Minimum Error Entropy Kalman Filter

Badong Chen®, Senior Member, IEEE, Lujuan Dang®, Yuantao Gu®, Senior Member, IEEE, Nanning Zheng®, Fellow, IEEE, and José C. Príncipe®, Fellow, IEEE

Abstract—To date, most linear and nonlinear Kalman filters (KFs) have been developed under the Gaussian assumption and the well-known minimum mean square error (MMSE) criterion. In order to improve the robustness with respect to impulsive (or heavy-tailed) non-Gaussian noises, the maximum correntropy criterion (MCC) has recently been used to replace the MMSE criterion in developing several robust Kalman-type filters. To deal with more complicated non-Gaussian noises such as noises from multimodal distributions, in this article, we develop a new Kalman-type filter, called minimum error entropy KF (MEE-KF), by using the minimum error entropy (MEE) criterion instead of the MMSE or MCC. Similar to the MCC-based KFs, the proposed filter is also an online algorithm with the recursive process, in which the propagation equations are used to give prior estimates of the state and covariance matrix, and a fixed-point algorithm is used to update the posterior estimates. In addition, the MEE extended KF (MEE-EKF) is also developed for performance improvement in the nonlinear situations. The high accuracy and strong robustness of MEE-KF and MEE-EKF are confirmed by experimental results.

Index Terms—Kalman filtering, minimum error entropy (MEE), non-Gaussian noises, robust estimation.

I. INTRODUCTION

Kalman filtering is a powerful technology for estimating the states of a dynamic system, which finds applications in many areas, including navigation, guidance, data integration, pattern recognition, tracking, and control systems [1]–[6]. The original Kalman filter (KF) was derived for a linear state-space model with Gaussian assumption [2], [7], [8]. To cope with nonlinear estimation problems, a variety of nonlinear extensions of the original KF have been proposed in the literature, including extended KF (EKF) [9], [10], unscented KF (UKF) [11], cubature KF ( CKF) [12], and many others [13]. However, most of these KFs are developed based on the popular minimum mean square error (MMSE) criterion and will face performance degradation in case of complicated noises, since in general MMSE is not a good choice for the estimation in non-Gaussian noises.

In recent years, to solve the performance degradation problem in heavy-tailed (or impulsive) non-Gaussian noises, some robust KFs have been developed by using certain non-MMSE criterion as the optimality criterion [14], [15]. Particularly, the maximum correntropy criterion (MCC) [16], [17] in information theoretic learning (ITL) [14], [15] has been successfully applied in Kalman filtering to improve the robustness against impulsive noises. Typical examples include the maximum correntropy-based KFs [18]–[27], maximum correntropy-based EKFs [28]–[30], maximum correntropy-based UKFs [31]–[34], maximum correntropy-based square-root CKFs [35], [36], and so on. Since correntropy is a local similarity measure and insensitive to large errors, these MCC-based filters are little influenced by large outliers compared with MMSE-based filters [16], [37].

The MCC is a nice choice for dealing with heavy-tailed non-Gaussian noises, but its performance may not be good when facing more complicated non-Gaussian noises, such as noises from multimodal distributions. The minimum error entropy (MEE) criterion [38], [39] is another important learning criterion in ITL, which has been successfully applied in robust regression, classification, system identification, and adaptive filtering [38]–[42]. Numerous experimental results show that MEE can outperform MCC in many situations although its computational complexity is a little higher [43], [44]. In addition, the superior performance and robustness of MEE have been proved in [45]. The goal of this article is to develop a new Kalman-type filter, called MEE-KF, by using the MEE as the optimality criterion. The proposed filter uses the propagation equations to obtain the prior estimates of the state and covariance matrix, and a fixed-point algorithm to update the posterior estimates and covariance matrix, recursively and online. To further improve the performance in the nonlinear situations, the MEE criterion is also incorporated into EKF, resulting in MEE extended KF (MEE-EKF).

The rest of this article is organized as follows. In Section II, we briefly review the KF algorithm and MEE criterion. In Section III, we develop the MEE-KF algorithm. Sections IV and V provide the computational complexity and convergence analysis, respectively. In Section VI, the MEE-EKF
is developed. The experimental results are presented in Section VII and finally, the conclusion is given in Section VIII.

II. BACKGROUND

A. Kalman Filter

Consider a linear dynamic system with unknown state vector \( x(k) \in \mathbb{R}^{n \times 1} \) and available measurement vector \( y(k) \in \mathbb{R}^{m \times 1} \). To estimate the state \( x(k) \), KF assumes a state-space model described by the following state and measurement equations:

\[
\begin{align*}
\dot{x}(k) &= Fx(k-1) + q(k-1) \\
y(k) &= Hx(k) + r(k)
\end{align*}
\]

(1)
(2)

where \( F \in \mathbb{R}^{n \times n} \) and \( H \in \mathbb{R}^{m \times n} \) are the state-transition matrix and measurement matrix, respectively. Here, the process noise \( q(k-1) \in \mathbb{R}^{n \times 1} \) and measurement noise \( r(k) \in \mathbb{R}^{m \times 1} \) are mutually independent, and satisfy

\[
\begin{align*}
E[q(k-1)] &= 0; E[r(k)] &= 0 \\
E[q(k-1)q^T(k-1)] &= Q(k-1) \\
E[r(k)r^T(k)] &= R(k)
\end{align*}
\]

(3)
(4)
(5)

where \( Q(k-1) \) and \( R(k) \) are the covariance matrices of \( q(k-1) \) and \( r(k) \), respectively. In general, the KF includes two steps.

1) Predict: The a-priori estimate \( \hat{x}(k|1) \) and the corresponding error covariance matrix \( P(k|1) \) are calculated by

\[
\begin{align*}
\hat{x}(k|1) &= F\hat{x}(k-1) \\
P(k|1) &= FP(k-1)F^T + Q(k-1).
\end{align*}
\]

(6)
(7)

2) Update: The a-posteriori estimate \( \hat{x}(k) \) and the corresponding error covariance matrix \( P(k) \) are obtained by

\[
\begin{align*}
\hat{x}(k) &= \hat{x}(k|1) + K(k)[y(k) - H\hat{x}(k|1)] \\
K(k) &= P(k|1)H^T[HPP(k|1)H^T + R(k)]^{-1}
\end{align*}
\]

(8)
(9)

\[
\begin{align*}
P(k) &= [I - K(k)H]P(k|1)\{I - K(k)H\}^T \\
&\quad + K(k)R(k)K^T(k)
\end{align*}
\]

(10)

where \( K(k) \) is the KF gain.

B. Minimum Error Entropy Criterion

Define the error of two random variables \( X \) and \( Y \) as \( e = Y - X \). In MEE, the error information \( e \) can be measured by Renyi’s entropy

\[
H_\alpha(e) = \frac{1}{1 - \alpha} \log V_\alpha(e)
\]

(11)

where \( \alpha (\alpha \neq 1, \alpha > 0) \) is the order of Renyi’s entropy, and \( V_\alpha(e) \) denotes the information potential defined by

\[
V_\alpha(e) = \int p^\alpha(x)dx = E[p^{\alpha-1}(e)]
\]

(12)

where \( p(\cdot) \) is the probability density function (PDF) of error \( e \) and \( E[\cdot] \) denotes the expectation operator. In practical applications, the PDF \( p(x) \) can be estimated by Parzen’s window approach [14]

\[
\hat{p}(x) = \frac{1}{N} \sum_{i=1}^{N} G_\sigma(x - e_i)
\]

(13)

where \( G_\sigma(x) = \exp(-x^2/2\sigma^2) \) denotes the Gaussian kernel with kernel size \( \sigma \); \( |e_i|_{i=1}^{N} \) are \( N \) error samples. Combining (12) and (13), one can obtain an estimate of the second order (\( \alpha = 2 \)) information potential \( V_2(e) \)

\[
\hat{V}_2(e) = \frac{1}{N} \sum_{i=1}^{N} \hat{p}(e_i) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} G_\sigma(e_i - e_j).
\]

(14)

Since the negative logarithmic function \( -\log \) is monotonically decreasing, minimizing the error entropy \( H_2(e) \) means maximizing the information potential \( \hat{V}_2(e) \).

III. MINIMUM ERROR ENTROPY KALMAN FILTER

A. Augmented Model

First, we denote the state prediction error as

\[
e(k|k-1) = x(k) - \hat{x}(k|k-1).
\]

(15)

Combining the above state prediction error with the measurement (2), one can obtain an augmented model

\[
\begin{align*}
\tilde{x}(k|k-1) &= [I \quad \theta_1] x(k) + \mu(k) \\
y(k) &= [\theta_2 \theta_1] \tilde{x}(k|k-1)
\end{align*}
\]

(16)

where \( \theta_1 \) denotes an \( n \times n \) identity matrix, and

\[
\mu(k) = [-e(k|k-1) \quad r(k)]^T
\]

(17)

is the augmented noise vector comprising of the state and measurement errors. Assuming that the covariance matrix of the augmented noise \( E[\mu(k)\mu(k)^T] \) is positive definite, we have

\[
E[\mu(k)\mu(k)^T] = \Theta(k)\Theta(k)^T
\]

(18)

\[
= \begin{bmatrix}
\Theta_p(k|k-1) & 0 \\
0 & \Theta_r(k)\Theta_c(k)^T
\end{bmatrix}
\]

where \( \Theta(k) \), \( \Theta_p(k|k-1) \), and \( \Theta_r(k) \) are obtained by the Cholesky decomposition of \( E[\mu(k)\mu(k)^T] \), \( P(k|k-1) \), and \( R(k) \), respectively. Multiplying both sides of (16) by \( \Theta^{-1}(k) \) gives

\[
d(k) = W(k)x(k) + e(k)
\]

(19)

where

\[
\begin{align*}
d(k) &= \Theta^{-1}(k) \begin{bmatrix}
\hat{x}(k|k-1) \\
y(k)
\end{bmatrix} \\
W(k) &= \Theta^{-1}(k) [I \quad \theta_1] \\
e(k) &= \Theta^{-1}(k) \begin{bmatrix}
-x(k) - \hat{x}(k|k-1) \\
r(k)
\end{bmatrix}
\end{align*}
\]

(20)
(21)
(22)

with \( d(k) = [d_1(k), d_2(k), \ldots, d_L(k)]^T \), \( W(k) = [w_1(k), w_2(k), \ldots, w_L(k)]^T \), \( e(k) = [e_1(k), e_2(k), \ldots, e_L(k)]^T \), and \( L = m + n \).
**B. Derivation of MEE-KF**

Based on (14), the cost function of MEE-KF is given by

$$J_L(x(k)) = \frac{1}{L} \sum_{i=1}^{L} \sum_{j=1}^{L} G_\sigma(e_j(k) - e_i(k)).$$  
(23)

Then, the optimal solution to $\hat{x}(k)$ is achieved by maximizing the cost function (23), that is

$$\hat{x}(k) = \arg \max_{x(k)} J_L(x(k))$$

$$= \arg \max_{x(k)} \frac{1}{L^2} \sum_{i=1}^{L} \sum_{j=1}^{L} G_\sigma(e_j(k) - e_i(k)).$$  
(24)

Setting the gradient of the cost function $J_L(x(k))$ regarding $x(k)$ to zero, we have

$$\frac{\partial}{\partial x(k)} J_L(x(k))$$

$$= \frac{1}{L^2 \sigma^2} \sum_{i=1}^{L} \sum_{j=1}^{L} \left( \frac{e_j(k) - e_i(k)}{G_\sigma(e_j(k) - e_i(k))} \right) w_j(k) - w_i(k)$$

$$= \Gamma_1 - \Gamma_2 - \Gamma_3 + \Gamma_4 = 2\Gamma_1 - 2\Gamma_3 = 0$$  
(25)

where

$$\Gamma_1 = \frac{1}{L^2 \sigma^2} \sum_{i=1}^{L} \sum_{j=1}^{L} e_j(k)G_\sigma(e_j(k) - e_i(k))w_j(k)^T$$

$$\Gamma_2 = \frac{1}{L^2 \sigma^2} \sum_{i=1}^{L} \sum_{j=1}^{L} e_j(k)G_\sigma(e_j(k) - e_i(k))w_j(k)^T$$

$$\Gamma_3 = \frac{1}{L^2 \sigma^2} \sum_{i=1}^{L} \sum_{j=1}^{L} e_i(k)G_\sigma(e_j(k) - e_i(k))w_j(k)^T$$

$$\Gamma_4 = \frac{1}{L^2 \sigma^2} \sum_{i=1}^{L} \sum_{j=1}^{L} e_i(k)G_\sigma(e_j(k) - e_i(k))w_j(k)^T.$$  
(26)

By using the matrix to denote the double summation, (25) can be rewritten as

$$\Gamma_1 - \Gamma_3 = \frac{1}{L^2 \sigma^2} W(k)^T \Phi(k) e(k) - \frac{1}{L^2 \sigma^2} W(k)^T \Phi(k) e(k) = 0$$  
(30)

where

$$[\Phi(k)]_{ij} = G_\sigma(e_j(k) - e_i(k))$$

$$\Phi(k) = \text{diag} \left( \sum_{i=1}^{L} G_\sigma(e_1(k) - e_i(k)), \ldots, \sum_{i=1}^{L} G_\sigma(e_L(k) - e_i(k)) \right).$$  
(31)

From (30), $x(k)$ can be solved by a fixed-point iteration

$$\hat{x}(k) = g(\hat{x}(k))$$

$$= (W(k)^T \Lambda(k) W(k))^{-1} (W(k)^T \Lambda(k) d(k))$$  
(33)

where

$$\Lambda(k) = \Psi(k) - \Phi(k) = \begin{bmatrix} \Lambda_x(k) & \Lambda_{xy}(k) \\ \Lambda_{yx}(k) & \Lambda_y(k) \end{bmatrix}$$  
(34)

$$\Lambda_x(k) = \begin{pmatrix} \Lambda_{x_1}(k), & \ldots, & \Lambda_{x_L}(k) \end{pmatrix}$$

$$\Lambda_{xy}(k) = \begin{pmatrix} \Lambda_{xy_1}(k), & \ldots, & \Lambda_{xy_L}(k) \end{pmatrix}$$

$$\Lambda_y(k) = \begin{pmatrix} \Lambda_{y_1}(k), & \ldots, & \Lambda_{y_L}(k) \end{pmatrix}.$$  
(35)

According to (20), (21), and (34), we arrive at

$$W(k)^T \Lambda(k) D(k)$$

$$= \left[ \Theta^{-1}_p(k[k-1])^T \Lambda_x + H^T \Theta^{-1}_r(k)^T \Lambda_y \right] \Theta^{-1}_p(k[k-1])$$

$$+ \left[ \Theta^{-1}_p(k[k-1])^T \Lambda_{xy} + H^T \Theta^{-1}_r(k)^T \Lambda_y \right] \Theta^{-1}_r(k)^T \Lambda_y.$$  
(36)

By (33) and (39) can be rewritten as

$$\hat{x}(k) = (\Omega_1 + \Omega_2 \Omega_3)^{-1} \Omega_1 \hat{x}(k[k-1]) + \Omega_2 y(k)$$

$$= (\Omega_1 + \Omega_2 \Omega_3)^{-1} \Omega_1 \hat{x}(k[k-1])$$

$$+ (\Omega_1 + \Omega_2 \Omega_3)^{-1} \Omega_2 y(k)$$  
(37)

where

$$\Omega_1 = \left[ \Theta^{-1}_p(k[k-1])^T \Lambda_x + H^T \Theta^{-1}_r(k)^T \Lambda_y \right] \Theta^{-1}_p(k[k-1])$$

$$\Omega_2 = \left[ \Theta^{-1}_p(k[k-1])^T \Lambda_{xy} + H^T \Theta^{-1}_r(k)^T \Lambda_y \right] \Theta^{-1}_r(k)^T \Lambda_y.$$  
(38)

By using the matrix inversion lemma

$$(A + BCD)^{-1} = A^{-1} - A^{-1} B (C^{-1} + DA^{-1} B)^{-1} D A^{-1}$$  
(39)

with the identifications

$$\Omega_1 \rightarrow A, \Omega_2 \rightarrow B, I_L \rightarrow C, \Omega_3 \rightarrow D$$  
(40)

one can reformulate (40) as

$$\hat{x}(k) = \hat{x}(k[k-1]) + K(k)(y(k) - H \hat{x}(k[k-1]))$$  
(41)

where

$$K(k) = \left[ \bar{P}(k[k-1]) + H^T \bar{P}(k[k-1]) \right]^{-1}$$  
(42)

$$\bar{P}(k[k-1]) = \bar{P}(k[k-1]) + H^T \bar{P}(k[k-1]) H$$  
(43)
Algorithm 1 MEE-KF

Step 1: Initialize the state priori estimate \( \hat{x}(k|0) \) and state prediction error covariance matrix \( P(k|0) \); set a proper kernel size \( \sigma \) and a small positive number \( \varepsilon \).

Step 2: Use Eqs. (6) and (7) to obtain \( \hat{x}(k|k-1) \) and \( P(k|k-1) \), respectively; use the Cholesky decomposition of \( \Theta(k) \) to obtain \( \Theta_p(k|k-1) \) and \( \Theta_r(k) \); use Eqs. (20) and (21) to obtain \( d(k) \) and \( W(k|-1) \), respectively.

Step 3: Let \( t = 1 \) and \( \tilde{x}(k|0) = \tilde{x}(k|k-1) \), where \( \tilde{x}(k) \) denotes the estimated state at the fixed-point iteration \( t \).

Step 4: Use available measurements \( \{ \gamma(k) \}_{k=1}^{T} \) to update:

\[
\hat{x}(k|t) = \hat{x}(k|k-1) + \tilde{K}(k)[\gamma(k) - H\tilde{x}(k|k-1)]
\]

with

\[
\tilde{e}_i(k) = d_i(k) - w(k)\hat{x}(k)_{t-1}
\]

\[
\tilde{\Lambda}(k) = \tilde{\Psi}(k) - \tilde{\Phi}(k) = \begin{bmatrix} \tilde{\Lambda}_x(k) & \tilde{\Lambda}_y(k) \\ \tilde{\Lambda}_y(k) & \tilde{\Lambda}_y(k) \end{bmatrix}
\]

\[
\tilde{P}(k|k-1) = \Theta_p^{-1}(k|k-1)^T\tilde{\Lambda}_x(k)\Theta_p^{-1}(k|k-1)
\]

\[
\tilde{P}_{xy}(k|k-1) = \Theta_r^{-1}(k|k-1)^T\tilde{\Lambda}_{xy}(k)\Theta_r^{-1}(k|k-1)
\]

\[
\tilde{R}(k) = \Theta_r^{-1}(k|k-1)^T\tilde{\Lambda}_y(k)\Theta_r^{-1}(k|k-1)
\]

\[
\tilde{K}(k) = \begin{bmatrix} \tilde{P}(k|k-1) + H^T\tilde{R}(k) \\ + (\tilde{P}_{xy}(k|k-1) + H^{TT}\tilde{R}(k))H \\ \end{bmatrix}^{-1} \begin{bmatrix} \tilde{P}_{xy}(k|k-1) + H^{TT}\tilde{R}(k) \end{bmatrix}
\]

Step 5: Compare \( \hat{x}(k) \) and \( \hat{x}(k)_{t-1} \):

\[
\frac{|\hat{x}(k) - \hat{x}(k)_{t-1}|}{|\hat{x}(k)_{t-1}|} \leq \varepsilon.
\]

If the above condition holds, set \( \hat{x}(k) = \hat{x}(k)_{t} \); otherwise, \( t + 1 \rightarrow t \) and continue to Step 6.

Step 6: Update \( k + 1 \rightarrow k \) and the posterior error covariance matrix by

\[
P(k) = [I - \tilde{K}(k)H]P(k|k-1)[I - \tilde{K}(k)H]^T + \tilde{K}(k)R(k)\tilde{K}(k)^T
\]

and return to Step 2.

### IV. COMPUTATIONAL COMPLEXITY

This section provides the comparison of the computational complexities of KF, maximum correntropy KF (MCKF) [18] and MEE-KF in terms of the floating-point operations.

The KF updates with (6)–(10), and the corresponding floating-point operations are given in Table I. From Table I, we can conclude that the computational complexity of KF is

\[
C_{\text{KF}} = 8n^3 + 10n^2m - n^2 + 6nm^2 - n + O(m^3).
\]

According to [18], the computational complexity of MCKF is

\[
C_{\text{MCKF}} = (2T + 8)n^3 + (4T + 6)n^2m + (2T - 1)n^2
+ (4T + 2)nm^2 + (3T - 1)nm + (4T - 1)n
+ 2Tm^2 + 2Tm + TO(n^3) + 2TO(m^3)
\]

where \( T \) denotes the fixed-point iteration number, which is relatively small in general as shown in simulations in Section VII.

The updates of MEE-KF involve (6), (7), (46)–(53), and (55), and the corresponding floating-point operations are shown in Table I. According to Table I, the computational complexity of MEE-KF is

\[
C_{\text{MEE-KF}} = (7T + 8)n^3 + (7T)m^3 + (19T + 6)n^2m - n^2
+ (15T + 2)nm^2 + TO(n^3) + TO(mnm) + Tm
+ (5T - 1)n + (6T - 1)nm + 2TO(m^3).
\]

The MEE-KF has an additional computational burden induced by the error entropy functions in comparison to KF, and has a slightly higher computational complexity than MCKF. In the sense of order of magnitude, the computational complexities of the MEE-KF, MCKF, and KF have no significant difference.

### V. CONVERGENCE ISSUE

This section provides a sufficient condition to ensure the convergence of the fixed-point iterations in MEE-KF, where the proof is similar to [39] and thus will not be provided here.
First, from (33), we can rewrite
\[ g(x(k)) = M_{ww}^{-1}M_{dw} \] (59)
with
\[ M_{ww}^{-1} = \sum_{i=1}^{L} \sum_{j=1}^{L} G_o(e_j(k) - e_i(k))[w_j(k) - w_i(k)]^T \]
and
\[ M_{dw} = \sum_{i=1}^{L} \sum_{j=1}^{L} G_o(e_j(k) - e_i(k))[d_j(k) - d_i(k)]. \]

Thus, an \( n \times n \) Jacobian matrix of \( g(x(k)) \) with respect to \( x(k) \) gives
\[
\nabla x(k)g(x(k)) = \frac{\partial}{\partial x(k)} M_{ww}^{-1} M_{dw} = \left[ \frac{\partial g(x(k))}{\partial x_1(k)} \frac{\partial g(x(k))}{\partial x_2(k)} \cdots \frac{\partial g(x(k))}{\partial x_L(k)} \right]
\]
\[ = -\frac{M_{ww}^{-1}}{2\sigma^2} \left( \sum_{i=1}^{L} \sum_{j=1}^{L} \xi_i(k) [w_j(k) - w_i(k)]^T \right) g(x(k)) \]
\[ + \frac{M_{ww}^{-1}}{2\sigma^2} \left( \sum_{i=1}^{L} \sum_{j=1}^{L} \xi_i(k) [d_j(k) - d_i(k)]^T \right) \]
(60)
with \( \xi_i(k) = (e_j(k) - e_i(k))[w_j(k) - w_i(k)]G_o(e_j(k) - e_i(k)) \). Define \( \| \cdot \|_p \) as an \( \ell_p \)-norm (\( p \geq 1 \)) of a vector or an induced norm of a matrix as \( \| A \|_p = \max_{\| x \| \neq 0} \| AW \|_p / \| W \|_p \). According to the proof in [39], the following theorem holds.

**Theorem 1**: If the kernel size satisfies \( \sigma \geq \max(\sigma_1, \sigma_2) \), we have
\[
\begin{align*}
\| g(x(k)) \|_1 & \leq \beta \\
\| \nabla x(k)g(x(k)) \|_1 & \leq \alpha < 1
\end{align*}
\]
where
\[
\beta > \rho = \sqrt{n} \frac{\sum_{i=1}^{L} \sum_{j=1}^{L} [d_j(k) - d_i(k)]}{\lambda_{\min} \left( \sum_{i=1}^{L} \sum_{j=1}^{L} [w_j(k) - w_i(k)]^T \right)}
\]
\( \lambda_{\min} \) is the minimum eigenvalue of the matrix \( M_{ww} \); and \( \sigma_1 \) and \( \sigma_2 \) are the solutions of \( \| g(x(k)) \|_1 = \beta \) and \( \| \nabla x(k)g(x(k)) \|_1 = \alpha \), respectively, i.e.,
\[
\| g(x(k)) \|_1 = \sqrt{n} \frac{\sum_{i=1}^{L} \sum_{j=1}^{L} [d_j(k) - d_i(k)]}{\lambda_{\min} \left( \sum_{i=1}^{L} \sum_{j=1}^{L} [w_j(k) - w_i(k)]^T \right)}
\]
(62)
with \( \xi_2(k) = G_o(\beta[\| w_j(k) - w_i(k) \|_1 + |d_j(k) - d_i(k)|] \), and
\[
\| \nabla x(k)g(x(k)) \|_1 = \gamma \sqrt{n} \frac{\sum_{i=1}^{L} \sum_{j=1}^{L} \xi_2(k) [w_j(k) - w_i(k)]^T}{\lambda_{\min} \left( \sum_{i=1}^{L} \sum_{j=1}^{L} [w_j(k) - w_i(k)]^T \right)}
\]
(63)
with
\[
\gamma = \sum_{i=1}^{L} \sum_{j=1}^{L} \left( \beta \| w_j(k) - w_i(k) \|_1 \| w_j(k) - w_i(k) \|_1 \right)^\frac{1}{2} \left( \beta \| w_j(k) - w_i(k) \|_1 \| w_j(k) - w_i(k) \|_1 \right)^\frac{1}{2} + \| d_j(k) - d_i(k) \|.
\]

According to Theorem 1 and Banach fixed-point theorem, the fixed-point algorithm induced by (33) can be guaranteed to converge to a unique fixed point in the range \( x(k) \in \{ \| x(k) \|_1 \leq \beta \} \) provided that the kernel size \( \sigma \) is larger than a certain value and an initial state vector satisfies \( \{ \| x(k) \|_1 \leq \beta \} \).

VI. **Minimum Error Entropy Extended Kalman Filter**

For a nonlinear system, the state and measurement equations can be described by
\[
x(k) = f(x(k - 1)) + q(k - 1)
\]
(64)
\[
y(k) = h(x(k)) + r(k)
\]
(65)
where \( f(\cdot) \) and \( h(\cdot) \) are the state-transition function and measurement (observation) function, respectively. The EKF is an important tool for dealing with the state estimation of a nonlinear system via approximating the nonlinear system by linear model. In EKF, the nonlinear functions \( f(x(k - 1)) \) and \( h(x(k)) \) can be approximated by the first-order Taylor series expansion at \( \hat{x}(k) \) and \( \hat{x}(k - 1) \), i.e.,
\[
f(x(k - 1)) \approx f(\hat{x}(k - 1)) + F(k - 1)(x(k - 1) - \hat{x}(k - 1))
\]
(66)
\[
h(x(k)) \approx h(\hat{x}(k - 1)) + H(k)(x(k) - \hat{x}(k - 1))
\]
(67)
where \( F(k - 1) \) and \( H(k) \) are the Jacobian matrices of \( f(\cdot) \) and \( h(\cdot) \), i.e.,
\[
F(k - 1) = \frac{\partial f(x(k - 1))}{\partial x} |_{x = \hat{x}(k - 1)}
\]
(68)
\[
H(k) = \frac{\partial h(x(k))}{\partial x} |_{x = \hat{x}(k - 1)}
\]
(69)

By substituting (66) and (67) into (64) and (65), the prediction and update equations of EKF can be obtained
\[
\hat{x}(k|k - 1) = f(\hat{x}(k - 1))
\]
(70)
\[
P(k|k - 1) = F(k - 1)P(k - 1)F^T(k - 1) + Q(k - 1)
\]
(71)
\[
K(k) = P(k|k - 1)H^T(k)[H(k)P(k|k - 1)H^T(k) + R(k)]^{-1}
\]
(72)
\[
\hat{x}(k) = \hat{x}(k|k - 1) + K(k)[y(k) - h(\hat{x}(k|k - 1))]
\]
(73)
\[
P(k) = [I - K(k)H(k)]P(k|k - 1)[I - K(k)H(k)]^T + K(k)R(k)K^T
\]
(74)
where \( K(k) \) is the gain of EKF. Similar to MEE-KF and maximum correntropy EKF (MCEKF) [28], the augmented model of MEE-EKF should be established. By substituting (67) into (65), the measurement equation is approximated by
\[
y(k) \approx h(\hat{x}(k|k - 1)) + H(k)(x(k) - \hat{x}(k|k - 1)) + r(k).n
\]
(75)

Then, replacing \( y(k) \) in (16) with that in (75), one can obtain the following augmented model:
\[
\begin{bmatrix}
\hat{x}(k|k - 1) \\
y(k) - h(\hat{x}(k|k - 1)) + H(k)(x(k) - \hat{x}(k|k - 1))
\end{bmatrix}
= \begin{bmatrix}
I_n \\
H(k)
\end{bmatrix} \begin{bmatrix}
x(k) + \mu(k)
\end{bmatrix}
\]
(76)
Algorithm 2 MEE-EKF

Step 1: Similar to the Step 1 in Algorithm 1.
Step 2: Use Eqs. (68), (70), and (71) to obtain \( \hat{x}(k|k-1) \) and \( P(k|k-1) \); use the Cholesky decomposition of \( \Theta(k) \) to obtain \( \Theta_p(k|k-1) \) and \( \Theta_e(k) \); use Eqs. (78) and (79) to obtain \( d(k) \) and \( W(k|k-1) \), respectively.
Step 3: Similar to the Step 3 in Algorithm 1.
Step 4: Use the available measurements \( \{y(k)\}_{k=1}^{N} \) to update:

\[
\hat{x}(k) = \hat{x}(k|k-1) + \tilde{K}(k)[y(k) - h(\hat{x}(k|k-1))] \tag{80}
\]

\[
\tilde{e}_t(k) = d_t(k) - w(k)\hat{x}(k|k-1) \tag{81}
\]

\[
\tilde{K}(k) = \left[ \tilde{P}(k|k-1) + H^T(k)\tilde{P}_y(k|k-1) + \tilde{P}_y(k|k-1)H(k) \right]^{-1}
\times \left[ \tilde{P}_y(k|k-1)H(k)\tilde{R}(k) \right], \tag{82}
\]

where the forms of \( \tilde{A}(k), \tilde{P}(k|k-1), \tilde{P}_y(k|k-1), \tilde{P}_y(k|k-1) \) and \( \tilde{R}(k) \) are the same as Eqs. (48)-(52), respectively.
Step 5: Similar to the Step 1 in Algorithm 1.
Step 6: Update \( k+1 \rightarrow k \) and the posterior covariance matrix:

\[
P(k) = [I - \tilde{K}(k)H(k)]P(k|k-1)[I - \tilde{K}(k)H(k)]^T
+ \tilde{K}(k)\tilde{R}(k)\tilde{R}^T(k) \tag{83}
\]

and return to Step 2.

where \( \mu(k) \) has the same form with (17). Similar to (19), we have

\[
d(k) = W(k)x(k) + e(k) \tag{77}
\]

where

\[
d(k) = \Theta^{-1}(k)\begin{bmatrix} \hat{x}(k|k-1) \\ y(k) - h(\hat{x}(k|k-1)) + H(k)\hat{x}(k|k-1) \end{bmatrix} \tag{78}
\]

\[
W(k) = \Theta^{-1}(k)\begin{bmatrix} I_n \\ H(k) \end{bmatrix} \tag{79}
\]

The forms of \( \Theta(k) \) and \( e(k) \) are the same as (18) and (22), respectively. Therefore, using (77) and the derivation similar to that of MEE-KF, one can obtain the MEE-EKF, as provided in Algorithm 2.

VII. EXPERIMENTAL RESULTS

In this section, the performances of the proposed MEE-KF and MEE-EKF are demonstrated in different scenarios through three examples, i.e., land vehicle navigation, tracking of autonomous driving [46]-[48], and prediction of infectious disease epidemics [49]. The results are compared with KF and MCKF in linear model, and EKF and MCEKF in nonlinear model. The performance evaluation index is defined by the mean square error (MSE)

\[
MSE = \frac{1}{l_{mc}} \sum_{l=1}^{l_{mc}} ME(l) \tag{84}
\]

where \( ME(l) = (1/N) \sum_{k=1}^{N} ||x(k) - \hat{x}(k)||^2; \hat{x}(k) \) is the estimate of \( x(k); \) and \( l_{mc} \) and \( N \) are the Monte Carlo runs and number of samples. In the simulations, the MSE is computed by averaging over 100 independent Monte Carlo runs.

A. Land Vehicle Navigation

In the first example, we assume a linear land vehicle navigation problem to illustrate the performance of MEE-KF.

1) Model: The state and measurement equations are described by

\[
x(k) = \begin{bmatrix} 1 & 0 & \Delta T & 0 \\ 0 & 1 & 0 & \Delta T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} x(k-1) + q(k-1) \tag{85}
\]

\[
y(k) = \begin{bmatrix} -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \end{bmatrix} x(k-1) + r(k) \tag{86}
\]

where \( x(k) = [x_1(k) \ x_2(k) \ x_3(k) \ x_4(k)]^T \) is the state vector with components being the north position, the east position, the north velocity, and the east velocity. The variables \( \theta = \pi/3 \) and \( \Delta T = 0.3 \) are the direction, and time interval of vehicle, respectively. The real state \( x(0) \), the prior estimate \( \hat{x}(1|0) \) and the corresponding error covariance matrix \( P(1|0) \) are initialized by

\[
x(0) = [0 \ 10\tan\theta \ 10]^T \tag{87}
\]

\[
\hat{x}(1|0) = [1 \ 1 \ 1 \ 1]^T \tag{88}
\]

\[
P(1|0) = \text{diag}[900 \ 900 \ 4 \ 4]. \tag{89}
\]

In the dynamic system, the process noises are assumed to be of the Gaussian distribution with \( q(k) \sim \mathcal{N}(0,0.001) \), where \( \mathcal{N}(0,0.001) \) denotes a zero-mean Gaussian distribution with variance 0.001.

2) Estimation Results of Different Algorithms: The following simulation results are presented to confirm the desirable performance of the MEE-KF. Here, we use 3000 samples (time steps) to calculate MSE.

First, the influence of the measurement noises on MSE is investigated. For the distribution of \( r(k) \), we consider four cases: 1) Gaussian noise: \( r(k) \sim \mathcal{N}(0,0.05) \); 2) Gaussian noise with outliers: \( r(k) \sim 0.99\mathcal{N}(0,0.001) + 0.01\mathcal{N}(0,1000) \), where \( \mathcal{N}(0,1000) \) is used to generate outliers; 3) mixture Gaussian noise: \( r(k) \sim 0.01\mathcal{N}(-0.1,0.001) + 0.99\mathcal{N}(0.1,1000) \); and 4) mixture Gaussian noise with outliers: \( r(k) \sim 0.48\mathcal{N}(-0.1,0.001) + 0.04\mathcal{N}(0,1000) + 0.48\mathcal{N}(0.1,0.001) \); the process noise is \( q(k) \sim \mathcal{N}(0,0.01) \) in this case.

Table II gives the estimation results of KF, MCKF, and MEE-KF in the presence of different measurement noises. The parameter settings of MCKF and MEE-KF are given in Table III, and all these parameters are set by trials to obtain the best performance. From Table II, we observe: 1) under the Gaussian noises, the performance of MCKF approaches that of KF, and KF preforms the optimal performance than MEE-KF and 2) when large outliers or mixture Gaussian noise [case 2, case 3], and case 4) are included in the measurement noise, the MCKF and MEE-KF show better performance than KF, and the MEE-KF achieves the best performance among the three algorithms.
Next, we investigate the influence of kernel size on the estimation performance. Table IV shows the MSE comparison of different filters with different kernel sizes, where the measurement noise is the same as case 4). The thresholds in MCKF and MEE-KF are set to $10^{-6}$. From Table IV, we find out that if the kernel size is too small or too large, the performance of MCKF and MEE-KF will become worse. Especially, for the small kernel size, the MCKF may diverge, while MEE-KF is more stable compared with MCKF. The desirable performance can be achieved by MEE-KF when the kernel size is around $1.5$.

Finally, the computing time of different filters at each iteration is summarized in Table V, which is measured with MATLAB 2016b running on i5-4590 and 3.30 GHz CPU. We see that the MEE-KF can achieve better performance at the cost of slightly higher computational burden than MCKF.

### B. Tracking of Autonomous Driving

This experiment shows the tracking of autonomous driving with the Kalman fusion algorithms [46]. The data source is from Udacity course self-driving car [47], which is measured by the lidar and radar sensors. In general, different sensors provide different types of information with different accuracies, especially, in different weather conditions [48]. In this example, the lidar sensor provides a good resolution about the position, while the radar sensor provides better accuracy about the velocity in poor weather compared to lidar. To merge the information from different sensors, the Kalman fusion algorithms are applied to different sensors to estimate effectively the trajectory of vehicle. Fig. 1 shows the Kalman fusion model, where $x(k-1)$ denotes the state variables at time $k-1$; $y_L(k)$ and $y_R(k)$ are the observations from lidar and radar sensors; $\hat{x}_L(k)$ and $\hat{x}_R(k)$ are the estimated states; and $\hat{x}(k)$ is the fusion state at time $k$. The detailed descriptions are presented in the following.

1) **Model:** For the tracked vehicle, the state can be described by a 4-dimensional vector $x(k) = [p_x(k), p_y(k), v_x(k), v_y(k)]^T$, where $p_x(k)$ and $p_y(k)$ denote the position information of $x$ and $y$ coordinates and $v_x(k)$ and $v_y(k)$ are the corresponding velocity information. The state equation can be expressed by

$$x(k) = \begin{bmatrix} 1 & 0 & \Delta T & 0 \\ 0 & 1 & 0 & \Delta T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} x(k-1) + q(k)$$

where $\Delta T = 0.1$ s is the time interval and $q(k)$ is the process noise with covariance matrix

$$Q(k) = \begin{bmatrix} \Delta T^2 & 0 & \Delta T^3 & 0 \\ 0 & \Delta T^2 & 0 & \Delta T^3 \\ \Delta T^3 & 0 & \Delta T^2 & 0 \\ 0 & \Delta T^3 & 0 & \Delta T^2 \end{bmatrix}.$$ 

The prior error covariance matrix is initialized as $P(1|0) = \text{diag}[1, 1, 1000, 1000]$. For the measurement equation, we consider two cases.

**Case 1:** When the measurements $y_L(k) = [p_x(k), p_y(k)]^T$ are acquired from the lidar sensor, the measurement equation can be described as

$$y_L(k) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} x(k) + r_L(k)$$

where $r_L(k)$ is the measurement noise.
TABLE IV
ESTIMATION RESULTS OF DIFFERENT ALGORITHMS WITH DIFFERENT KERNEL SIZES

| Algorithms | σ    | MSE of x1     | MSE of x2     | MSE of x3     | MSE of x4     |
|------------|------|---------------|---------------|---------------|---------------|
| KF         | N/A  | 1.2031 ± 0.3572 | 1.1811 ± 0.2139 | 0.3901 ± 0.6712 | 0.3622 ± 0.2057 |
| MCKF       | σ = 3.0 | N/A         | N/A           | N/A           | N/A           |
|            | σ = 4.0 | 0.8213 ± 0.5663 | 0.8437 ± 0.3573 | 0.3773 ± 0.7391 | 0.3298 ± 0.2330 |
|            | σ = 5.0 | 0.8423 ± 0.4875 | 0.8121 ± 0.2310 | 0.3610 ± 0.0495 | 0.3260 ± 0.0352 |
|            | σ = 10.0 | 1.0576 ± 0.3515 | 1.0366 ± 0.2103 | 0.3766 ± 0.6477 | 0.3471 ± 0.2066 |
| MEE-KF     | σ = 1.0 | 1.1688 ± 0.2709 | 1.1748 ± 0.2570 | 0.3895 ± 0.6414 | 0.3613 ± 0.2667 |
|            | σ = 1.5 | 0.6258 ± 0.2942 | 0.5376 ± 0.4943 | 0.3296 ± 0.4930 | 0.2896 ± 0.2441 |
|            | σ = 2.0 | 0.6833 ± 0.3411 | 0.6671 ± 0.6711 | 0.3490 ± 0.4930 | 0.4213 ± 0.2453 |
|            | σ = 3.0 | 0.8410 ± 0.2367 | 0.9269 ± 0.2604 | 0.3895 ± 0.4748 | 0.5876 ± 0.3474 |
|            | σ = 5.0 | 1.0381 ± 0.2719 | 1.1381 ± 0.3910 | 0.4391 ± 0.4750 | 0.7279 ± 0.2161 |
|            | σ = 10 | 1.1912 ± 0.2995 | 1.2142 ± 0.3956 | 0.4628 ± 0.4698 | 0.7814 ± 0.2204 |

TABLE V
AVERAGE COMPUTING TIME (SEC) FOR EACH ITERATION

| Algorithms | KF | MCKF | MEE-KF |
|------------|----|------|--------|
| Computing time (sec) | 0.000066 | 0.000204 | 0.000232 |

TABLE VI
ESTIMATION RESULTS OF DIFFERENT ALGORITHMS FOR TRAJECTORY 1 AND TRAJECTORY 2

| Algorithms | MSE of trajectory 1 | MSE of trajectory 2 |
|------------|----------------------|----------------------|
| KF/EKF     | 0.5408               | 0.008387             |
| MCKF/MCEKF | 0.5270               | 0.008184             |
| MEE-KF/MEE-EKF | 0.1554               | 0.007789             |

where $r_L(k)$ is the measurement noise from lidar sensor with covariance matrix $R_L(k) = \text{diag}[0.0025, 0.0025]$. For the linear state-space model in (90) and (91), the KF, MCKF and MEE-KF can be employed.

**Case 2.** The measurements $y_R(k) = [\rho(k), \varphi(k), \bar{\rho}(k)]^T$ from the radar sensor involve three components in polar coordinates, i.e., the range $\rho(k)$, the angle $\varphi(k)$ between $\rho(k)$ and $x$ coordinates axis, and the range rate $\bar{\rho}(k)$. The measurement equation can then be established by

$$y_R(k) = h(x(k)) + r_R(k)$$

where the nonlinear function

$$h(x(k)) = \begin{bmatrix} \sqrt{p_x^2(k) + p_y^2(k)} \\ \arctan\left(\frac{p_y(k)}{p_x(k)}\right) \\ p_x(k)v_x(k) + p_y(k)v_y(k) \\ \sqrt{p_x^2(k) + p_y^2(k)} \end{bmatrix}$$

can be derived by a mapping from the polar coordinates to the cartesian coordinates; $r_R(k)$ is the measurement noise from radar sensor with covariance matrix $R_R(k) = \text{diag}[0.09, 0.05, 0.09]$. To estimate the state $x(k)$ in the cartesian coordinates, one should transform the measurements $y_R(k) = [\rho(k), \varphi(k), \bar{\rho}(k)]^T$ as state vector $x(k) = [p_x(k), p_y(k), v_x(k), v_y(k)]^T$.

$$p_x(k) = \rho(k) \cos(\varphi(k)), \quad p_y(k) = \rho(k) \sin(\varphi(k))$$

$$v_x(k) = \bar{\rho}(k) \cos(\varphi(k)), \quad v_y(k) = \bar{\rho}(k) \sin(\varphi(k)).$$

2) Estimation Results of Different Algorithms: The compared algorithms are performed for two trajectories, i.e., trajectory 1 and trajectory 2. Figs. 2 and 3 show the tracking results of trajectory 1 and trajectory 2, respectively. Table VI presents the MSE of different Kalman fusion algorithms, i.e., the KF/EKF, MCKF/MCEKF, and MEE-KF/MEE-EKF. In this simulation, the kernel sizes are set to 1.66 in MEE-EKF, 20.0 in MEE-KF, 15.0 in MCE-EKF, and 20.0 in MCE-KF. From Table VI, one can see that the MEE-KF/MEE-EKF fusion method exhibits the best performance among all compared algorithms.
C. Prediction of Infectious Disease Epidemics

In this part, the prediction of epidemic [49] is used to validate the effectiveness of the proposed MEE-EKF. The disease control center (DCC) provides the influenza-like-illness (ILI) data [49] of the United States, a traditional information source for the epidemic estimation.

1) Model: The state process of an epidemic is based on the susceptible-infected-recovered (SIR) model [49], given by

\[
\begin{align*}
    s(k+1) &= s(k) - \delta_1 s(k)i(k) \quad (95) \\
    i(k+1) &= i(k) + \delta_1 s(k)i(k) - \delta_2 i(k) \quad (96) \\
    b(k+1) &= b(k) + \delta_2 i(k) \quad (97)
\end{align*}
\]

where \(s(k), i(k), \) and \(b(k)\) are the susceptible, infected, and recovered crowd density, respectively, with \(s(k) + i(k) + b(k) = 1\) for all \(k\). The parameter \(\delta_1\) denotes the rate about the spread of illness and \(\delta_2\) is the rate about recovery from infection. Since the recovered \(b(k)\) is obtained by \(b(k) = 1 - s(k) - i(k)\), the state of epidemic at time \(k\) can be described by \(x(k) = [s(k), i(k)]^T\). Therefore, the state equation can be written by

\[
x(k) = f(x(k-1)) + q(k) \quad (98)
\]

where the nonlinear function \(f(x(k-1))\) is given in (95) and (96), and \(Q(k) = 10^{-8}I_n\).

In the state space model, the measurement equation specifies how the observed data depend on the state of epidemic. When monitoring an epidemic, the unknown \(s(k), i(k), \) and \(b(k)\) are regarded as hidden states of the model, and the observed data are acquired via syndromic surveillance system. Due to only \(i(k)\) can be obtained by the syndromic surveillance system, the measurement equation can be established by

\[
y(k) = Hx(k) + r(k) \quad (99)
\]

where the matrix \(H = [0 \ 1]\) denotes the measurement matrix and \(R(k) = 10^{-6}\).

2) Estimation Results of Different Algorithms: Fig. 4 plots 52 weeks density of people infected in 2010 year. Table VII gives the estimated MSE by EKF, MCEKF, and MEE-EKF algorithms. In Fig. 4, a peak around the 25th week occurs, which means that the number of people infected arrives at maximum. In the SIR model, the contact rate \(\delta_1\) and recovery time \(\delta_2\) are set to 0.35 and 0.11, respectively. The kernel sizes of MEE-EKF and MCEKF are 20.0, and 16.0, respectively. From Fig. 4 and Table VII, the MEE-EKF provides a much better prediction for the density of people infected than EKF and MCEKF.

VIII. Conclusion

In this article, we propose the MEE-KF with a fixed-point iteration. Unlike the original KF based on the well-known MMSE criterion and the MCKF based on the MCC, the MEE-KF is developed by using the MEE criterion as the optimality criterion. With an appropriate kernel size, the MEE-KF algorithm can achieve better performance than the KF and MCKF, especially, when the underlying system is disturbed by some complicated non-Gaussian noises. The computational complexity of the MEE-KF is provided, and a sufficient condition for the convergence of the fixed-point iteration is given. Furthermore, the MEE-EKF is also developed to deal with the problem of state estimation of a nonlinear system in non-Gaussian noises. Simulations on land vehicle navigation, tracking of autonomous driving and prediction of infectious disease epidemics have confirmed the desirable performance of the proposed MEE-KF and MEE-EKF. How to select the kernel size \(\sigma\) and how to adjust the error mean are the keys in practical applications. The future work will be concerned to deal with these problems.

References

[1] R. E. Kalman, “A new approach to linear filtering and prediction problem,” Trans. ASME D J. Basic Eng., vol. 82, no. 1, pp. 35–45, Mar. 1960.
[2] G. Welch and G. Bishop, An Introduction to the Kalman Filter, Univ. North Carolina at Chapel Hill, Chapel Hill, NC, USA, 2001.
[3] J. K. Uhlman, “Algorithms for multiple target tracking,” Amer. Sci., vol. 80, no. 2, pp. 128–141, Sep. 1997.
[4] J. Lu, T. Zhang, F. Hu, and Q. Hao, “Preprocessing design in pyroelectric infrared sensor-based human-tracking system: On sensor selection and calibration,” IEEE Trans. Syst., Man, Cybern., Syst., vol. 47, no. 2, pp. 263–275, Feb. 2017.
[5] S. Y. Chen, “Kalman filter for robot vision: A survey,” IEEE Trans. Ind. Electron., vol. 59, no. 11, pp. 4409–4420, Nov. 2012.
[6] X. Yang, W.-A. Zhang, and L. Yu, “A bank of decentralized extended information filters for target tracking in event-triggered WSNs,” IEEE Trans. Syst., Man, Cybern., Syst., to be published, doi: 10.1109/TSMC.2018.2883706.
[7] Y. Huang, Y. Zhang, Z. Wu, N. Li, and J. Chambers, “A novel adaptive Kalman filter with inaccurate process and measurement noise covariance matrices,” IEEE Trans. Autom. Control, vol. 63, no. 2, pp. 594–601, Feb. 2018.
[8] W. Li, G. Wei, D. Ding, Y. Liu, and F. E. Alsaadi, “A new look at boundedness of error covariance of Kalman filtering,” IEEE Trans. Syst., Man, Cybern., Syst., vol. 48, no. 2, pp. 309–314, Feb. 2018.
[9] S. J. Julier and K. Uhlmann, “A new extension of the Kalman filter to nonlinear systems,” in Proc. 11th Int. Symp. Aerosp./Defence Sens. Simulat. Controls, pp. 182–193, 1997.
Lujuan Dang received the B.S. degree in information science and technology from Northwest University, Xi’an, China, in 2015, and the M.S. degree in electronic and information engineering from Southwest University, Chongqing, China, in 2018. She is currently pursuing the Ph.D. degree with the Institute of Artificial Intelligence and Robotics, Xi’an Jiaotong University, Xi’an. Her current interests focus on adaptive filtering and information theoretic learning.

Yuantao Gu (S’02–M’03–SM’17) received the B.S. degree in electronic engineering from Xi’an Jiaotong University, Xi’an, China, in 1998, and the Ph.D. degree (with Hons.) in electronic engineering from Tsinghua University, Beijing, China, in 2003. He joined the Faculty of Tsinghua University in 2003, where he is currently a Tenured Associate Professor with the Department of Electronic Engineering. He was a Visiting Scientist with Microsoft Research Asia, Beijing, from 2005 to 2006, the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA, USA, from 2012 to 2013, and the Department of Electrical Engineering and Computer Science, University of Michigan at Ann Arbor, Ann Arbor, MI, USA, in 2015. His current research interests include high-dimensional statistics, sparse signal recovery, temporal-space and graph signal processing, related topics in wireless communications, and information networks.

Dr. Gu was a recipient of the Best Paper Award of IEEE GlobalSIP in 2015, the Award for Best Presentation of Journal Paper of ChinaSIP in 2015, and the Zhang Si-Ying (CCDC) Outstanding Young Paper Award (with his student) in 2017. He has been an Associate Editor of the IEEE Transactions on Signal Processing since 2013, was a Handling Editor for the EURASIP Digital Signal Processing from 2015 to 2017, and has been an Elected Member of the IEEE Signal Processing Theory and Methods Technical Committee since 2017.

Nanning Zheng (SM’94–F’06) received the graduation degree from the Department of Electrical Engineering, Xi’an Jiaotong University, Xi’an, China, in 1975, the M.S. degree in information and control engineering from Xi’an Jiaotong University in 1981, and the Ph.D. degree in electrical engineering from Keio University, Yokohama, Japan, in 1985. He is currently a Professor and the Director of the Institute of Artificial Intelligence and Robotics, Xi’an Jiaotong University. His current research interests include computer vision, pattern recognition, machine vision and image processing, neural networks, and hardware implementation of intelligent systems.

Prof. Zheng is the Chinese Representative on the Governing Board of the International Association for Pattern Recognition. He also serves as an Executive Deputy Editor of Chinese Science Bulletin. He became a member of the Chinese Academy of Engineering in 1999.

José C. Príncipe (M’83–SM’90–F’00) received the bachelor’s degree in electrical engineering from the University of Porto, Porto, Portugal, in 1972, and the master’s and Ph.D. degrees from the University of Florida, Gainesville, FL, USA, in 1979. He is currently a Distinguished Professor of Electrical and Computer Engineering and Biomedical Engineering with the University of Florida, Gainesville, FL, USA, where he teaches advanced signal processing, machine learning, and artificial neural networks modeling. He is the Eckis Professor of ECE and the Founder and the Director of the Computational NeuroEngineering Laboratory (CNEL), University of Florida, Gainesville, FL, USA. He is also an Adjunct Professor with the Institute of Artificial Intelligence and Robotics, Xi’an Jiaotong University, Xi’an, China. The CNEL Lab has been studying signal and pattern recognition principles based on information theoretic criteria (entropy and mutual information). He has authored or coauthored over 800 publications. He directed 95 Ph.D. dissertations and 65 master’s theses. He wrote in 2000 an interactive electronic book entitled Neural and Adaptive Systems (Wiley) and more recently coauthored several books on Brain Machine Interface Engineering (Morgan and Claypool), Information Theoretic Learning (Springer), and Kernel Adaptive Filtering (Wiley). His current research interests include processing of time varying signals with adaptive neural models.

Dr. Príncipe is a member of the Advisory Board of the University of Florida Brain Institute and the Advisory Board of the University of Florida Brain Institute. He was the Chair of the Technical Committee on Neural Networks of the IEEE Signal Processing Society, the President of the International Neural Network Society, and the Editor-in-Chief of the IEEE Transactions on Biomedical Engineering.