Proceedings of the ASME 2017 Dynamic Systems and Control Conference
DSCC 2017
October 11-13, 2017, Tysons Corner, Virginia, USA

MAXIMUM CORRENTROPY CRITERION CONSTRAINED KALMAN FILTER

Seyed Fakoorian†, Mahmoud Moosavi, Reza Izanloo, Vahid Azimi, and Dan Simon
Department of Electrical
Engineering and Computer Science
Cleveland State University
Cleveland, Ohio 44115
† Corresponding author: s.fakoorian@csuohio.edu

ABSTRACT

Non-Gaussian noise may degrade the performance of the Kalman filter because the Kalman filter uses only second-order statistical information, so it is not optimal in non-Gaussian noise environments. Also, many systems include equality or inequality state constraints that are not directly included in the system model, and thus are not incorporated in the Kalman filter. To address these combined issues, we propose a robust Kalman-type filter in the presence of non-Gaussian noise that uses information from state constraints. The proposed filter, called the maximum correntropy criterion constrained Kalman filter (MCC-CKF), uses a correntropy metric to quantify not only second-order information but also higher-order moments of the non-Gaussian process and measurement noise, and also enforces constraints on the state estimates. We analytically prove that our newly derived MCC-CKF is an unbiased estimator and has a smaller error covariance than the standard Kalman filter under certain conditions. Simulation results show the superiority of the MCC-CKF compared with other estimators when the system measurement is disturbed by non-Gaussian noise and when the states are constrained.

1 INTRODUCTION

The Kalman filter is the most powerful tool for the state estimation of linear dynamic systems because of its simplicity and optimality. Even for nonlinear systems, extended and sigma-point Kalman filters are common state-estimation methods (for example, the extended Kalman filter (EKF) and the unscented Kalman filter (UKF) [1--3]). However, the performance of Kalman filtering methods may deteriorate if the state estimates of the system are required to satisfy constraints or if the system is disturbed by non-Gaussian noise [4]. In non-Gaussian noise settings, higher-order information of the process and measurement noise is not used in the Kalman filter [5].

Several Kalman-type filters have been introduced for the state estimation of the discrete-time constrained systems. In these filters, equality or inequality state constraints are incorporated with Kalman filter algorithms to provide better filtering performance. In [4] a number of constrained Kalman filters for both linear and nonlinear systems are surveyed, such as the perfect measurement approach [6], estimate projection [7], the projected UKF [8], and others. Although many constrained filters have been proposed, estimate projection is the most commonly used approach for constraint enforcement [9]. In [10] Kalman filtering with quadratic equality state constraints is developed using Lagrangian multipliers. The most recent developments in constrained Kalman filtering are evaluated in [11], where a constrained Kalman-like filter for continuous-time constrained systems is considered with respect to a group of orthonormal matrices.

The performance of the aforementioned filters may break down if the system is perturbed by non-Gaussian noise, since the Kalman filter uses minimum mean square error (MMSE) as the optimality criterion. This criterion is sensitive to large outliers and heavy-tailed non-Gaussian noise. Even though the UKF is robust to non-Gaussian noise because it is based on sequential Monte Carlo sampling [12], computational effort is the main disadvantage of the UKF.

A Kalman-type filter that is robust to non-Gaussian noise and large outliers, the maximum correntropy criterion Kalman filter (MCC-KF), was introduced in [13,14] and further developed in [15]. In MCC-KF, we use correntropy instead of MMSE as the optimality criterion in development of the filter so that the filter can account for higher-order information from the process and measurement noise.

In this paper we generalize the results in [14] to accommodate equality state constraints for linear dynamic systems, and we propose a new Kalman-type filter which is robust to non-Gaussian measurement noise and which can satisfy state equality constraints. We call this filter the maximum correntropy criterion constrained Kalman filter (MCC-CKF). We mathematically prove that our newly derived MCC-CKF has a smaller error covariance than the unconstrained standard Kalman filter under certain conditions. In this paper we first review and rederive the estimate projection method for the constrained Kalman filter (CKF) using the minimum-variance unbiased (MVU) approach. Then we improve the CKF for non-Gaussian noise by deriving the MCC-CKF. Although the measurement update of the MCC-CKF is not projection-based, the time update of the MCC-CKF is the same as in the standard KF (and also the same as in the projection-based CKF). Finally, we present two examples showing that the MCC-CKF outperforms other popular estimation algorithms.

This paper is organized as follows. In Section II we rederive the estimate projection-based CKF. In Section III we propose a new cost function and derive the MCC-CKF. Section IV provides simulation results and compares different constrained estimators in the presence of non-Gaussian noise. Section V concludes the paper and suggests future research.

2 ESTIMATE PROJECTION WITH THE KALMAN FILTER ESTIMATE

Consider the linear stochastic discrete-time system

$$\begin{aligned} x_k &= F x_{k-1} + w_k \\ y_k &= H x_k + v_k \end{aligned} \tag{1}$$

where $x_k \in \mathbb{R}^n$ and $y_k \in \mathbb{R}^m$ 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 $Q_k = E\left[w_k w_k^T\right]$ and $R_k = E\left[v_k v_k^T\right]$), and the matrices F and H are the state transition and measurement matrices, which we assume to be constant for the sake of notational simplicity. The Kalman filter is given as follows [16, Chapter 5]:

$$\hat{x}_0 = E\left[x_0\right] \tag{2}$$

$$P_0 = E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T] \tag{3}$$

$$\hat{x}_k^- = F\hat{x}_{k-1} \tag{4}$$

$$P_{k|k-1} = F P_{k-1|k-1} F^T + Q_k \tag{5}$$

$$K_k = \left(P_{k|k-1}^{-1} + H^T R_k^{-1} H\right)^{-1} H^T R_k^{-1} \tag{6}$$

$$\hat{x}_k = \hat{x}_k^- + K_k\left(y_k - H\hat{x}_k^-\right) \tag{7}$$

$$P_{k|k} = \left(I - K_k H\right) P_{k|k-1} \left(I - K_k H\right)^T + K_k R_k K_k^T \tag{8}$$

for $k = 1, 2, \ldots$, where $E[\cdot]$ is the expected value operation, and $K_k \in \mathbb{R}^{n\times m}$ is the Kalman gain. $\hat{x}_k^-$ is the *a priori* estimate of the state x_k, which is based on measurements up to and including time $k-1$, and has covariance $P_{k|k-1}$. $\hat{x}_k$ is the *a posteriori* estimate of the state x_k, which is based on measurements up to and including time k, and has covariance $P_{k|k}$. It has been proven that the Kalman filter is an unbiased filter and minimizes the estimation error covariance at each time step if both measurement and process noise are Gaussian and uncorrelated [16, Chapter 5].

Now assume that our system satisfies the equality constraint

$$D x_k = d \tag{9}$$

where D is a known matrix and d is a known vector. In this case we want to find a state estimate $\hat{x}_k^c$ which satisfies the constraint

$$D\hat{x}_k^c = d \tag{10}$$

One way to incorporate this linear equality in the Kalman filter is to project the unconstrained estimate of the Kalman filter onto the constraint surface. In [4], the following estimate was derived to satisfy this constraint:

$$\hat{x}_k^c = \text{argmin}_x \left(x - \hat{x}_k\right)^T W \left(x - \hat{x}_k\right), \text{ such that } Dx = d \tag{11}$$

where W is a positive-definite user-specified weighting matrix. It was also shown that $W = P_{k|k}^{-1}$ results in the minimum variance constrained estimate.

Here we use the projection approach to rederive the constrained filter based on the MVU approach. This constrained filter will be compared later in this paper with the MCC-CKF, which will be derived in the next section. We consider

$$\hat{x}_k^c = \hat{x}_k - M_k\left(D\hat{x}_k - d\right) \tag{12}$$

where $\hat{x}_k^c$ is the constrained estimate and the optimal gain M_k is to be determined.

Theorem 2.1: Let $\hat{x}_k$ be an unbiased estimator of x_k; that is, $E[x_k - \hat{x}_k] = E[\xi_k] = 0$. Then the constrained state estimate (12) is unbiased; that is, $E[\hat{x}_k^c] = E[x_k]$.

Proof: From (9) and (12) we can write

$$
\begin{aligned}
E[x_k - \hat{x}_k^c] &= E[x_k - \hat{x}_k + M_k(D\hat{x}_k - d)] \\
&= E[\xi_k + M_k(D\hat{x}_k - d - Dx_k + d)] \\
&= E[\xi_k - M_k D\xi_k] \\
&= (I - M_k D)E[\xi_k]
\end{aligned} \tag{13}
$$

Since $E[\xi_k] = 0$, (12) is unbiased estimator of x_k for any value of M_k. QED

Theorem 2.2: Let M_k be given by

$$
M_k = P_{k|k} D^T \left(D P_{k|k} D^T\right)^{-1} \tag{14}
$$

then $\hat{x}_k^c$ is the MVU estimate of x_k.

Proof: The covariance of the constrained estimation error $P_{k|k}^c = E[\xi_k^c \xi_k^{cT}]$, where $\xi_k^c = x_k - \hat{x}_k^c$. Thus from (13) we can write

$$
\begin{aligned}
P_{k|k}^c &= (I - M_k D)E[\xi_k \xi_k^T](I - M_k D)^T \\
&= (I - M_k D)P_{k|k}(I - M_k D)^T
\end{aligned} \tag{15}
$$

The MVU estimate $\hat{x}_k^c$ is obtained by minimizing the trace of (15) with respect to M_k:

$$
\frac{\partial \mathrm{Tr} P_{k|k}^c}{\partial M_k} = 0 \tag{16}
$$

Solving (16) results in the optimal gain M_k of (14), which yields the MVU estimate $\hat{x}_k^c$. QED

Note that the constrained Kalman filter (CKF) (12) with optimal gain (14) is the same as the CKF which is obtained by solving (11) and setting $W = P_{k|k}^{-1}$. It has been proven in [7] that the covariance of the CKF is smaller than that of unconstrained Kalman filter.

3 CONSTRAINED FILTERING IN THE PRESENCE OF NON-GAUSSIAN NOISE

In this section, we propose a new cost function and use the results of the previous section to derive a new constrained Kalman-type filter which is robust when the system is disturbed by non-Gaussian noise or shot noise. The new Kalman-type filtering algorithm is called the MCC-CKF since it uses correntropy instead of MMSE as the optimality criterion, which enables the filter to use higher-order information from the noise signals. The properties of the MCC have been discussed in [17, 18]. To derive the MCC-CKF we define the following cost function:

$$
\begin{aligned}
J_m =& G_\sigma\left(\|\hat{x}_k - \hat{x}_k^-\|_{P_{k|k-1}^{-1}}\right) + \frac{1}{m}\sum_{j=1}^{m} G_\sigma\left(\|y_{jk} - H_j \hat{x}_k\|_{R_{jjk}^{-1}}\right) + \\
& \frac{1}{2}\|D\hat{x}_k - d\|_{W^{-1}}^2
\end{aligned} \tag{17}
$$

where $G_\sigma(\|\cdot\|) = \exp\left(\frac{-\|\cdot\|^2}{2\sigma^2}\right)$, σ is the user-specified bandwidth (kernel size), and W is a positive-definite user-specified weighting matrix. R_{jjk} is the j-th element of diagonal matrix R_k; and H_j is the j-th row of matrix H. This cost function comprises three terms, where the third term contributes to the state constraint, and the first two terms measure correntropy, which addresses non-Gaussian noise or shot noise. The MCC-CKF can be derived by solving

$$
\frac{\partial J_m}{\partial \hat{x}_k} = 0 \tag{18}
$$

$$
\begin{aligned}
& C_k^s P_{k|k-1}^{-1}\left(\hat{x}_k - \hat{x}_k^-\right) - H^T C_k^m R_k^{-1}\left(y_k - H\hat{x}_k\right) - \\
& \sigma^2 D^T W^{-1}\left(D\hat{x}_k - d\right) = 0
\end{aligned} \tag{19}
$$

where

$$
C_k^s = G_\sigma\left(\|\hat{x}_k - \hat{x}_k^-\|_{P_{k|k-1}^{-1}}\right) \tag{20}
$$

$$
\begin{aligned}
C_k^m = \mathrm{Diag}\Big(& G_\sigma\left(\|y_{1k} - H_1\hat{x}_k\|_{R_{k11}^{-1}}\right), \ldots, \\
& G_\sigma\left(\|y_{mk} - H_m\hat{x}_k\|_{R_{kmm}^{-1}}\right)\Big)
\end{aligned} \tag{21}
$$

To derive the MCC-CKF we approximate $\hat{x}_k \simeq \hat{x}_k^-$ in the Gaussian kernel function of (20). With this approximation $C_k^s \simeq 1$, so we have

$$
\begin{aligned}
& \left(P_{k|k-1}^{-1} + H^T C_k^m R_k^{-1} H - \sigma^2 D^T W^{-1} D\right)\hat{x}_k = P_{k|k-1}^{-1}\hat{x}_k^- + \\
& H^T C_k^m R_k^{-1} y_k - \sigma^2 D^T W^{-1} d
\end{aligned} \tag{22}
$$

Then we add and subtract $\left(H^T C_k^m R_k^{-1} H - \sigma^2 D^T W^{-1} D\right)\hat{x}_k^-$ on the right side of (22) to obtain

$$
\begin{aligned}
& \left(P_{k|k-1}^{-1} + H^T C_k^m R_k^{-1} H - \sigma^2 D^T W^{-1} D\right)\hat{x}_k = P_{k|k-1}^{-1}\hat{x}_k^- + \\
& H^T C_k^m R_k^{-1} y_k - \sigma^2 D^T W^{-1} d + \left(H^T C_k^m R_k^{-1} H - \sigma^2 D^T W^{-1} D\right)\hat{x}_k^- -
\end{aligned}
$$

$$\left(H^T C_k^m R_k^{-1} H - \sigma^2 D^T W^{-1} D\right) \hat{x}_k^- \tag{23}$$

Equation (23) can be simplified as

$$\begin{aligned}&\left(P_{k|k-1}^{-1} + H^T C_k^m R_k^{-1} H - \sigma^2 D^T W^{-1} D\right) \hat{x}_k = \left(P_{k|k-1}^{-1} + \right.\\ &\left. H^T C_k^m R_k^{-1} H - \sigma^2 D^T W^{-1} D\right) \hat{x}_k^- + \\ &H^T C_k^m R_k^{-1} \left(y_k - H\hat{x}_k^-\right) + \sigma^2 D^T W^{-1} \left(D\hat{x}_k^- - d\right)\end{aligned} \tag{24}$$

Now we approximate $\hat{x}_k \simeq \hat{x}_k^-$ in the Gaussian kernel function of (21). We can then write the constrained state estimate update equation as

$$\hat{x}_k = \hat{x}_k^- + L_{1k}\left(y_k - H\hat{x}_k^-\right) + L_{2k}\left(D\hat{x}_k^- - d\right) \tag{25}$$
$$L_{1k} = \lambda_k^{-1} H^T C_k^m R_k^{-1} \tag{26}$$
$$L_{2k} = \lambda_k^{-1} \sigma^2 D^T W^{-1} \tag{27}$$
$$\lambda_k = \left(P_{k|k-1}^{-1} + H^T C_k^m R_k^{-1} H - \sigma^2 D^T W^{-1} D\right) \tag{28}$$
$$\begin{aligned}C_k^m = \mathrm{Diag}&\left(G_\sigma\left(\|y_{1k} - H_1\hat{x}_k^-\|_{R_{k11}^{-1}}\right), \ldots,\right.\\ &\left. G_\sigma\left(\|y_{mk} - H_m\hat{x}_k^-\|_{R_{kmm}^{-1}}\right)\right)\end{aligned} \tag{29}$$

We can write the covariance update equation from (25) as follows [16, Chapter 5]:

$$P_{k|k} = \left(I - L_{1k}H + L_{2k}D\right) P_{k|k-1} \left(I - L_{1k}H + L_{2k}D\right)^T + L_{1k}R_k L_{1k}^T \tag{30}$$

The MCC-CKF is summarized in Table 1.

This new estimator can be applied to constrained systems that are perturbed by non-Gaussian noise. It can be seen that when the j-th measurement is disturbed by large outliers or shot noise, the j-th element of the diagonal matrix C_k^m, from (21), goes to zero and prevents the divergence of the estimator; the innovation term of that measurement, $r_{jk} = y_{jk} - H_j\hat{x}_k$, goes to infinity, but $G_\sigma\left(\|r_{jk}\|_{R_{kjj}^{-1}}\right)$ goes to zero more quickly.

Theorem 3.1: Let $\hat{x}_k^-$ be an unbiased estimator of x_k; that is, $E[x_k - \hat{x}_k^-] = E[\xi_k^-] = 0$. Then the *a posteriori* MCC-CKF (25) is also an unbiased estimator; that is, $E[x_k - \hat{x}_k] = E[\xi_k^m] = 0$.

Proof: From (1) and (25) we can write

$$\begin{aligned}E\left[\xi_k^m\right] &= E\left[x_k - \hat{x}_k^- - L_{1k}\left(y_k - H\hat{x}_k^-\right) - L_{2k}\left(D\hat{x}_k^- - d\right)\right]\\ &= E\left[\xi_k^- - L_{1k}\left(H\xi_k^- + v_k\right) + L_{2k}D\xi_k^-\right]\\ &= \left(I - L_{1k}H + L_{2k}D\right) E\left[\xi_k^-\right] - L_{1k}E\left[v_k\right]\\ &= 0\end{aligned} \tag{31}$$

QED

Theorem 3.2: The MCC constrained state estimate (25) has a smaller error covariance than that of the unconstrained state estimate (7) if $W = I$ and σ satisfies

$$\sigma^2 D^T D = H^T C_k^m R_k^{-1} H - H^T R_k^{-1} H \tag{32}$$

where $\sigma \neq 0$ and D is full rank.

Proof: The covariance matrix of the MCC-CKF (30) can be written as

$$P_{k|k} = \left(I - \beta_k\right) P_{k|k-1} \left(I - \beta_k\right)^T + L_{1k}R_k L_{1k}^T \tag{33}$$

where $\beta_k = L_{1k}H - L_{2k}D$. Substituting (26) and (27) in the equation for β_k gives

$$\begin{aligned}\beta_k = &\left(P_{k|k-1}^{-1} + H^T C_k^m R_k^{-1} H - \sigma^2 D^T D\right)^{-1} \times\\ &\left(H^T C_k^m R_k^{-1} H - \sigma^2 D^T D\right)\end{aligned} \tag{34}$$

Substituting (32) into (34) yields

$$\beta_k = \left(P_{k|k-1}^{-1} + H^T R_k^{-1} H\right)^{-1} H^T R_k^{-1} H \tag{35}$$

TABLE 1: Maximum correntropy criterion constrained Kalman filter (MCC-CKF)

Initialization:
$\hat{x}_0 = E\left[x_0\right]$, $P_0 = E\left[\left(x_0 - \hat{x}_0\right)\left(x_0 - \hat{x}_0\right)^T\right]$
Prior estimation:
$\hat{x}_k^- = F\hat{x}_{k-1}$
$P_{k\|k-1} = FP_{k\|k-1}F^T + Q_k$
Posteriori estimation:
$C_k^m = \mathrm{Diag}\left(G_\sigma\left(\\|y_{1k} - H_1\hat{x}_k^-\\|_{R_{k11}^{-1}}\right), \ldots, G_\sigma\left(\\|y_{mk} - H_m\hat{x}_k^-\\|_{R_{kmm}^{-1}}\right)\right)$
$\lambda_k = P_{k\|k-1}^{-1} + H^T C_k^m R_k^{-1} H - \sigma^2 D^T W^{-1} D$
$L_{1k} = \lambda_k^{-1} H^T C_k^m R_k^{-1}$
$L_{2k} = \lambda_k^{-1} \sigma^2 D^T W^{-1}$
$\hat{x}_k = \hat{x}_k^- + L_{1k}\left(y_k - H\hat{x}_k^-\right) + L_{2k}\left(D\hat{x}_k^- - d\right)$
$P_{k\|k} = \left(I - L_{1k}H + L_{2k}D\right) P_{k\|k-1} \left(I - L_{1k}H + L_{2k}D\right)^T + L_{1k}R_k L_{1k}^T$

Similarly, $L_{1k}R_kL_{1k}^T$ can be written as

$$L_{1k}R_kL_{1k}^T = \left(P_{k|k-1}^{-1} + H^TR_k^{-1}H\right)^{-1} H^TC_k^mR_k^{-1}C_k^mH\times \left(P_{k|k-1}^{-1} + H^TR_k^{-1}H\right)^{-1} \tag{36}$$

When we compare β_k in (35) with K_kH from (6), we can see that $\beta_k = K_kH$. When we compare $L_{1k}R_kL_{1k}^T$ in (36) with $K_kR_kK_k^T$ from (8), we can see that $L_{1k}R_kL_{1k}^T \leq K_kR_kK_k^T$ because $C_k^m \in [0,1]$. Therefore, the covariance of the MCC-CKF (30) is smaller than that of the KF (8). QED

4 SIMULATION RESULTS

To illustrate the effectiveness of the proposed filter for systems with state constraints in the presence of shot noise, we provide two examples. The first is a simple linear benchmark navigation problem [14], and the second example is a pendulum with a nonlinear constraint which demonstrates that our proposed filter is effective for linearization-based nonlinear filtering.

The kernel size σ plays a significant role in the behavior of correntropy filters. A larger value of σ reduces the effect of the correntropy for the rejection of non-Gaussian noise since $\sigma \to \infty$ results in $C_k^m \to 1$ [15]. In general, the kernel size can be set manually by trial and error, or by using an optimization algorithm; for instance, particle swarm optimization. An adaptive rule for kernel size selection was introduced for a different filter in [13], where they heuristically set the kernel size at each time instant to the innovation term. In this paper, we chose $\sigma = 1$ by trial and error in both examples.

Example 1 – Linear vehicle tracking

Consider a land-based vehicle navigation problem [4]. 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 are given as

$$\begin{aligned} x_k &= \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} x_{k-1} + \begin{bmatrix} 0 \\ 0 \\ \Delta t \sin\theta \\ \Delta t \cos\theta \end{bmatrix} u_k + w_k \\ y_k &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} x_k + v_k \end{aligned} \tag{37}$$

where Δt is the discretization step size and u_k is a known acceleration input. The sample period $\Delta t = 3$ sec and the heading angle $\theta = 60$ deg. The system and estimator are initialized as

$$x_0^T = \hat{x}_0^T = \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix},\ P_0 = \mathrm{Diag}\{4,4,3,3\}$$

If we know that the vehicle is traveling on a road with a heading of θ, then constraint $Dx = d$ holds, where

$$\begin{aligned} D &= \begin{bmatrix} 1 & -\tan\theta & 0 & 0 \\ 0 & 0 & 1 & -\tan\theta \end{bmatrix} \\ d &= \begin{bmatrix} 0 & 0 \end{bmatrix}^T \end{aligned}$$

Suppose that the process noise $w(t)$ and the measurement noise $v(t)$ are Gaussian, but shot noise is imposed on the first measurement.

$$\begin{aligned} w(t) &\sim N(0,Q) \\ v(t) &\sim N(0,R) + \text{shot noise} \end{aligned}$$

where $Q = \mathrm{Diag}\{0.1,0.1,0.1,0.1\}$ and $R = \mathrm{Diag}\{0.1,0.1\}$. The shot noise can be modeled as an impulsive function as shown in Fig. 1. We use 100 Monte Carlo simulations to quantify estimation performance and we plot the RMSE over 100 Monte Carlo simulations for the third state in Fig. 2. We see that the MCC-CKF performs substantially better than the KF and the CKF.

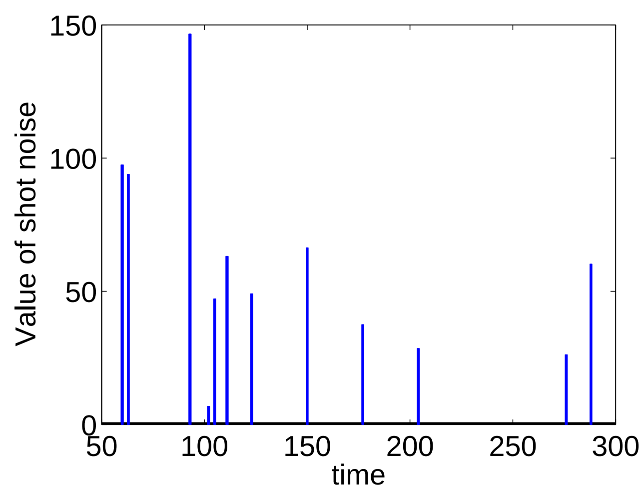

FIGURE 1: Shot noise in the first measured output of Example 1

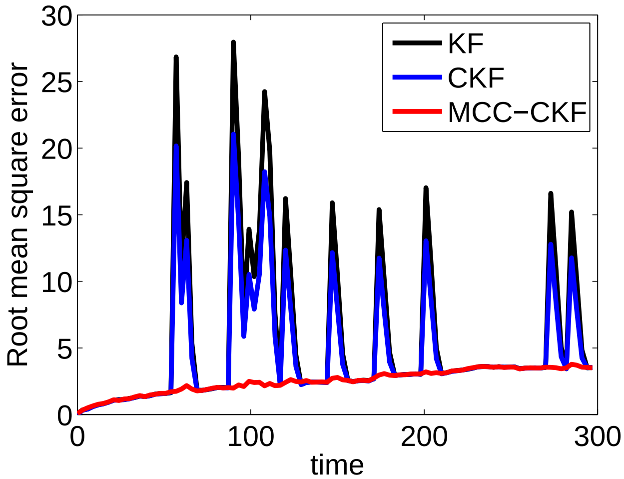

FIGURE 2: Comparison of RMSE in the presence of shot noise for Example 1, where RMSE is taken over 100 Monte Carlo simulations. The third state is shown here for illustration purposes.

Example 2 – A simple pendulum

A discretized model of a pendulum is given as follows [19]:

$$\begin{aligned} \theta_k &= \theta_{k-1} + T\omega_{k-1} \\ \omega_k &= \omega_{k-1} - \left(T\frac{g}{L}\right)\sin\theta_{k-1} \\ y_k &= \omega_k + v_k + \text{shot noise} \end{aligned} \tag{38}$$

where θ is angular position, ω is angular velocity, T is the discretization step size, g is the acceleration due to gravity, and L is the pendulum length. Using noisy measurements of the angular velocity of the pendulum, we want to obtain state estimates that satisfy the conservation of energy, which is a nonlinear constraint on the states θ and ω:

$$-mgL\cos\theta_k + mL^2\omega^2/2 = c \tag{39}$$

where c is some constant and m is the pendulum mass. We use $L = 1$, $m = 1$, $g = 9.81$ and $T = 0.05$ sec. We also use the initial state $x_0^T = \begin{bmatrix}\theta_0 & \omega_0\end{bmatrix} = \begin{bmatrix}\pi/4 & \pi/50\end{bmatrix}$ with zero initial estimation error. We do not consider process noise in the system simulation, but we use covariance matrix $Q = \mathrm{Diag}\{0.007^2, 0.007^2\}$ to enhance the responsiveness of the Kalman filter to measurements. The Gaussian measurement noise $v(t)$ has covariance $R = 0.01$, and we assume that the system measurement is also disturbed by impulsive shot noise.

We can perform a Taylor series expansion to linearize the nonlinear constraint and system dynamics and then apply the CKF and the MCC-CKF for constrained state estimation. This idea is similar to that which is used in the EKF. The accuracy of several estimators for this problem are compared in Table 2, where 100 Monte Carlo simulations are considered for each RMS estimation error. It can be seen that the MCC-CKF works well for this system and achieves the smallest RMSE among the filters in the table.

TABLE 2: Comparison of the RMSE of the state estimates of Example 2 in the presence of shot noise over 100 Monte Carlo simulations

	state 1	state 2
unconstrained EKF	1.22	1.15
perfect measurement	0.85	0.81
projected UKF	0.18	0.12
CKF	0.91	0.85
MCC-CKF	**0.09**	**0.09**

5 CONCLUSION AND FUTURE WORKS

We first rederived the projection-based CKF and then we used the correntropy-based MCC, along with a new cost function, to derive a robust Kalman-like constrained filter called the MCC-CKF. This filter can be used for state estimation of systems with state constraints in the presence of non-Gaussian noise. Although this filter is not a projection-based method, the projection-based CKF and the MCC-CKF both use the same equation for the time update. We mathematically proved that this filter achieves a smaller estimation error covariance than the standard Kalman filter under certain conditions. Simulation results showed that the proposed MCC-CKF achieved smaller estimation errors that other estimators for both a linear and a nonlinear example.

As future work we will generalize the kernel size theorem to derive more broadly based results for the estimation error covariance. We will also optimize the kernel size using an adaptive mechanism to ensure that the error covariance of the filter is as small as possible. We will also generalize the filter to include a more broad class of weighting functions W. Future work will also attempt to derive the MCC-CKF for continuous-time systems. We will also combine the MCC-CKF approach with other state estimators such as the UKF and moving horizon estimation (MHE).

ACKNOWLEDGMENT

This work was supported by National Science Foundation grant 1344954.

REFERENCES

[1] Reif, K., Gunther, S., E.Yaz, and Unbehauen, R., 2000. ‘‘Stochastic stability of the continuous-time extended Kalman filter’’. *IEE Proc. Control Theory Application,* ***147***(1), pp. 45--72.

[2] Sarkka, S., 2007. ‘‘On unscented Kalman filtering for state estimation of continuous-time nonlinear systems’’. *IEEE Transactions on Automatic Control,* ***52***(9), pp. 1631--1641.

[3] Julier, S., Uhlmann, K., and Durrant, H., 1995. ‘‘A new approach for filtering nonlinear systems’’. In American Control Conference, pp. 1628--1632.

[4] Simon, D., 2010. ‘‘Kalman filtering with state constraints: a survey of linear and nonlinear algorithms’’. *IET Control Theory & Applications,* ***4***(8), pp. 1303--1318.

[5] Plataniotis, K. N., Androutsos, D., and Venetsanopoulos, A. N., 1997. ‘‘Nonlinear filtering of non-Gaussian noise’’. *Journal of Intelligent and Robotic Systems,* ***19***(2), pp. 207--231.

[6] Wang, L.-S., Chiang, Y.-T., and Chang, F.-R., 2002. ‘‘Filtering method for nonlinear systems with constraints’’. *IEE Proceedings-Control Theory and Applications,* ***149***(6), pp. 525--531.

[7] Simon, D., and Chia, T. L., 2002. ‘‘Kalman filtering with state equality constraints’’. *IEEE Transactions on Aerospace and Electronic Systems,* ***38***(1), pp. 128--136.

[8] Teixeira, B. O., Torres, L. A., Aguirre, L. A., and Bernstein, D. S., 2008. ‘‘Unscented filtering for interval-constrained nonlinear systems’’. In IEEE Conference on Decision and Control, pp. 5116--5121.

[9] Hewett, R. J., Heath, M. T., Butala, M. D., and Kamalabadi, F., 2010. ‘‘A robust null space method for linear equality constrained state estimation’’. *IEEE Transactions on Signal Processing,* ***58***(8), pp. 3961--3971.

[10] Wang, D., Li, M., Huang, X., and Li, J., 2014. ‘‘Kalman filtering for a quadratic form state equality constraint’’. *Journal of Guidance, Control, and Dynamics,* ***37***(3), pp. 951--958.

[11] Ruiter, A. H., and Forbes, J. R., 2017. ‘‘Continuous-time Kalman filtering on the orthogonal group $o(n)$’’. *International Journal of Robust and Nonlinear Control*.

[12] Doucet, A., De Freitas, N., and Gordon, N., 2001. ‘‘An introduction to sequential Monte Carlo methods’’. In *Sequential Monte Carlo Methods in Practice*. Springer, pp. 3--14.

[13] Cinar, G. T., and Principe, J. C., 2012. ‘‘Hidden state estimation using the correntropy filter with fixed point update and adaptive kernel size’’. In International Joint Conference on Neural Networks, pp. 1--6.

[14] Izanloo, R., Fakoorian, S. A., Sadoghi, H., and Simon, D., 2016. ‘‘Kalman filtering based on the maximum correntropy criterion in the presence of non-Gaussian noise’’. In 50^{th} Annual Conference on Information Science and Systems, pp. 530--535.

[15] Chen, B., Liu, X., Zhao, H., and Príncipe, J. C., 2017. ‘‘Maximum correntropy Kalman filter’’. *Automatica,* ***76***, pp. 70--77.

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

[17] Liu, W., Pokharel, P. P., and Príncipe, J. C., 2007. ‘‘Correntropy: Properties and applications in non-Gaussian signal processing’’. *IEEE Transactions on Signal Processing,* ***55***(11), pp. 5286--5298.

[18] He, R., Zheng, W.-S., and Hu, B.-G., 2011. ‘‘Maximum correntropy criterion for robust face recognition’’. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* ***33***(8), pp. 1561--1576.

[19] Teixeira, B. O., Chandrasekar, J., Tôrres, L. A., Aguirre, L. A., and Bernstein, D. S., 2009. ‘‘State estimation for linear and non-linear equality-constrained systems’’. *International Journal of Control,* ***82***(5), pp. 918--936.