Kalman filtering based on the maximum correntropy criterion in the presence of non-Gaussian noise

Reza Izanloo
Dept. of Computer Engineering
Ferdowsi University of Mashhad, Iran
rezaizanloo_88@stu.um.ac.ir

Seyed Abolfazl Fakoorian
Dept. of Electrical Engr. and Computer Science
Cleveland State University
Cleveland, Ohio
s.fakoorian@csuohio.edu

Hadi Sadoghi Yazdi
Dept. of Computer Engineering
Ferdowsi University of Mashhad, Iran
h-sadoghi@um.ac.ir

Dan Simon
Dept. of Electrical Engr. and Computer Science
Cleveland State University
Cleveland, Ohio
d.j.simon@csuohio.edu

***Abstract*—State estimation in the presence of non-Gaussian noise is discussed. Since the Kalman filter uses only second-order signal information, it is not optimal in non-Gaussian noise environments. The maximum correntropy criterion (MCC) is a new approach to measure the similarity of two random variables using information from higher-order signal statistics. The correntropy filter (C-Filter) uses the MCC for state estimation. In this paper we first improve the performance of the C-Filter by modifying its derivation to obtain the modified correntropy filter (MC-Filter). Next we use the MCC and weighted least squares (WLS) to propose an MCC filter in Kalman filter form, which we call the MCC-KF. Simulation results show the superiority of the MCC-KF compared with the C-Filter, the MC-Filter, the unscented Kalman filter, the ensemble Kalman filter, and the Gaussian sum filter, in the presence of two different types of non-Gaussian disturbances (shot noise and Gaussian mixture noise).**

Keywords—state estimation; maximum correntropy criterion (MCC); Kalman filter; non-Gaussian noise

I. INTRODUCTION

Adaptive filters are powerful tools for state estimation and system identification and have been successfully applied in many fields, such as fault diagnosis, target tracking, and so on [1]. There are many adaptive filters that are based on state-space models. The Kalman filter is the most common adaptive filter because of its optimality, versatility, and simplicity [2], [3]. If the measurement and process noises are Gaussian, the Kalman filter is the optimal estimator. However, in non-Gaussian noise environments the performance of the Kalman filter can break down [4]. Many filtering methods have been presented to cope with non-Gaussian noise. In general, there are three approaches to improve the robustness of state estimation in the presence of non-Gaussian noise.

The first approach develops filters for systems with both outliers and non-Gaussian noise. In these filters, noise distributions such as heavy-tailed distributions and t-distributions are considered [5], [6]. However, these distributions can be difficult to analytically handle more than one dimension, which limits the applicability of these filters. Also, these filters require high computational effort and therefore are significantly more difficult to implement in real time than Kalman filters [7].

The second approach to handling non-Gaussian noise is the multiple-model (MM) filter [8]. In this approach, a non-Gaussian distribution is approximated with a finite sum of Gaussian distributions that represents different modes [9]. The state posterior of the probability density function (pdf) is then formed as a weighted sum of Gaussians. For instance, the Gaussian sum filter (GSF) uses a bank of Kalman filters [10], [11]. The computational complexity is the main drawback of MM filters since the number of modes increases exponentially with the number of filters in the bank [9].

The third approach to handling non-Gaussian noise involves sequential Monte Carlo (MC) sampling, which enables the approximation of any probability distribution [12]. Particle filters are commonly used MC techniques which approximate the posterior distribution of the state with a set of random samples with associated weights [9], [13]. The ensemble Kalman filter (EnKF) is a recursive estimator which is related to the particle filter and which approximates the state estimate with a finite set of randomly selected samples [14]. The unscented Kalman filter (UKF) is also similar to the particle filter, but uses a deterministic sampling technique to capture the mean and covariance of the state estimate with a minimal set of sample points, called sigma points [15]. Computational effort is the main problem of statistical sampling methods, and often results in unacceptable performance in real-time applications [7].

To reduce the aforementioned drawbacks, [16] introduces the correntropy filter (C-Filter) for state estimation and shows that the C-Filter outperforms the Kalman filter in the presence of non-Gaussian noise. The C-Filter uses higher order information from the signals and nonlinear statistics of the data for estimation. The main contribution of this paper is to introduce new and improved algorithms on the basis of [16] for state estimation in the presence of non-Gaussian noise.

This paper is organized as follows. In Section II we briefly review the Kalman filter based on WLS. In Section III we consider the correntropy criterion and its properties for two random variables. In Section IV we review the C-Filter for state estimation, and then propose two new filters: the modified correntropy filter (MC-Filter), and the maximum correntropy criterion Kalman filter (MCC-KF). Section V provides simulation results and compares different estimators in the presence of shot noise and Gaussian mixture noise. In Section VI we conclude the paper and suggest future research.

II. Kalman Filter Based on WLS Method

Consider the linear stochastic discrete-time dynamic system

$$x_k = Fx_{k-1} + w_k \tag{1}$$

$$y_k = Hx_k + v_k \tag{2}$$

where $x_k \in R^n$ and $y_k \in R^m$ are the state vector and measurement vector respectively, and w_k and v_k are the process and measurement noise respectively (zero mean with covariance matrices Q_k and R_k respectively).

Several approaches can be used to derive the Kalman filter equations: for example, the orthogonality principle, the maximum a posterior (MAP) approach, the unbiased minimum variance (UMV) approach, and the weighted least square (WLS) method. In WLS, the quadratic objective function is

$$J = \frac{1}{2}(y_k - H\hat{x}_k)^T R_k^{-1}(y_k - H\hat{x}_k) + \frac{1}{2}(\hat{x}_k - F\hat{x}_{k-1})^T P_{k|k-1}^{-1}(\hat{x}_k - F\hat{x}_{k-1}) \tag{3}$$

where

$$P_{k|k-1} = E[e_k^- e_k^{-T}],\ e_k^- = x_k - \hat{x}_k^-$$

$$R_k = E[v_k v_k^T],\ Q_k = E[w_k w_k^T]$$

The Kalman filter can be derived by solving

$$\frac{\partial J}{\partial \hat{x}_k} = 0 \tag{4}$$

The Kalman filter equations are given as follows:

$$\hat{x}_k^- = F\hat{x}_{k-1} \tag{5}$$

$$P_{k|k-1} = FP_{k-1|k-1}F^T + Q_k \tag{6}$$

$$K_k = (P_{k|k-1}^{-1} + H^T R_k^{-1} H)^{-1} H^T R_k^{-1} \tag{7}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - H\hat{x}_k^-) \tag{8}$$

$$P_{k|k} = (I - K_k H)P_{k|k-1}(I - K_k H)^T + K_k R_k K_k^T \tag{9}$$

where K_k and P_k are the Kalman gain and estimation error covariance matrix respectively. The Kalman filter is optimal in the case of Gaussian noise. However, it is sub-optimal for other noise distributions because it only uses second-order information from the measurements. So there is a need for a method that uses higher-order information, like the maximum correntropy criterion (MCC).

III. Correntropy Measure

The correntropy criterion has received much attention in machine learning, pattern recognition, and non-Gaussian noise signal processing, especially in the case of large outliers [17], [18]. In information theoretic learning and kernel methods, correntropy is defined as a statistical metric of the similarity between two random variables [19]-[22]. Cross-correntropy is a generalization of correlation between a pair of scalar random variables that measures not only the second-order information but also higher-order moments of the joint probability density function [16], [23]. The cross-correntropy of two scalar random variables X and Y is defined as

$$C_\sigma(X,Y) = E[k_\sigma(X,Y)] = \iint k_\sigma(x,y) f_{XY}(x,y)\,dx\,dy \tag{10}$$

where $E[.]$ denotes the expectation operator, $k_\sigma(.,.)$ denotes a positive definite kernel function satisfying Mercer's theory, and $f_{XY}(.,.)$ denotes the joint density function of X and Y. Since the joint distribution function is not accessible and only a finite number of data points N are available in practice, we use the sample estimator

$$\hat{C}_\sigma(X,Y) = \frac{1}{N}\sum_{i=1}^{N} k_\sigma(x_i, y_i) \tag{11}$$

We use the Gaussian kernel in our formulation so (11) is written as follows:

$$\hat{C}_\sigma(X,Y) = \frac{1}{N}\sum_{i=1}^{N} G_\sigma(x_i - y_i) \tag{12}$$

where $G_\sigma(x_i - y_i) = \exp\left(-\frac{\|x_i - y_i\|^2}{2\sigma^2}\right)$ with kernel size (bandwidth) σ. From (12), we see that the Gaussian correntropy function is positive and bounded and reaches its maximum if and only if $X = Y$, in which case the maximum correntropy criterion (MCC) is realized. As mentioned earlier, correntropy uses higher-order moments of the signals, a property which can be shown using a Taylor series expansion of the Gaussian kernel function:

$$C_\sigma(X,Y) = \frac{1}{\sqrt{2\pi}\sigma}\sum_{n=0}^{\infty}\frac{(-1)^n}{2^n\sigma^{2n}n!}E[(X-Y)^{2n}] \tag{13}$$

It can be seen that correntropy involves the sum of all the even moments of the random variable $X - Y$ when the Gaussian kernel is used.

IV. State Estimation Using the Correntropy Criterion

The aforementioned properties of correntropy motivated the definition of a new estimation cost function consisting of Gaussian kernel functions, and resulted in the correntropy filter [16]. In the C-Filter we optimize the following cost function:

$$J_m = G_\sigma(\|y_k - H\hat{x}_k\|) + G_\sigma(\|\hat{x}_k - F\hat{x}_{k-1}\|) \tag{14}$$

The C-Filter formulation is derived by minimizing (14) with respect to $\hat{x}_k$:

$$\frac{\partial J_m}{\hat{x}_k} = 0 \tag{15}$$

$$\frac{1}{\sigma^2} G_\sigma(\|y_k - H\hat{x}_k\|) H^T (y_k - H\hat{x}_k)$$

$$-\frac{1}{\sigma^2} G_\sigma(\|\hat{x}_k - F\hat{x}_{k-1}\|)(\hat{x}_k - F\hat{x}_{k-1}) = 0 \quad (16)$$

$$\hat{x}_k = F\hat{x}_{k-1} + \frac{G_\sigma(\|y_k - H\hat{x}_k\|)}{G_\sigma(\|\hat{x}_k - F\hat{x}_{k-1}\|)} H^T (y_k - H\hat{x}_k) \quad (17)$$

The C-Filter is derived [16] by approximating $\hat{x}_k \approx \hat{x}_k^-$ on the right-hand side of (17). That is, $G_\sigma(\|\hat{x}_k - F\hat{x}_{k-1}\|) \approx G_\sigma(\|\hat{x}_k^- - F\hat{x}_{k-1}\|) = 1$, and so (17) is written as

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - H\hat{x}_k^-) \quad (18)$$

$$K_k = G_\sigma(\|y_k - H\hat{x}_k^-\|) H^T \quad (19)$$

Equations (18) and (19) comprise the C-Filter. We can also write the covariance propagation of the C-Filter by using [25, Chap. 5], although it does not affect the estimation.

A. Modified Correntropy Filter

In this section we introduce the modified correntropy filter (MC-Filter) by modifying (17). Recall that the C-Filter was derived by approximating $\hat{x}_k \approx \hat{x}_k^-$ on the right-hand side of (17). To derive the MC-Filter we approximate $\hat{x}_k \approx \hat{x}_k^-$ only in the Gaussian kernel functions of (17) but not in the innovation term, which gives

$$\hat{x}_k + G_\sigma(\|y_k - H\hat{x}_k^-\|) H^T H\hat{x}_k = F\hat{x}_{k-1} + G_\sigma(\|y_k - H\hat{x}_k^-\|) H^T y_k \quad (20)$$

By adding and subtracting $G_\sigma(\|y_k - H\hat{x}_k^-\|) H^T H\hat{x}_k^-$ on the right-hand side of (20) we obtain

$$(I + G_\sigma(\|y_k - H\hat{x}_k^-\|) H^T H)\hat{x}_k = F\hat{x}_{k-1} + G_\sigma(\|y_k - H\hat{x}_k^-\|) H^T y_k + G_\sigma(\|y_k - H\hat{x}_k^-\|) H^T H\hat{x}_k^- - G_\sigma(\|y_k - H\hat{x}_k^-\|) H^T H\hat{x}_k^- \quad (21)$$

$$(I + G_\sigma(\|y_k - H\hat{x}_k^-\|) H^T H)\hat{x}_k = (I + G_\sigma(\|y_k - H\hat{x}_k^-\|) H^T H)\hat{x}_k^- + G_\sigma(\|y_k - H\hat{x}_k^-\|) H^T (y_k - H\hat{x}_k^-) \quad (22)$$

The state estimate update equation is then obtained as

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - H\hat{x}_k^-) \quad (23)$$

$$K_k = (I + G_\sigma(\|y_k - H\hat{x}_k^-\|) H^T H)^{-1} \times G_\sigma(\|y_k - H\hat{x}_k^-\|) H^T \quad (24)$$

As in the C-Filter, we can write the covariance propagation using [25, Chap. 5]. However, as in the C-Filter, the covariance is not used in the state estimate calculation. The MC-Filter is summarized in Algorithm 1.

Initialization
$\hat{x}_0 = E[x_0], P_0 = E[e_0 e_0^T]$
Prior estimation
$\hat{x}_k^- = F\hat{x}_{k-1}$ $P_{k\|k-1} = FP_{k-1\|k-1}F^T + Q_k$
Posterior estimation
$L_k = G_\sigma(\|y_k - H\hat{x}_k^-\|)$
$K_k = (I + L_k H^T H)^{-1} L_k H^T$
$\hat{x}_k = \hat{x}_k^- + K_k(y_k - H\hat{x}_k^-)$
$P_{k\|k} = (I - K_k H) P_{k\|k-1} (I - K_k H)^T + K_k R_k K_k^T$

Algorithm 1. Modified correntropy filter (MC-Filter)

B. Kalman filter based on maximum correntropy and WLS

We now introduce a new cost function to make the C-Filter more robust in the presence of non-Gaussian noise. We know that the weighting matrices in WLS, R_k^{-1} and $P_{k|k-1}^{-1}$, result in minimum-variance estimation, and the use of correntropy in the C-Filter results in the use of higher-order moments. We can combine these ideas from MCC and WLS to define a new objective function:

$$J_m = G_\sigma\left(\|y_k - H\hat{x}_k\|_{R_k^{-1}}\right) + G_\sigma\left(\|\hat{x}_k - F\hat{x}_{k-1}\|_{P_{k|k-1}^{-1}}\right) \quad (25)$$

To minimize this objective function, we compute its derivative with respect to $\hat{x}_k$ to obtain

$$-\frac{1}{\sigma^2} G_\sigma\left(\|y_k - H\hat{x}_k\|_{R_k^{-1}}\right) H^T R_k^{-1} (y_k - H\hat{x}_k) + \frac{1}{\sigma^2} G_\sigma\left(\|\hat{x}_k - F\hat{x}_{k-1}\|_{P_{k|k-1}^{-1}}\right) P_{k|k-1}^{-1} (\hat{x}_k - F\hat{x}_{k-1}) = 0 \quad (26)$$

$$P_{k|k-1}^{-1}\hat{x}_k - P_{k|k-1}^{-1} F\hat{x}_{k-1} = L_k H^T R_k^{-1} (y_k - H\hat{x}_k) \quad (27)$$

$$L_k = \frac{G_\sigma\left(\|y_k - H\hat{x}_k\|_{R_k^{-1}}\right)}{G_\sigma\left(\|\hat{x}_k - F\hat{x}_{k-1}\|_{P_{k|k-1}^{-1}}\right)} \quad (28)$$

To obtain the state estimate we add and subtract the term $(L_k H^T R_k^{-1} H\hat{x}_k)$ on the right-hand side of (27) to obtain

$$(P_{k|k-1}^{-1} + L_k H^T R_k^{-1} H)\hat{x}_k = P_{k|k-1}^{-1}\hat{x}_k^- + L_k H^T R_k^{-1} y_k - L_k H^T R_k^{-1} H\hat{x}_k^- + L_k H^T R_k^{-1} H\hat{x}_k^- \quad (29)$$

$$(P_{k|k-1}^{-1} + L_k H^T R_k^{-1} H)\hat{x}_k = \left(P_{k|k-1}^{-1} + L_k H^T R_k^{-1} H\right)\hat{x}_k^- + L_k H^T R_k^{-1} (y_k - H\hat{x}_k^-) \quad (30)$$

The new estimator is derived from (30) as follows:

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - H\hat{x}_k^-) \quad (31)$$

$$K_k = (P_{k|k-1}^{-1} + L_k H^T R_k^{-1} H)^{-1} L_k H^T R_k^{-1} \quad (32)$$

$$P_{k|k} = (I - K_k H) P_{k|k-1} (I - K_k H)^T + K_k R_k K_k^T \quad (33)$$

As in the C-Filter and the MC-Filter, the covariance update of (33) is based on the standard covariance update equation of the predictor-corrector estimator of (31) for any gain matrix K_k [25, Chap. 5]. This new estimator, which is called the maximum correntropy criterion Kalman filter (MCC-KF) uses correntropy to deal with non-Gaussian disturbances and also uses the covariance in the gain matrix calculation, as summarized in Algorithm 2.

Initialization :
$\hat{x}_0 = E[x_0], P_0 = E[e_0 e_0^T]$
Prior estimation :
$\hat{x}_k^- = F\hat{x}_{k-1}$ $P_{k\|k-1} = FP_{k-1\|k-1}F^T + Q_k$
Posterior estimation :
$L_k = \dfrac{G_\sigma\left(\|y_k - H\hat{x}_k^-\|_{R_k^{-1}}\right)}{G_\sigma\left(\|\hat{x}_k^- - F\hat{x}_{k-1}\|_{P_{k\|k-1}^{-1}}\right)}$
$K_k = (P_{k\|k-1}^{-1} + L_k H^T R_k^{-1} H)^{-1} L_k H^T R_k^{-1}$
$\hat{x}_k = \hat{x}_k^- + K_k(y_k - H\hat{x}_k^-)$
$P_{k\|k} = (I - K_k H) P_{k\|k-1} (I - K_k H)^T + K_k R_k K_k^T$

Algorithm 2. Maximum correntropy criterion Kalman filter (MCC-KF)

The innovation term $r_k = (y_k - H\hat{x}_k^-)$ in (31) diverges in the presence of shot noise or large outliers, but K_k from (32) prevents the divergence of the estimator. The performance of the MCC-KF in the presence of shot noise is shown as follows:

$$\lim_{r_k \to \infty} y_k = \infty \quad (34)$$

$$\lim_{y_k \to \infty} L_k = 0 \quad (35)$$

$$\lim_{L_k \to 0} K_k = 0 \quad (36)$$

$$\lim_{K_k \to 0} \hat{x}_k = \hat{x}_k^- \quad (37)$$

We see that, contrary to the Kalman filter, measurements with very large outliers will not affect the state estimate.

V. SIMULATION

We consider a benchmark navigation problem [24] to illustrate the new MCC-based estimators. The first two state components are the north and east positions of a land vehicle, and the last two components are the north and east velocities. The velocity of the vehicle is in the direction of θ, an angle measured clockwise from due east. A position-measuring device provides a noisy measurement of the vehicle's north and east positions. The vehicle dynamics and measurement is given as follows:

$$x_k = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} x_{k-1} + \begin{bmatrix} 0 \\ 0 \\ \Delta t \sin\theta \\ \Delta t \cos\theta \end{bmatrix} u_k + w_k$$

$$y_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} x_k + v_k$$

where Δt is the discretisation step size and u_k is a known acceleration input. The sample period Δt is 3 sec and the heading angle θ is set to a constant 60 deg. The system and estimator are initialized as

$$x_0{}^T = \hat{x}_0{}^T = [1 \quad 1 \quad 0 \quad 0], P_0 = \mathrm{Diag}([4, 4, 3, 3])$$

We consider two different cases for the system and measurement disturbances. For each case, we use 100 Monte Carlo simulations to quantify estimation performance.

Case 1: All four elements of w_k, along with the first element of v_k, are comprised of Gaussian noise plus shot noise. Fig. 1 represents the shot noise that is applied to the first measurement. The shot noise that is used for w_k is similar but is not shown here.

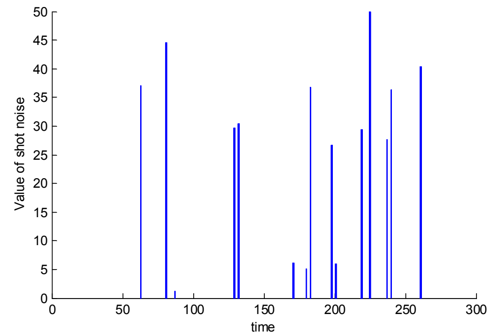

Fig. 1. Case 1 – shot noise in first measured output

The noise terms are expressed as follows:

$$w_k = N(\mu_x, Q) + \text{Shot noise}$$
$$v_k = N(\mu_y, R) + \text{Shot noise}$$

where

$$\mu_x^T = [0 \quad 0 \quad 0 \quad 0], \ \mu_y^T = [0 \quad 0]$$
$$Q = \mathrm{Diag}([0.1, 0.1, 0.1, 0.1]), \ R = \mathrm{Diag}([0.1, 0.1])$$

The root mean square error (RMSE taken over 100 Monte Carlo simulations) for the third state is plotted in Fig. 2, which shows the clear improvement of the MCC-Filter compared to the C-Filter for this scenario.

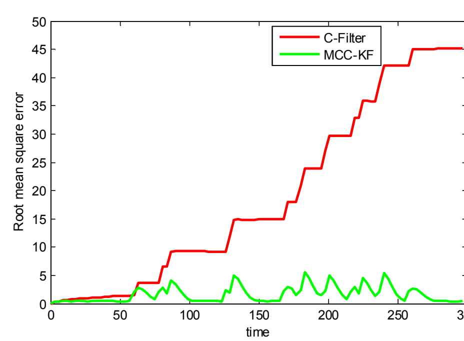

Fig. 2. Case 1 – Comparison of RMSE in the presence of shot noise, where RMSE is taken over 100 Monte Carlo simulations. The third state is shown here for illustration purposes.

The accuracy of several estimators for this problem are compared in Table I, which shows normalized cost functions which are defined as follows:

$$C_j = \sum_{i=1}^{n} \frac{RMSE_{ij}}{\max_{t\in[0,T]} RMSE_{ij}} \quad i = 1,2,\ldots,n \text{ and } j = 1,2,\ldots,s \qquad (38)$$

where $RMSE_{ij}$ is the root mean square error of the i^{th} state using the j^{th} estimator method, n=4 is the number of states, s=6 is the number of estimator methods, and T is the simulation length.

TABLE I. Case 1 – Comparison of RMSE cost values in the presence of shot noise

	RMSE				Cost
	State 1	State 2	State 3	State 4	
UKF	2.16	2.16	0.93	0.94	1.11
GSF	4.17	4.13	1.15	1.16	1.64
EnKF	3.69	3.41	2.47	2.49	1.31
C-Filter	44.81	43.58	20.16	20.55	3.32
MC-Filter	41.14	41.84	20.16	20.23	3.25
MCC-KF	**2.16**	**2.01**	**0.84**	**0.81**	**1.03**

Case 2 : Both w_k and v_k are Gaussian mixture noise as shown in Fig. 3.

$$w_k = \alpha N(\mu_{x1}, Q_1) + (1-\alpha)N(\mu_{x2}, Q_2)$$

$$v_k = \alpha N(\mu_{y1}, R_1) + (1-\alpha)N(\mu_{y2}, R_2)$$

where

$$\mu_{x1}^T = [-3 \quad -3 \quad -3 \quad -3],\ \mu_{x2}^T = [2 \quad 2 \quad 2 \quad 2]$$

$$\mu_{y1}^T = [2 \quad 2],\ \mu_{y2}^T = [-2 \quad -2]$$

$$Q_1 = Q_2 = Q,\ R_1 = R_2 = R, \text{ and } \alpha = 0.5$$

The accuracy of the MCC-KF is compared with the C-Filter and several other filters in Fig. 4 and Table II. We see that in Case 2, as in Case 1, the MCC-KF has significantly better performance than the other filters, including the C-Filter.

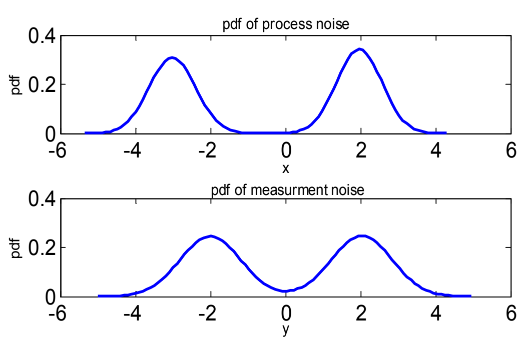

Fig. 3. Case 2 – probability density function of Gaussian mixture noise

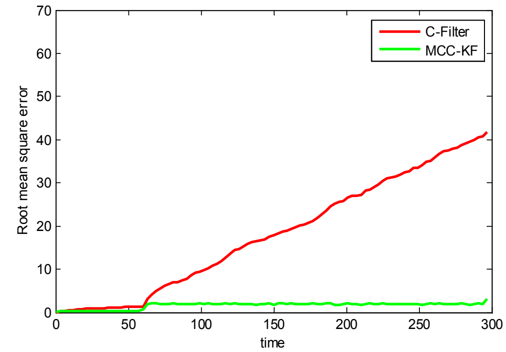

Fig. 4. Case 2 – Comparison of RMSE in the presence of Gaussian mixture noise, where RMSE is taken over 100 Monte Carlo simulations. The third state is shown here for illustration purposes.

In both Case 1 (shot noise) and Case 2 (Gaussian mixture noise), the MCC-KF also has better performance than the UKF, the GSF, and EnKF.

TABLE II. Case 2 – Comparison of RMSE cost values in the presence of Gaussian mixture noise

	RMSE				Cost
	State 1	State 2	State 3	State 4	
UKF	2.98	2.96	1.80	1.82	1.15
GSF	4.17	4.10	2.03	2.03	1.25
EnKF	4.32	4.39	2.55	2.57	1.24
C-Filter	41.05	41.22	20.25	20.18	3.05
MC-Filter	40.11	39.87	20.25	19.96	2.84
MCC-KF	**2.98**	**2.95**	**1.63**	**1.61**	**1.09**

VI. CONCLUSION AND FUTURE WORK

We modified the C-Filter to derive two new correntropy-based state estimators: the modified correntropy filter (MC-Filter) and the maximum correntropy criterion Kalman filter (MCC-KF). The proposed filters have the same structure as the Kalman filter, and also use high-order statistics to improve state estimation. The simulation results show that the MCC-KF has lower estimation errors than the UKF, the GSF, and the EnKF. The MCC-KF does not require the use of multiple filters or sigma points, and so it has lower computational effort than the UKF, GSF, and EnKF. The new MCC-KF achieves performance that is 68% and 65% better than the C-Filter in the presence of shot noise and Gaussian mixture noise respectively.

Kernel size plays a significant role in the behavior of correntropy filters. Future work will focus on finding optimal rules for kernel size selection using, for instance, particle swarm optimization. Future work will also attempt to find analytical conditions that guarantee the stability of the MCC-KF. We can also combine the newly-derived MCC-KF with other state estimation paradigms such as constrained state estimation, moving horizon estimation, particle filtering, and so on.

The Matlab code that was used to generate the simulation results in this paper can be downloaded from http://embeddedlab.csuohio.edu/Correntropy to replicate the results in this paper.

References

[1] A. Bryson and Y. Ho, Applied Optimal Control. Wiley New York, 1975.

[2] S. A. Fakoorian, D. Simon, H. Richter and V. Azimi, "Ground reaction force estimation in prosthetic legs with an extended Kalman filter," IEEE Int. Systems Conference, Orlando, Florida, US, April 2016.

[3] L. Ljung, "Asymptotic behavior of the extended Kalman filter as parameters estimator for linear systems," IEEE Trans. Automatic Control, vol. 24, no. 1, pp. 36–50, 1979.

[4] K.N. Plataniontis, D. Androutsos and A. Venetsanopoulos, "Nonlinear filtering of non-Gaussian noise," Journal of Intelligent and Robotic Systems, vol. 19, pp. 207–231, 1997.

[5] H. Sorenson and D. Alspach, "Recursive Bayesian estimation using Gaussian sums," Automatica, vol. 7, no. 4, pp. 465–479, 1971.

[6] A. Harvey and A. Luati, "Filtering with heavy tails," Journal of the American Statistical Association, vol. 109, no.507, pp. 1112-1122, 2014.

[7] G. Agamennoni, I. J. Neito and E. M. Nebot, "An outlier-robust Kalman filter," IEEE Int. Conf. Robotics and Automation, Shanghai, China, pp. 1551-1558, 2011.

[8] D. Magill, "Optimal adaptive estimation of sampled stochastic processes," IEEE Trans. Automatic Control, vol. 10, no. 4, pp. 434-439, 1965.

[9] I. Bilik, J. Tabrikian, "MMSE-based filtering in presence of non-Gaussian system and measurement noise," IEEE Trans. Aerospace and Electronic systems, vol. 46, no. 3, pp. 1153-1170, 2010.

[10] D. L. Alspach and H. and Sorenson, "Non-linear Bayesian estimation using Gaussian sum approximation," IEEE Trans. Automatic Control, vol. 17, pp. 439-448, 1972.

[11] H. Jayesh and P. M. Djuric, "Gaussian sum particle filter," IEEE Trans. Signal Processing, vol. 51, no. 10, pp. 2602-2612, 2003.

[12] A. Doucet, N. de Freitas, and N. Gordon, An Introduction to Sequential Monte Carlo Methods, New York: Springer, 2001.

[13] P. M. Djuric et al, "Particle filtering," IEEE Signal Processing Magazine, pp. 9-38, 2003

[14] J. Curn, D. Marinescu1, G. Lacey and V. Cahill, "Estimation with non-white Gaussian observation noise using a generalised Ensemble Kalman filter," IEEE International Symposium on Robotic and Sensors Environments, Magdeburg, Germany, pp. 85-90, 2012

[15] S. Julier, S. Uhlmann, "Unscented filtering and nonlinear estimation," Proc. IEEE, vol. 92, no. 3, pp. 401–422, 2004

[16] G. T. Cinar and J. C. Principe, "Hidden state estimation using the Correntropy Filter with fixed point update and adaptive kernel size," IEEE World Congress on Computational Intelligence, Brisbane, Australia, pp. 1-6, 2012.

[17] I. Santamaria and J. C. Principe, "Generalized correlation Function: Definition, properties, and application to blind equalization," IEEE Trans. Signal Processing, vol. 54, no. 6, 2006.

[18] A. Garde, L. Sornmo, R. Jane, B. F. Giraldo, "Correntropy-based analysis of respiratory patterns in patients with chronic heart failure," 31st Annual International Conference of the IEEE EMBS, Minneapolis, Minnesota, pp. 4687-4690, 2009.

[19] L. Weifeng, P. P. Pokharel and J. C. Principe, "Correntropy: Properties and applications in non-Gaussian signal processing," IEEE Trans. Signal Processing, vol. 58, no. 11, pp. 5286-5298, 2007.

[20] R. He, W. Zheng and B. Hu, "Maximum correntropy criterion for robust face recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 33, no. 8, pp. 1561-1586, 2011.

[21] L. Weifeng, P. P. Pokharel and J. C. Principe, "Correntropy: A Localized Similarity Measure," International Joint Conference on Neural Networks, Vancouver, Canada, pp. 4919-4924, 2006.

[22] J. Principe, Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives, Springer Verlag, 2010.

[23] M. Rao, S. Seth, J. Xu, Y. Chen and J. C. Principe, "A test of independence based on a generalized correlation function," Signal Processing, vol. 91, no. 1, pp. 15-27, 2011.

[24] D. Simon, "Kalman filtering with state constraints: a survey of linear and nonlinear algorithms," IET Control Theory & Applications, vol. 4, no. 8, pp. 1303-1318, 2010.

[25] D. Simon, Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches, John Wiley & Sons, 2006.