFIR-1($\sigma = 9$)	2.184	0.850
MCFIR-1($\sigma_{adptive}$)	1.691	0.774

TABLE 2. ARMSE of MCFIR-2 with distinct kernel size.

Filters	$ARMSE_p$ (m)	$ARMSE_v$ (m/s)
MCFIR-2($\sigma = 2$)	1.729	0.780
MCFIR-2($\sigma = 3$)	1.714	0.776
MCFIR-2($\sigma = 5$)	1.813	0.794
MCFIR-2($\sigma = 9$)	2.191	0.853
MCFIR-2($\sigma_{adptive}$)	1.700	0.774

TABLE 3. ARMSE of different filters.

Filters	$ARMSE_p$ (m)	$ARMSE_v$ (m/s)
KF	3.035	0.920
OKF	1.543	0.719
UFIR	3.641	1.135
HKF	2.243	0.857
MCKF	2.026	0.835
RSTKF	1.834	0.821
MCFIR-1	1.691	0.774
MCFIR-2	1.700	0.774

similar performance while confirming that the MCFIR-1 and MCFIR-2 filter is nearly equivalent as the bias accumulation is not apparent.

The standard unbiased FIR filter, is severely corrupted by measurement outliers and cannot ensure reliable estimations, which implies that the standard unbiased FIR filter cannot retain robustness in excessive measurement outliers. By comparing the unbiased FIR filter with the proposed MCFIR filters, it can be found that the proposed filters obtain superior performance, which confirms that the MCFIR filter is an effective FIR filter to cope with state estimate problem in dynamic process corrupted by non-Gaussian noises.

2) The RMSEs of different filters corresponding to different outlier percentages are shown in Fig. 2. The ARMSE of different filters growth with outlier density, and the proposed filters always obtained a lower ARMSE than the other existing filters.

3) To verify the robustness of FIR filters to model uncertain errors, an additional interference $[0.1; 0.3\sin(kT); 0.3; -0.2\cos(kT)]$ was added into the system transform matrix in [240,260]. As shown in Fig. 4, the MCFIR filters recovered to normal status within fewer steps, which indicates that the MCFIR filters had better robustness to model uncertain interference than other robust Kalman filters.

4) The horizon length N is one of the most important parameters that determine the characteristics of FIR filters. To compensate the bias accumulation caused by a longer N on MCFIR-1, Fig. 5 shows that an appropriate forget factor might improve this problem to some extent.
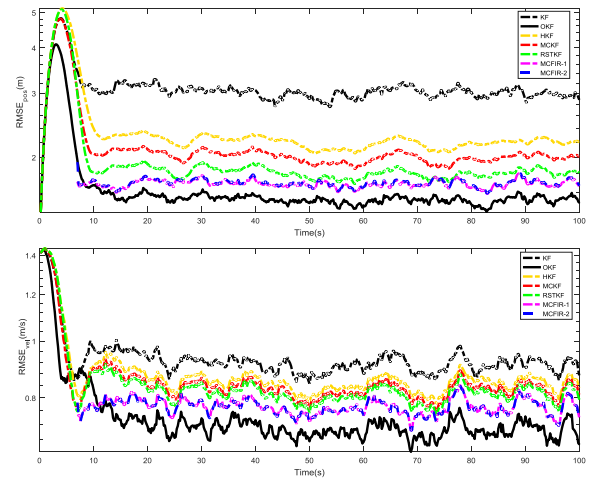

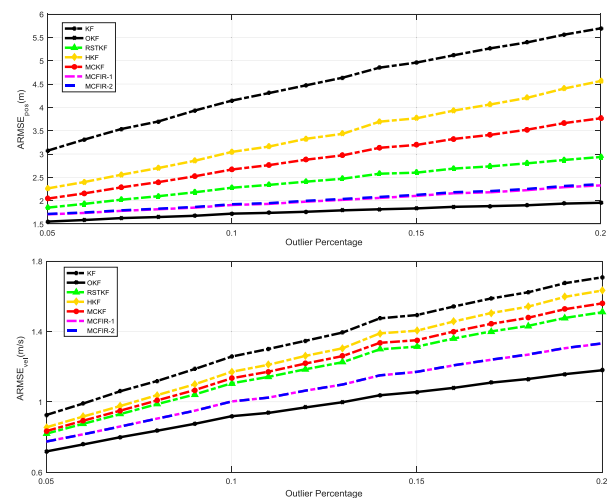**FIGURE 2.** RMSE of the existing filters and the proposed filters.

FIGURE 3. ARMSE of filters with different outlier percentage.
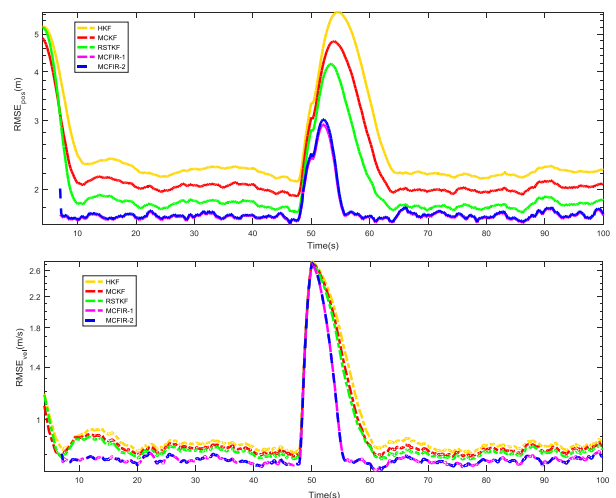

FIGURE 4. RMSE of different filters with temporary uncertainties.

Nevertheless, the MCFIR-2 avoids filtering divergence caused by the excessive bias accumulation, which is a significant advantage compared with MCFIR-1. To further explore

the adaptive kernel method proposed in this work, we plotted the ARMSE-N curves of the MCFIR filters with $\sigma = 5$ and adaptive σ in Fig. 5. The σ adaptive MCFIR filters always kept a lower ARMSE than filters with a fixed σ.
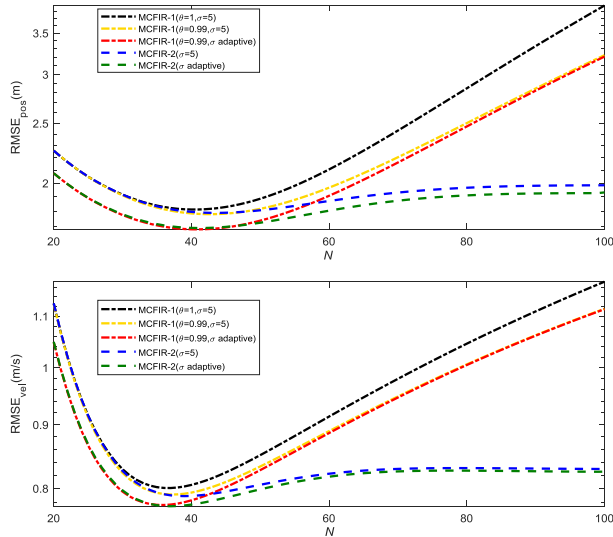

FIGURE 5. ARMSE of MCFIR filter varies with *N*.

Computational time is another essential characteristic of the filters, and the average total computational time of different filters is evaluated. Table 4 shows that the standard Kalman filter (0.009s) is the fastest, and the robust Kalman filters: HKF (0.175s), MCKF (0.220s), and RSTKF (0.207s) requires more computation time than Kalman filter. Then, the batch form (0.243s) and iterative form (0.292s) UFIR filter are the fastest among the FIR filters, and MCFIR-1 filter with batch (0.245s) and iterative (0.351s) form need a little additional computational time, which consumes extra time for correntropy function related calculations. As for the bias-constrained MCFIR-2 filters with batch (0.271s) and iterative (0.981s) form, complex matrix operations bring more computational burdens. As the matrix operations are optimized in simulation programs, and the iterative FIR filters have higher complexity and loop times, making the iterative FIR filters consume a longer running time. Owing to the FIR filter structure, there are many factors that might affect its computational speed, such as horizon length, data pre-processing method and so on. In practical applications, the

TABLE 4. Average computational time of different filters.

Filters	Time (s)
KF	0.009
HKF	0.175
MCKF	0.220
RSTKF	0.207
UFIR (Batch /Iterative)	0.243/0.292
MCFIR-1(Batch /Iterative)	0.245/0.351
MCFIR-2(Batch /Iterative)	0.271/0.981

TABLE 5. ARMSE of different filters.

Filters	$\mathrm{ARMSE_p}$ (m)	$\mathrm{ARMSE_v}$(m/s)	$\mathrm{ARMSE_a}$ (m/s²)
KF	5.381	2.962	1.097
OKF	2.518	1.824	0.931
UFIR	6.450	4.475	1.472
HKF	3.530	2.451	1.046
MCKF	3.227	2.355	1.035
MCFIR-1($\sigma = 5$)	3.100	2.323	1.031
MCFIR-2($\sigma = 5$)	3.106	2.343	1.036
MCFIR-1($\sigma_{\mathrm{adptive}}$)	2.852	2.181	1.008
MCFIR-2($\sigma_{\mathrm{adptive}}$)	2.865	2.202	1.013

computational complexity can be simplified by the adaptive horizon and offline computation to improve filtering efficiency.

5) For the 2-D constant turn-rate model used in the above cases, which was applied to generate the basic trajectory in the horizontal plane, we made relevant deformations and obtained a 3-D coordinated turn (3DCT) model including the nine-dimensional state vector with the position, velocity, and acceleration elements to further verify the applicability of the proposed filters [35]. If the target performs a mixture of horizontal and vertical turn maneuvers, the 3DCT model has higher tracking accuracy.

The sate transform matrix for the 3DNCT model can be reformulated as

$$\mathbf{A}_{3DCT} = \begin{bmatrix} 1 & \sin(\alpha T)/\alpha & (1-\cos(\alpha T))/\alpha^2 \\ 0 & \cos(\alpha T) & \sin(\alpha T)/\alpha \\ 0 & -\sin(\alpha T)/\alpha & \cos(\alpha T) \end{bmatrix}, \quad (59)$$

$$\mathbf{A} = \mathrm{diag}[\mathbf{A}_{3DCT}\mathbf{A}_{3DCT}\mathbf{A}_{3DCT}], \quad (60)$$

$$\mathbf{\Gamma} = \begin{bmatrix} T^3/6 & T^2/2 & T & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & T^3/6 & T^2/2 & T & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & T^3/6 & T^2/2 & T \end{bmatrix}^T \quad (61)$$

$\mathbf{Q} = \mathbf{diag}[0.05, 0.1, 0.1]$ and $\mathbf{R} = \mathbf{diag}[10\,10\,10]$, The state vector $\mathbf{x}_k = [x\dot{x}\ddot{x}y\dot{y}\ddot{y}z\dot{z}\ddot{z}]^T$ denotes the Cartesian coordinates, velocities, and corresponding accelerations, and the initial state $\mathbf{x}_0 = [0\,0\,0.1\,0\,0.1\,0\,0.1]^T$. It was assumed that $\hat{x}_0$ sampled form $\mathbf{x}_0$ with a covariance $\mathbf{\Gamma Q \Gamma}^T$ in each simulation.

The process and measurement noise vectors are generated from $\mathbf{v}_k \sim 0.95N(0, \mathbf{R}) + 0.05N(0, 100\mathbf{R})$ and $\mathbf{w}_k \sim 0.95N(0, \mathbf{Q}) + 0.05N(0, 100\mathbf{Q})$ respectively.

To maintain the consistency with above cases, each filter takes the same parameters as the above examples for comparison. In this condition, the RSTKF appears divergence caused by mathematical problems, and it's not included in the comparison.

The simulation results confirmed again that the proposed MCFIR filters significantly outperformed the existing robust filters, while the system process and measurement were severely disturbed by non-Gaussian noises. Moreover, this example confirmed that the proposed filters' tuning parameters had excellent compatibility, and it is not always necessary to search for new optimal tuning parameters for the proposed filter.

VI. CONCLUSION

A novel robust MCFIR filter was proposed in this paper, the filter operates as a finite memory form. It takes advantage of the robust characteristic of maximum correntropy method to attenuate the interference of non-Gaussian noise. An adaptive kernel size algorithm applied for the MCFIR filter was also proposed to ensure the filter operates with the appropriate kernel size. Simulation results confirmed that the proposed MCFIR filter obtain superior characteristics to existing robust filters in linear state-space systems corrupted with non-Gaussian noise.

APPENDIX A

DERIAVATION OF ITERATIVE UNBIASED MCFIR

Representing the inverse of the modified GNPG as $\tilde{\mathbf{G}}_k^{-1} = \tilde{\mathbf{C}}_{k,s}^T \mathbf{\Omega}_{k,s} \tilde{\mathbf{C}}_{k,s}$ which can be decomposed into the sum of several small matrices in the past N steps

$$\tilde{\mathbf{G}}_k^{-1} = \sum_{i=s}^{k} \mathcal{A}_{k,i}^{-T} \mathbf{C}_i^T \mathbf{\Omega}_i \mathbf{C_i} \mathcal{A}_{k,i}^{-1}, \qquad (A.1)$$

as $i = k$, then $\mathcal{A}_{k,s}^{-T} \mathbf{C}_k^T \mathbf{\Omega}_k \mathbf{C}_k \mathcal{A}_{k,s}^{-1} = \mathbf{C}_k^T \mathbf{\Omega}_k \mathbf{C}_k$ and

$$\tilde{\mathbf{G}}_k^{-1} = \mathbf{C}_k^T \mathbf{\Omega}_k \mathbf{C}_k + \Sigma_{i=s}^{k-1} \mathcal{A}_{k,i}^{-T} \mathbf{C}_i^T \mathbf{\Omega}_i \mathbf{C}_i \mathcal{A}_{k,i}^{-1}, \qquad (A.2)$$

since $\tilde{\mathbf{G}}_{k-1}^{-1} = \Sigma_{i=s}^{k-1} \mathbf{A}_{k-1,i}^{-T} \mathbf{C}_i^T \mathbf{\Omega}_i \mathbf{C}_i \mathbf{A}_{k-1,i}^{-1}$, combine the above equations then

$$\tilde{\mathbf{G}}_k = \left(\mathbf{C}_k^T \mathbf{\Omega}_k \mathbf{C}_k + (\mathbf{A}_k \tilde{\mathbf{G}}_{k-1} \mathbf{A}_k^T)^{-1} \right)^{-1}, \qquad (A.3)$$

Similarly, we rewrite $\tilde{\mathbf{C}}_{k,s}^T \mathbf{\Omega}_{k,s} \mathbf{Y}_{k,s}$ as

$$\tilde{\mathbf{C}}_{k,s}^T \mathbf{\Omega}_{k,s} \mathbf{Y}_{k,s} = \mathbf{C}_k^T \mathbf{\Omega}_k \mathbf{y}_k + \mathbf{A}_k^{-T} \tilde{\mathbf{C}}_{k-1,s}^T \mathbf{\Omega}_{k-1,s} \mathbf{Y}_{k-1,s}, \qquad (A.4)$$

then

$$\tilde{\mathbf{x}}_k = \tilde{\mathbf{G}}_k (\mathbf{C}_k^T \mathbf{\Omega}_k \mathbf{y}_k + \mathbf{A}_k^{-T} \tilde{\mathbf{C}}_{k-1,s}^T \mathbf{\Omega}_{k-1,s} \mathbf{Y}_{k-1,s}), \quad (A.5)$$

take step at $k-1$, $\tilde{\mathbf{C}}_{k-1,s}^T \mathbf{\Omega}_{k-1,s} \mathbf{Y}_{k-1,s} = \tilde{\mathbf{G}}_{k-1}^{-1} \tilde{\mathbf{x}}_{k-1}$

$$\tilde{\mathbf{x}}_k = \tilde{\mathbf{G}}_k (\mathbf{C}_k^T \mathbf{\Omega}_k \mathbf{y}_k + \mathbf{A}_k^{-T} \tilde{\mathbf{G}}_{k-1}^{-1} \tilde{\mathbf{x}}_{k-1}), \qquad (A.6)$$

substitute $\tilde{\mathbf{G}}_{k-1}$ of (27) to the (30) equation, we could obtain the final recursive Kalman like form as follow

$$\tilde{\mathbf{x}}_k = \mathbf{A}_k \tilde{\mathbf{x}}_{k-1} + \tilde{\mathbf{G}}_k \mathbf{C}_k^T \mathbf{\Omega}_k (\mathbf{y}_k - \mathbf{C}_k \mathbf{A}_k \tilde{\mathbf{x}}_{k-1}). \qquad (A.7)$$

REFERENCES

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, Mar. 1960.

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

[3] E. A. Wan and R. van der Merwe, "The unscented Kalman filter for nonlinear estimation," in *Proc. IEEE Adapt. Syst. Signal Process., Commun., Control Symp.*, Oct. 2000, pp. 153–158.

[4] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," *IEEE Trans. Autom. Control*, vol. 54, no. 6, pp. 1254–1269, Jun. 2009.

[5] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, 2002.

[6] H. Zhu, H. Leung, and Z. He, "A variational Bayesian approach to robust sensor fusion based on student-t distribution," *Inf. Sci.*, vol. 221, pp. 201–214, Feb. 2013.

[7] H. Nurminen, T. Ardeshiri, R. Piche, and F. Gustafsson, "Robust inference for state-space models with skewed measurement noise," *IEEE Signal Process. Lett.*, vol. 22, no. 11, pp. 1898–1902, Nov. 2015.

[8] Y. Huang, Y. Zhang, N. Li, and J. Chambers, "Robust student's t based nonlinear filter and smoother," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 5, pp. 2586–2596, Oct. 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] P. J. Huber, *Robust Statistics*. Berlin, Germany: Springer, 2011.

[11] L. Chang, K. Li, and B. Hu, "Huber's M-estimation-based process uncertainty robust filter for integrated INS/GPS," *IEEE Sensors J.*, vol. 15, no. 6, pp. 3367–3374, Jun. 2015.

[12] B. Kovacevic, Z. Durovic, and S. Glavaski, "On robust Kalman filtering," *Int. J. Control*, vol. 56, no. 3, pp. 547–562, Sep. 1992.

[13] M. A. Gandhi and L. Mili, "Robust Kalman filter based on a generalized maximum-likelihood-type estimator," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2509–2520, May 2010.

[14] L. Chang, B. Hu, G. Chang, and A. Li, "Robust derivative-free Kalman filter based on Huber's M-estimation methodology," *J. Process Control*, vol. 23, no. 10, pp. 1555–1561, Nov. 2013.

[15] L. Chang, B. Hu, G. Chang, and A. Li, "Multiple outliers suppression derivative-free filter based on unscented transformation," *J. Guid., Control, Dyn.*, vol. 35, no. 6, pp. 1902–1906, Nov. 2012.

[16] X. Liu, H. Qu, J. Zhao, and B. Chen, "State space maximum correntropy filter," *Signal Process.*, vol. 130, pp. 152–158, Jan. 2017.

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

[18] G. Wang, Y. Zhang, and X. Wang, "Iterated maximum correntropy unscented Kalman filters for non-Gaussian systems," *Signal Process.*, vol. 130, pp. 87–94, Oct. 2019.

[19] H. Wang, H. Li, W. Zhang, J. Zuo, and H. Wang, "Maximum correntropy derivative-free robust Kalman filter and smoother," *IEEE Access*, vol. 6, pp. 70794–70807, 2018, doi: 10.1109/ACCESS.2018.2880618.

[20] B. Chen, J. Wang, H. Zhao, N. Zheng, and J. C. Principe, "Convergence of a fixed-point algorithm under maximum correntropy criterion," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1723–1727, Oct. 2015.

[21] G. Wang, N. Li, and Y. Zhang, "Maximum correntropy unscented Kalman and information filters for non-Gaussian measurement noise," *J. Franklin Inst.*, vol. 354, no. 18, pp. 8659–8677, Dec. 2017.

[22] X. Liu, H. Qu, J. Zhao, and P. Yue, "Maximum correntropy square-root cubature Kalman filter with application to SINS/GPS integrated systems," *ISA Trans.*, vol. 80, pp. 195–202, Sep. 2018.

[23] F. Huang, J. Zhang, and S. Zhang, "Adaptive filtering under a variable kernel width maximum correntropy criterion," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 64, no. 10, pp. 1247–1251, Oct. 2017, doi: 10.1109/TCSII.2017.2671339.

[24] S. Fakoorian, R. Izanloo, A. Shamshirgaran, and D. Simon, "Maximum correntropy criterion Kalman filter with adaptive kernel size," in *Proc. IEEE Nat. Aerosp. Electron. Conf. (NAECON)*, Dayton, OH, USA, Jul. 2019, pp. 581–584, doi: 10.1109/NAECON46414.2019.9057886.

[25] B. Chen, X. Wang, N. Lu, S. Wang, J. Cao, and J. Qin, "Mixture correntropy for robust learning," *Pattern Recognit.*, vol. 79, pp. 318–327, Jul. 2018.

[26] H. Wang, W. Zhang, J. Zuo, and H. Wang, "Outlier-robust Kalman filters with mixture correntropy," *J. Franklin Inst.*, vol. 357, no. 8, pp. 5058–5072, May 2020, doi: 10.1016/j.jfranklin.2020.03.042.

[27] W. H. Kwon, P. S. Kim, and S. H. Han, "A receding horizon unbiased FIR filter for discrete-time state space models," *Automatica*, vol. 38, no. 3, pp. 545–551, Mar. 2002.

[28] W. H. Kwon, P. S. Kim, and P. Park, "A receding horizon Kalman FIR filter for discrete time-invariant systems," *IEEE Trans. Autom. Control*, vol. 44, no. 9, pp. 1787–1791, 1999.

[29] Y. S. Shmaliy, "An iterative Kalman-like algorithm ignoring noise and initial conditions," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2465–2473, Jun. 2011.

[30] S. Zhao, Y. S. Shmaliy, and F. Liu, "Fast Kalman-like optimal unbiased FIR filtering with applications," *IEEE Trans. Signal Process.*, vol. 64, no. 9, pp. 2284–2297, May 2016.

[31] S. Zhao, Y. S. Shmaliy, and C. K. Ahn, "Iterative maximum likelihood FIR estimation of dynamic systems with improved robustness," *IEEE/ASME Trans. Mechatronics*, vol. 23, no. 3, pp. 1467–1476, Jun. 2018.

[32] S. Zhao, Y. Ma, and B. Huang, "Robust FIR state estimation of dynamic processes corrupted by outliers," *IEEE Trans. Ind. Informat.*, vol. 15, no. 1, pp. 139–147, Jan. 2019.

[33] S. Zhao, Y. S. Shmaliy, C. K. Ahn, and F. Liu, "Adaptive-horizon iterative UFIR filtering algorithm with applications," *IEEE Trans. Ind. Electron.*, vol. 65, no. 8, pp. 6393–6402, Aug. 2018.

[34] F. Ramirez-Echeverria, A. Sarr, and Y. S. Shmaliy, "Optimal memory for discrete-time FIR filters in state-space," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 557–561, Feb. 2014.

[35] J. D. Glass, W. D. Blair, and Y. Bar-Shalom, "IMM estimators with unbiased mixing for tracking targets performing coordinated turns," in *Proc. IEEE Aerosp. Conf.*, Big Sky, MT, USA, Mar. 2013, pp. 1–10.

XUYOU LI received the B.E. and M.E. degrees in communication engineering and the Ph.D. degree in navigation, guidance, and control from the Department of Electronic Engineering, College of Automation, Harbin Engineering University, Harbin, China, in 1987, 1990, and 2002, respectively. From 2002 to 2005, he was a Postdoctoral Fellow with the Harbin Institute of Technology, Harbin. He is currently a Professor of precision instruments and machinery with Harbin Engineering University. His research interests include fiber optical gyroscope, fiber optical sensors, strapdown inertial navigation systems, and integrated navigation.

QINGWEN MENG received the B.S. degree from the Department of Equipment Engineering, Shenyang Ligong University, China, in 2014. She is currently pursuing the Ph.D. degree with the Department of Automation, Harbin Engineering University, Harbin, China. She is also a joint Ph.D. Student with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada. Her current research interests include signal processing, information fusion, and their applications in navigation technology, such as inertial navigation and integrated navigation.

YANDA GUO received the B.S. degree from the Department of Automation, Harbin Engineering University, Harbin, China, in 2016, where he is currently pursuing the Ph.D. degree. His current research interests include signal processing, process control, and their applications in navigation technology, such as strapdown inertial navigation systems and integrated navigation.

• • •