Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

On the Stable Cholesky Factorization-based Method for the Maximum Correntropy Criterion Kalman Filtering ⋆

Maria V. Kulikova *

* *CEMAT (Center for Computational and Stochastic Mathematics), Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 LISBOA, Portugal (email: maria.kulikova@ist.utl.pt).*

Abstract: This paper continues the research devoted to the design of numerically stable square-root implementations for the maximum correntropy criterion Kalman filtering (MCC-KF). In contrast to the previously obtained results, here we reveal the first robust (with respect to round-off errors) method within the Cholesky factorization-based approach. The method is formulated in terms of square-root factors of the *covariance* matrices, i.e. it belongs to the covariance-type filtering methodology. Additionally, a numerically stable orthogonal transformation is utilized at each iterate of the algorithm for accurate propagation of the Cholesky factors involved. The results of numerical experiments illustrate a superior performance of the novel MCC-KF implementation compared to both the conventional algorithm and its previously published Cholesky-based variant.

Keywords: Maximum Correntropy Criterion, Kalman filter, Cholesky decomposition.

1. INTRODUCTION

One of the most recent solution to the problem of "distributionally" robust filtering (i.e. when the actual distribution deviates from the "nominal" one) is obtained under the so-called maximum correntropy criterion (MCC). More precisely, when the classical state-space models are examined, the "nominal" distribution is assumed to be Gaussian, and the goal is to enhance the underlying filter performance in terms of estimation accuracies in a case of the presence of outliers. The obtained MCC-based estimators with the Kalman filtering (KF) like structure (i.e. the first two moments are computed, only) are derived for both the linear stochastic systems in Liu et al. (2007); Chen et al. (2014, 2015); Cinar and Príncipe (2011, 2012); Izanloo et al. (2016); Chen et al. (2017); Liu et al. (2017b); Kulikova (2017); Fakoorian et al. (2019) and the nonlinear state-space models in Liu et al. (2016); Kulikov and Kulikova (2018a); Liu et al. (2017a); Qin et al. (2017); Wang et al. (2017, 2016); Kulikov and Kulikova (2020). One of the obtained estimators is called the maximum correntropy criterion Kalman filter (MCC-KF) as proposed in Izanloo et al. (2016). This estimation method is widely used for solving practical problems; e.g., see Yang and Huang (2017, 2018) and many other studies.

Recently, the numerical robustness issues of the mentioned MCC-KF estimator have been investigated in Kulikova (2019). The key problem in this research area is to derive a stable (in a finite precision arithmetic) square-root MCC-KF implementation methods, which are demanded for solving applications with high reliability requirements as discussed in Grewal and Kain (2010); Grewal (2019) and many other works. In this paper, we focus on the traditional square-root strategy that is based on the Cholesky factorization applied to covariance matrices involved in the filter; see Kailath et al. (2000); Simon (2006); Grewal and Andrews (2015). Unfortunately, the previously discovered Cholesky-based MCC-KF method in Kulikova (2019) is shown to possess a poor performance in ill-conditioned state estimation scenario. Thus, the goal of this research is to answer a question whether it is possible or not to find a more reliable MCC-KF implementation within the discussed class of Cholesky factorization-based algorithms. We answer positively this question and derive the first stable square-root MCC-KF method that outperforms for estimation accuracy both the conventional MCC-KF algorithm and its previously published Cholesky-based variant.

More precisely, the novel estimator is designed in terms of the *lower* triangular Cholesky factors of the error *covariance matrices*, i.e. it is of covariance-type filtering methodology. Additionally, stable orthogonal rotations are utilized as far as possible for propagating the involved Cholesky factors. The performance of the examined MCC-KF implementation methods are studied by using a sixth-order radar tracking system example.

2. PROBLEM STATEMENT

Consider a linear discrete-time state-space model

$$x_k = F_{k-1}x_{k-1} + G_{k-1}w_{k-1}, \quad k \geq 1, \tag{1}$$
$$y_k = H_k x_k + v_k \tag{2}$$

where the system matrices $F_k \in \mathbb{R}^{n\times n}$, $G_k \in \mathbb{R}^{n\times q}$, $H_k \in \mathbb{R}^{m\times n}$ and the noises' covariances $Q_k \in \mathbb{R}^{q\times q}$ ($Q_k > 0$) and

⋆ The author acknowledges the financial support of the Portuguese FCT — *Fundação para a Ciência e a Tecnologia*, through the projects UIDB/04621/2020 and UIDP/04621/2020 of CEMAT/IST-ID, Center for Computational and Stochastic Mathematics, Instituto Superior Técnico, University of Lisbon.

$R_k \in \mathbb{R}^{m\times m}$ ($R_k \geq 0$) are known. The vectors $x_k \in \mathbb{R}^n$ and $y_k \in \mathbb{R}^m$ are, respectively, the hidden dynamic state to be estimated and the available measurement vector. The random variables x_0, w_k and v_k are assumed to satisfy

$$\mathbf{E}\{x_0\} = \bar{x}_0, \quad \mathbf{E}\left\{(x_0-\bar{x}_0)(x_0-\bar{x}_0)^\top\right\} = \Pi_0,$$
$$\mathbf{E}\{w_k\} = \mathbf{E}\{v_k\} = 0, \quad \mathbf{E}\left\{w_k x_0^\top\right\} = \mathbf{E}\left\{v_k x_0^\top\right\} = 0,$$
$$\mathbf{E}\left\{w_k v_k^\top\right\} = 0, \quad \mathbf{E}\left\{w_k w_j^\top\right\} = Q_k\delta_{kj},$$
$$\mathbf{E}\left\{v_k v_j^\top\right\} = R_k\delta_{kj}$$

where the symbol δ_{kj} is the Kronecker delta function, and the initial mean $\bar{x}_0$ and error covariance $\Pi_0 \geq 0$ are known.

The minimum *linear* expected mean square error (MSE) estimator derived for the examined state-space model (1), (2) is known as the Kalman filter (KF); see Theorem 9.2.1 in Kailath et al. (2000). In case of Gaussian uncertainties in the model, the minimum expected MSE estimate belongs to a class of linear functions and, hence, being a linear estimator the KF provides the optimal estimate in the MSE sense; e.g., see Aravkin et al. (2017). However, in non-Gaussian settings, the classical KF exhibits only sub-optimal behavior under the minimum expected MSE estimation criterion.

The concept of *correntropy* (that is a similarity measure of two random variables) has become a very popular technique in the past few years in the realm of designing the "distributionally" robust estimators, i.e. when the actual distribution deviates from the "nominal" one. For the examined state-space models, the "nominal" distribution is assumed to be Gaussian, and the goal is to enhance the classical KF performance in a case of outliers appearance. One of the resulted estimators is called the maximum correntropy criterion Kalman filter (MCC-KF) as proposed in Izanloo et al. (2016). In general, the maximum correntropy cost function is used in the related estimation problem as follows, e.g., see the details in Chapter 5 in Comminiello and Príncipe (2018): an estimator of unknown state $X \in \mathbb{R}$ can be defined as a function of observations $Y \in \mathbb{R}^m$, i.e. $\hat{X} = g(Y)$ where g is solved by maximizing the correntropy between X and $\hat{X}$, i.e.

$$\arg\max_{g\in G} V(X,\hat{X}) = \arg\max_{g\in G} \mathbf{E}\left\{k_\sigma\Big(X - g(Y)\Big)\right\} \tag{3}$$

where G stands for the collection of all measurable functions of Y, $k_\sigma(\cdot)$ is a kernel function and $\sigma > 0$ is the kernel size (bandwidth). As mentioned above, the "nominal" distribution is assumed to be Gaussian and, hence, we explore the Gaussian kernel in the related estimation problem given as follows:

$$k_\sigma(X-\hat{X}) = \exp\left\{-(X-\hat{X})^2/(2\sigma^2)\right\}. \tag{4}$$

It is not difficult to see that the MCC cost (3) with kernel (4) reaches its maximum if and only if $X = \hat{X}$.

In summary, the discussed MCC-KF technique for estimating the unknown dynamic state $\hat{x}_{k|k}$ in the state-space model (1), (2) with the Gaussian kernel $k_\sigma(\cdot)$ from (4) is derived by maximizing the following cost function:

$$J(k) = k_\sigma(\|\hat{x}_{k|k} - F_{k-1}\hat{x}_{k-1|k-1}\|_{P^{-1}_{k|k-1}}) + k_\sigma(\|y_k - H_k\hat{x}_{k|k}\|_{R_k^{-1}}). \tag{5}$$

The stated optimization problem has been solved by Cinar and Príncipe (2012); Izanloo et al. (2016) where the arisen nonlinear equation was resolved with respect to $\hat{x}_{k|k}$ by utilizing a fixed point rule with one iterate, only. More precisely, this approach yields the following filtering recursion:

$$\hat{x}_{k|k} = F_{k-1}\hat{x}_{k-1|k-1} + K_k(y_k - H_k\hat{x}_{k|k-1}) \tag{6}$$

where the gain matrix is computed by

$$K_k = \lambda_k\left(P^{-1}_{k|k-1} + \lambda_k H_k^\top R_k^{-1}H_k\right)^{-1} H_k^\top R_k^{-1} \tag{7}$$

and λ_k is a scalar adjusting weight as suggested in Izanloo et al. (2016)

$$\lambda_k = \frac{k_\sigma(\|y_k - H_k\hat{x}_{k|k-1}\|_{R_k^{-1}})}{k_\sigma(\|\hat{x}_{k|k-1} - F_{k-1}\hat{x}_{k-1|k-1}\|_{P^{-1}_{k|k-1}})}. \tag{8}$$

Finally, the recursion for the state estimate in (6) is combined with the symmetric Joseph stabilized formula for calculating the filter error covariance matrix, which has been derived for the classical KF; e.g., see Kailath et al. (2000); Simon (2006); Grewal and Andrews (2015). The pseudo-code in Algorithm 1 summarizes the discussed MCC-KF suggested in Izanloo et al. (2016).

Algorithm 1. MCC-KF (*conventional MCC-KF*)

Initialization: ($k = 0$) $\hat{x}_{0|0} = \bar{x}_0$ and $P_{0|0} = \Pi_0$.
Time Update: ($k = \overline{1,N}$)
1 $\hat{x}_{k|k-1} = F_{k-1}\hat{x}_{k-1|k-1}$;
2 $P_{k|k-1} = F_{k-1}P_{k-1|k-1}F_{k-1}^\top + G_{k-1}Q_{k-1}G_{k-1}^\top$;
Measurement Update: ($k = \overline{1,N}$)
3 Compute λ_k by formula (8);
4 $K_k = \lambda_k\left(P^{-1}_{k|k-1} + \lambda_k H_k^\top R_k^{-1}H_k\right)^{-1} H_k^\top R_k^{-1}$;
5 $P_{k|k} = (I - K_kH_k)P_{k|k-1}(I - K_kH_k)^\top + K_kR_kK_k^\top$;
6 $\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - H_k\hat{x}_{k|k-1})$.

Algorithm 1 is said to be of the *covariance* type presented in the *conventional* form, because it recursively processes the *full* error *covariance* matrices $P_{k|k-1}$ and $P_{k|k}$ at each iterate of the filter. Taking into account that any covariance matrix is a symmetric matrix, it makes sense to propagate only half of them. The traditional way utilized in the KF community is the use of Cholesky factorization. It implies the decomposition $P = SS^\top$ and, then, the underlying filtering recursion is re-derived in terms of propagating the Cholesky factor S, only; see Morf and Kailath (1975); Park and Kailath (1995); Bierman and Thornton (1977) and many other studies. This computational approach is also known to improve the numerical stability of any conventional KF-like implementation in a finite precision arithmetic, because it ensures the symmetric form and positive (semi-) definiteness of the original matrix P (while the recovering by backward multiplication $SS^\top = P$) despite the influence of roundoff errors; see Kailath et al. (2000); Simon (2006); Grewal and Andrews (2015). Recently, the Cholesky factorization-based method has been derived in Kulikova (2019). However, the numerical stability of the suggested method is still poor as illustrated by the results of numerical experiments presented in the cited paper. The question to be answered in this paper is the following: whether it is possible or not to find a more reliable MCC-KF implementation within the discussed Cholesky-based class of methods.

3. MCC-KF CHOLESKY-BASED FILTERING

We start our research with the MCC-KF square-root method previously designed in Kulikova (2019). The goal is to improve the method in terms of its stability with respect to roundoff errors. It is worth noting here that the first Cholesky factorization-based MCC-KF implementation derived in the cited paper is formulated in terms of the upper triangular Cholesky factors. For readers' convenience, we first re-formulate it in terms of the lower triangular matrices and, next, we propose its robust alternative. More precisely, let's consider the Cholesky factorization of a symmetric positive definite matrix A in the form $A = A^{1/2}A^{\top/2}$ where the factor $A^{1/2}$ is a *lower* triangular matrix with positive diagonal entries. The square-root MCC-KF algorithms imply the following essential features: (i) the factorization is only performed for $\Pi_0 > 0$; (ii) the filtering equation in Algorithm 1 are then re-derived for propagating the lower triangular matrices $P_{k|k-1}^{1/2}$ and $P_{k|k}^{1/2}$; (iii) numerically stable orthogonal rotations are utilized for updating the involved Cholesky factors that additionally provide the *array* form suitable for parallel implementation. Finally, it is important that the adjusting weight λ_k defined in (8) is a nonnegative value and, hence, a square root exists.

To design the mathematically equivalent analogue of Algorithm 1, we factorize its equations as follows:

$$P_{k|k-1} = F_{k-1}P_{k-1|k-1}F_{k-1}^{\top} + G_{k-1}Q_{k-1}G_{k-1}^{T} = AA^{\top}$$
$$= [F_{k-1}P_{k-1|k-1}^{1/2}, G_{k-1}Q_{k-1}^{1/2}][F_{k-1}P_{k-1|k-1}^{1/2}, G_{k-1}Q_{k-1}^{1/2}]^{\top}$$

where an orthogonal rotation, say Q, is applied to get the corresponding lower triangular post-array, R, by transformation $AQ = R$, i.e. we have

$$RR^{\top} = (AQ)(AQ)^{\top} = AA^{\top} = [X,\ 0][X,\ 0]^{\top} = P_{k|k-1}$$

and, hence, we conclude that $X := P_{k|k-1}^{1/2}$ is the resulted square-root factor, which we are looking for. This value is pulled out from the post-array, if required.

Next, the following formulas have been proved for the filter gain K_k calculation in Kulikova (2017):

$$K_k = \lambda_k \hat{P}_{k|k} H_k^{\top} R_k^{-1}$$
$$= \lambda_k \left(P_{k|k-1}^{-1} + \lambda_k H_k^{\top} R_k^{-1} H_k\right)^{-1} H_k^{\top} R_k^{-1}. \qquad (9)$$
$$K_k = \lambda_k P_{k|k-1} H_k^{\top} R_{e,k}^{-1}$$
$$= \lambda_k P_{k|k-1} H_k^{\top} (\lambda_k H_k P_{k|k-1} H_k^{\top} + R_k)^{-1}. \qquad (10)$$

It is not difficult to see that the MCC-KF estimator involves equation (9) for computing K_k in line 4 of Algorithm 1. To derive the square-root form, we consider the term in the brackets, i.e. $\hat{P}_{k|k}^{-1} = P_{k|k-1}^{-1} + \lambda_k H_k^{\top} R_k^{-1} H_k$, and factorize it in the form $AA^{\top}$ where the post array is obtained by transformation $R = AQ$. Thus, we get

$$AA^{\top} = P_{k|k-1}^{-1} + \lambda_k H_k^{\top} R_k^{-1} H_k$$
$$= [P_{k|k-1}^{-\top/2},\ \lambda_k^{1/2} H_k^{\top} R_k^{-\top/2}][P_{k|k-1}^{-\top/2},\ \lambda_k^{1/2} H_k^{\top} R_k^{-\top/2}]^{\top}$$
$$= RR^{\top} = \hat{P}_{k|k}^{-1} = [\hat{P}_{k|k}^{-\top/2},\ 0][\hat{P}_{k|k}^{-\top/2},\ 0]^{\top}.$$

When the resulted block $[\hat{P}_{k|k}^{-\top/2}]$ is read-off from the post-array, the gain matrix is calculated as follows:

$$K_k = \lambda_k P_{k|k} H_k^{\top} R_k^{-1} = \lambda_k \Big([\hat{P}_{k|k}^{-\top/2}][\hat{P}_{k|k}^{-\top/2}]^{\top}\Big)^{-1} H_k^{\top} R_k^{-1}.$$

Finally, although the Cholesky factor $\hat{P}_{k|k}^{1/2}$ is already available in the MCC-KF algorithm, the estimator implementation implies its re-calculation by using the so-called Joseph stabilized equation for updating the error covariance matrix $P_{k|k}$ at the last line of Algorithm 1. To distinguish the matrix $P_{k|k}$ used in the gain K_k calculation by formula (9) from the matrix obtained from the Joseph stabilized equation, we use notation $\hat{P}_{k|k}$ and $P_{k|k}$ for these two cases, respectively. It is worth noting here that the Joseph stabilized formula ensures the symmetric form of error covariance matrix $P_{k|k}$ and, hence, it is recognized to be the preferable implementation strategy of any conventional filtering algorithm. In a similar way, we factorize

$$AA^{\top} = (I - K_k H_k) P_{k|k-1} (I - K_k H_k)^{\top} + K_k R_k K_k^{\top}$$
$$= [(I - K_k H_k) P_{k|k-1}^{1/2},\ K_k R_k^{1/2}]$$
$$\times [(I - K_k H_k) P_{k|k-1}^{1/2},\ K_k R_k^{1/2}]^{\top}$$
$$= RR^{\top} = P_{k|k} = [P_{k|k}^{1/2},\ 0][P_{k|k}^{1/2},\ 0]^{\top}.$$

Having summarized the formulas above, we obtain the following Cholesky-based square-root MCC-KF method.

Algorithm 1a. SR MCC-KF (*Cholesky-based MCC-KF*)

INITIALIZATION: $(k = 0)$

Apply Cholesky factorization: $\Pi_0 = \Pi_0^{1/2}\Pi_0^{\top/2}$;

Set initial values: $\hat{x}_{0|0} = \bar{x}_0$, $P_{0|0}^{1/2} = \Pi_0^{1/2}$;

TIME UPDATE: $(k = \overline{1,N})$

1 Compute $\hat{x}_{k|k-1}$ in line 1 of Algorithm 1;

2 Build the pre-array and lower triangularize it

$$\underbrace{\Big[F_{k-1}P_{k-1|k-1}^{1/2},\ G_{k-1}Q_{k-1}^{1/2}\Big]}_{\text{Pre-array } \mathbb{A}} \mathbb{Q}_1 = \underbrace{\Big[P_{k|k-1}^{1/2},\ 0\Big]}_{\text{Post-array } \mathbb{R}};$$

Read-off from post-array the factor $P_{k|k-1}^{1/2}$;

MEASUREMENT UPDATE: $(k = \overline{1,N})$

3 Compute λ_k by formula (8);

4 Build the pre-array and lower triangularize it

$$\underbrace{\Big[P_{k|k-1}^{-\top/2},\ \lambda_k^{1/2} H_k^{\top} R_k^{-\top/2}\Big]}_{\text{Pre-array } \mathbb{A}} \mathbb{Q}_2 = \underbrace{\Big[\hat{P}_{k|k}^{-\top/2},\ 0\Big]}_{\text{Post-array } \mathbb{R}};$$

Read-off from post-array the factor $[\hat{P}_{k|k}^{-\top/2}]$;

5 Compute $K_k = \lambda_k [\hat{P}_{k|k}^{-\top/2}]^{-\top} [\hat{P}_{k|k}^{-\top/2}]^{-1} H_k^{\top} R_k^{-1}$;

6 Calculate the state $\hat{x}_{k|k}$ in line 6 of Algorithm 1;

7 Build the pre-array and lower triangularize it

$$\underbrace{\Big[(I - K_k H_k) P_{k|k-1}^{1/2},\ K_k R_k^{1/2}\Big]}_{\text{Pre-array } \mathbb{A}} \mathbb{Q}_3 = \underbrace{\Big[P_{k|k}^{1/2},\ 0\Big]}_{\text{Post-array } \mathbb{R}};$$

Read-off from post-array the factor $P_{k|k}^{1/2}$.

It is not difficult to see that the numerical behaviour of Algorithm 1a heavily depends on a condition number of matrices $P_{k|k-1}^{1/2} \in \mathbb{R}^{n\times n}$ and $\hat{P}_{k|k}^{1/2} \in \mathbb{R}^{n\times n}$ because their inversion is required in Algorithm 1a. The method can be improved by avoiding the matrix inversion operation.

We construct an alternative robust Cholesky-based variant by using equation (10) for computing the gain matrix

K_k. As a result, our new square-root method requires the inverse of the lower triangular matrix $R_{e,k}^{1/2} \in \mathbb{R}^{m\times m}$, only. Indeed, we factorize equation $R_{e,k} = \lambda_k H_k P_{k|k-1} H^\top + R_k$ in a similar way as shown above and, then, utilize a stable orthogonal transformation for updating the resulted square-root factors, i.e. we get

$$\begin{aligned} AA^\top &= \lambda_k H_k P_{k|k-1} H^\top + R_k \\ &= [\lambda_k^{1/2} H_k P_{k|k-1}^{1/2},\ R_k^{1/2}][\lambda_k^{1/2} H_k P_{k|k-1}^{1/2},\ R_k^{1/2}]^\top \\ &= RR^\top = R_{e,k} = [R_{e,k}^{1/2},\ 0][R_{e,k}^{1/2},\ 0]^\top, \end{aligned}$$

and the gain matrix K_k is then calculated by equation (10) by using the available value $[R_{e,k}^{1/2}]$ as follows:

$$K_k = \lambda_k P_{k|k-1} H_k^\top R_{e,k}^{-1} = \lambda_k P_{k|k-1} H_k^\top [R_{e,k}^{1/2}]^{-\top} [R_{e,k}^{1/2}]^{-1}.$$

Thus, we summarize an alternative Cholesky-based implementation for the MCC-KF estimator.

Algorithm 1b. SR MCC-KF

Initialization: Repeat from Algorithm 1a.
Time Update: Repeat from Algorithm 1a.
Measurement Update: ($k = \overline{1, N}$)

1 Compute λ_k by formula (8);
2 Build the pre-array and lower triangularize it

$$\underbrace{\left[\lambda_k^{1/2} H_k P_{k|k-1}^{1/2},\ R_k^{1/2}\right]}_{\text{Pre-array } \mathbb{A}} \mathbb{Q}_2 = \underbrace{\left[R_{e,k}^{1/2},\ 0\right]}_{\text{Post-array } \mathbb{R}};$$

Read-off from post-array the factor $[R_{e,k}^{1/2}]$;
3 Compute $K_k = \lambda_k P_{k|k-1} H_k^\top [R_{e,k}^{1/2}]^{-\top} [R_{e,k}^{1/2}]^{-1}$;
4 Compute $\hat{x}_{k|k}$ by formula in line 6 of Algorithm 1;
5 Use formula in line 7 of Algorithm 1a to find $P_{k|k}^{1/2}$.

As can be seen, the suggested square-root Algorithm 2b requires the innovation covariance matrix $R_{e,k}^{1/2}$ inversion, only. Thus, it is expected to possess a better numerical behavior compared to the conventional implementation in Algorithm 1 and the Cholesky-based Algorithm 1a. Next section provides the results of numerical experiments.

4. NUMERICAL EXPERIMENTS

We first wish to justify a theoretical derivation of the presented square-root MCC-KF Algorithm 1b.

Example 1. In the Radar Tracking Example from (Grewal and Andrews, 2015, p. 227), the signals are processed by the filter in order to determine the position of maneuvering airborne objects. The system state is defined as follows: $x_k = [r_k,\ \dot{r}_k,\ U_k^1,\ \theta_k,\ \dot{\theta}_k,\ U_k^2]^\top$ where r_k is the range of the vehicle at time t_k, $\dot{r}_k$ is the range rate of the vehicle at time t_k, U_k^1 is the maneuvering-correlated state noise, θ_k is the bearing of the vehicle at time t_k, $\dot{\theta}_k$ is the bearing rate of the vehicle at time t_k, U_k^2 is the maneuvering-correlated state noise. The model dynamics is given as follows:

$$x_k = \begin{bmatrix} 1 & T & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & \rho & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & T & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & \rho \end{bmatrix} x_{k-1} + \begin{bmatrix} 0 \\ 0 \\ w_{k-1}^1 \\ 0 \\ 0 \\ w_{k-1}^2 \end{bmatrix},$$

where $\rho = 0.5$ is correlation coefficient and $T = 10$ is the sampling period in seconds. The measurements are provided by

$$y_k = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} x_k + \begin{bmatrix} v_k^1 \\ v_k^2 \end{bmatrix}$$

The MCC-KF estimators are tested in the presence of, the so-called, impulsive noise scenario, which is used for simulating the outliers in the examined Gaussian state-space model:

$$\begin{aligned} w_k &\sim \mathcal{N}(0, Q) + \texttt{Shot noise}(20\% \text{ are corrupted}), \\ v_k &\sim \mathcal{N}(0, R) + \texttt{Shot noise}(20\% \text{ are corrupted}) \end{aligned}$$

where the covariance matrices Q and R are given by

$$Q = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_1^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_2^2 \end{bmatrix} \quad \text{and} \quad R = \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_\theta^2 \end{bmatrix}$$

with $\sigma_r^2 = (1000\ \text{m})^2$, $\sigma_\theta^2 = (0.017\ \text{rad})^2$, $\sigma_1^2 = (103/3)^2$ and $\sigma_2^2 = 1.3 \times 10^{-8}$. Finally, the initial values for each estimator to be examined are the following: $x_0 \sim \mathcal{N}(\bar{x}_0, \Pi_0)$ where $\bar{x}_0 = 0$ and

$$\Pi_0 = \begin{bmatrix} \sigma_r^2 & \frac{\sigma_r^2}{T} & 0 & 0 & 0 & 0 \\ \frac{\sigma_r^2}{T} & \frac{2\sigma_r^2}{T^2} + \sigma_1^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_1^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_\theta & \frac{\sigma_\theta}{T} & 0 \\ 0 & 0 & 0 & \frac{\sigma_\theta}{T} & \frac{2\sigma_\theta}{T^2} + \sigma_1^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_2^2 \end{bmatrix}.$$

To simulate the impulsive noise (the shot noise), we follow the approach suggested in Izanloo et al. (2016). The corresponding Matlab routine `Shot_noise` is also published recently in Kulikova (2020). The magnitude of each impulse is chosen randomly from the uniform discrete distribution in the interval $[0, 5]$. The time instances where the outliers occur are also chosen randomly on the interval $t_k \in [21, 300]$ by using the uniform distribution. When the 'exact' solution $x_{exact}(t_k)$ and the measurements $y_k = y(t_k)$ are simulated on the examined interval $t_k \in [1, 300]$, we solve the inverse problem, i.e. the dynamic state is to be estimated from the observed signal $y(t_k)$ by various filtering algorithms. They are all tested within equal conditions, i.e. they utilize the same filters' initials, the same measurements and the same noise covariances. Finally, the experiment is repeated for $M = 100$ times and the root mean square error (RMSE) is calculated in these Monte Carlo runs.

Fig. 1 illustrates the total RMSE in all six components of the dynamic state averaged over $M = 100$ Monte Carlo runs. As can be seen, all implementation methods produce the same result, i.e. they work with the same estimation accuracy. This substantiates an algebraic equivalence of all MCC-KF algorithms under examination and the correctness of our theoretical derivations presented in Section 3. Meanwhile, the difference in numerical behaviour of the conventional Algorithm 1, the previously suggested square-root Algorithm 1a and the newly derived square-

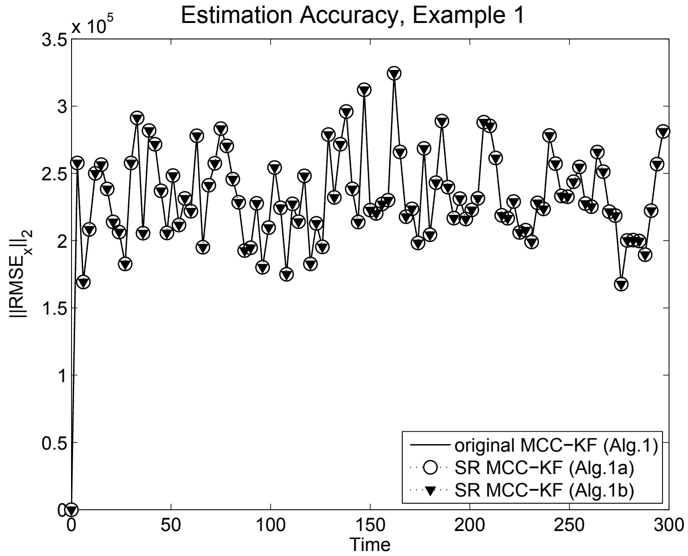

Fig. 1. The total $\|\text{RMSE}_{x_i}(t_k)\|_2$ $(i = 1, \ldots, 6)$ calculated for each MCC-KF estimator in case of radar tracking scenario in Example 1.

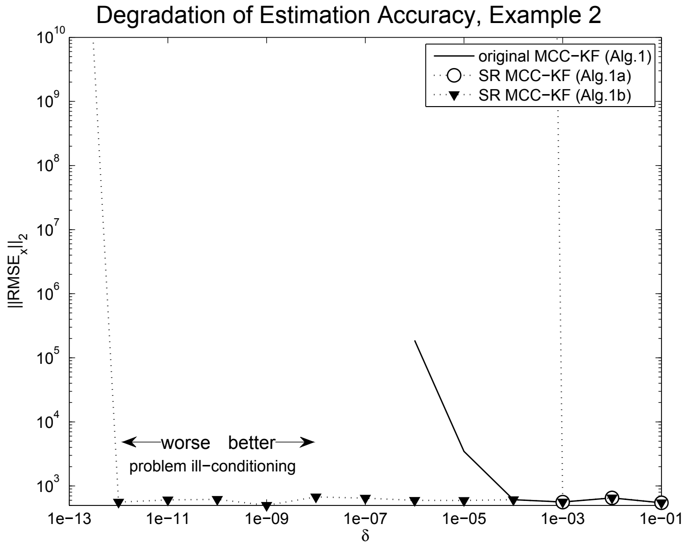

Fig. 2. The degradation of total RMSE for each MCC-KF estimator under examination in case of ill-conditioned radar tracking cases in Example 2.

root Algorithm 1b can be observed in ill-conditioned state estimation scenario. For that, we consider Example 2.

Example 2. Consider the radar tracking problem in Example 1 with Gaussian uncertainties only, where the dynamic state is observed through the following measurement scheme:

$$y_k = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1+\delta \end{bmatrix} x_k + \begin{bmatrix} v_k^1 \\ v_k^2 \end{bmatrix}$$

where $x_0 \sim \mathcal{N}(0, I_6)$ and $v_k \sim \mathcal{N}(0, \delta^2 I_2)$. The parameter δ is used for simulating roundoff effect, i.e. it is assumed to be $\delta^2 < \epsilon_{roundoff}$, but $\delta > \epsilon_{roundoff}$ and $\epsilon_{roundoff}$ stands for the unit roundoff error.

Measurement scheme in Example 2 allows for simulating various ill-conditioned scenarios including the continuous-discrete estimation methods as discussed in Kulikov and Kulikova (2017, 2018b, 2019). More precisely, when the ill-conditioning parameter δ tends to machine precision limit, $\delta \to \epsilon_{roundoff}$, we observe a degradation of the underlying Riccati-type recursion. The discussed numerical instability of the conventional filtering methods is arisen from the matrix inversion $R_{e,k}$ that becomes almost singular after a few filtering steps; see the third reason of the KF divergence in (Grewal and Andrews, 2015, p. 288). Within the square-root implementation approach, the inverse of its triangular Cholesky factor is required instead.

Fig. 2 illustrates a degradation of the resulted estimation accuracies when $\delta \to \epsilon_{roundoff}$ for each MCC-KF implementation under examination. As it was anticipated the square-root MCC-KF Algorithm 1b is the most robust implementation method among three algorithms examined. We observe that all MCC-KF methods work equally accurate and with small estimation errors in well-conditioned scenarios, i.e. when δ is large, which corresponds to ill-conditioned parameter $\delta = 10^{-1}$ and $\delta = 10^{-2}$. However, while the problem ill-conditioning increases the conventional Algorithm 1 and the square-root Algorithm 1a fail to solve the stated problem. Indeed, for $\delta \leq 10^{-4}$ they both provide either a very large total RMSE or yield `NaN` (it states for 'Not a Number' in MATLAB), which is not plotted. It is worth noting here that the conventional MCC-KF implementation in Algorithm 1 degrades even a bit slower than the previously published square-root Algorithm 1a. A possible explanation for such behaviour is that the conventional MCC-KF implementation implies the Joseph stabilized equation for the error covariance matrix calculation. Indeed, the MCC-KF estimator in Algorithm 1 fails at $\delta = 10^{-5}$ because the resulted accuracies are very large, i.e. it means no correct digits in the obtained estimates. Meanwhile, the square-root Algorithm 1a fails a bit faster, i.e. at $\delta = 10^{-4}$. Finally, the newly-derived square-root method in Algorithm 1b is the most stable implementation method. It works accurately until the ill-conditioned parameter $\delta = 10^{-13}$, i.e. it is able to manage the ill-conditioned scenarios of Example 2. This is the only one robust Cholesky-based implementation existed for the MCC-KF estimator in engineering literature so far.

5. CONCLUSION

In this paper, the problem of designing the numerically stable square-root methods for the maximum correntropy criterion Kalman filter is discussed. The first robust (with respect to roundoff errors) square-root MCC-KF algorithm has been found within the class of Cholesky factorization-based implementations. Nowadays, this is the only one reliable Cholesky-based method existed for the MCC-KF estimator in engineering literature. The square-root solution is proposed for the MCC-KF filtering for a case of the scalar adjusting parameter involved. The derivation of a square-root solution within the Cholesky decomposition for the MCC-KF methods with matrix-type adjusting weights involved is an open question for a future research.

REFERENCES

Aravkin, A., Burke, J.V., Ljung, L., Lozano, A., and Pillonetto, G. (2017). Generalized Kalman smoothing: Modeling and algorithms. *Automatica*, 86, 63–86.

Bierman, G.J. and Thornton, C.L. (1977). Numerical comparison of Kalman filter algorithms: Orbit determination case study. *Automatica*, 13(1), 23–35.

Chen, B., Liu, X., Zhao, H., and Príncipe, J.C. (2017). Maximum Correntropy Kalman Filter. *Automatica*, 76, 70–77.

Chen, B., Wang, J., Zhao, H., Zheng, N., and Príncipe, J.C. (2015). Convergence of a fixed-point algorithm under maximum correntropy criterion. *IEEE Signal Processing Letters*, 22(10), 1723–1727.

Chen, B., Xing, L., Liang, J., Zheng, N., and Príncipe, J.C. (2014). Steady-state mean-square error analysis for adaptive filtering under the maximum correntropy criterion. *IEEE Signal Processing Letters*, 21(7), 880–884.

Cinar, G.T. and Príncipe, J.C. (2011). Adaptive background estimation using an information theoretic cost for hidden state estimation. In *The 2011 International Joint Conference on Neural Networks (IJCNN)*, 489–494.

Cinar, G.T. and Príncipe, J.C. (2012). Hidden state estimation using the Correntropy filter with fixed point update and adaptive kernel size. In *The 2012 International Joint Conference on Neural Networks (IJCNN)*, 1–6.

Comminiello, D. and Príncipe, J.C. (2018). *Adaptive Learning Methods for Nonlinear System Modeling*. Elsevier, 1st edition edition.

Fakoorian, S., Mohammadi, A., Azimi, V., and Simon, D. (2019). Robust Kalman-type filter for non-Gaussian noise: Performance analysis with unknown noise covariances. *Journal of Dynamic Systems, Measurement, and Control*, 141(9), 091011.

Grewal, M.S. (2019). Practical design and implementation methods for Kalman filtering for mission critical applications. *Navigation*, 66(1), 239–249.

Grewal, M.S. and Andrews, A.P. (2015). *Kalman Filtering: Theory and Practice using MATLAB*. John Wiley & Sons, New Jersey, 4-th edition edition.

Grewal, M.S. and Kain, J. (2010). Kalman filter implementation with improved numerical properties. *IEEE Transactions on Automatic Control*, 55(9), 2058–2068.

Izanloo, R., Fakoorian, S.A., Yazdi, H.S., and Simon, D. (2016). Kalman filtering based on the maximum correntropy criterion in the presence of non-Gaussian noise. In *2016 Annual Conference on Information Science and Systems (CISS)*, 500–505.

Kailath, T., Sayed, A.H., and Hassibi, B. (2000). *Linear Estimation*. Prentice Hall, New Jersey.

Kulikov, G.Yu. and Kulikova, M.V. (2017). Square-root Kalman-like filters for estimation of stiff continuous-time stochastic systems with ill-conditioned measurements. *IET Control Theory & Applications*, 9(11), 1420–1425.

Kulikov, G.Yu. and Kulikova, M.V. (2018a). Estimation of maneuvering target in the presence of non-Gaussian noise: A coordinated turn case study. *Signal Process.*, 145, 241–257.

Kulikov, G.Yu. and Kulikova, M.V. (2018b). Moore-Penrose-pseudo-inverse-based Kalman-like filtering methods for estimation of stiff continuous-discrete stochastic systems with ill-conditioned measurements. *IET Control Theory & Applications*, 12(16), 2205–2212.

Kulikov, G.Yu. and Kulikova, M.V. (2019). Numerical robustness of extended Kalman filtering based state estimation in ill-conditioned continuous-discrete nonlinear stochastic chemical systems. *International Journal of Robust and Nonlinear Control*, 5(29), 1377–1395.

Kulikov, G.Yu. and Kulikova, M.V. (2020). A comparative study of Kalman-like filters for state estimation of turning aircraft in presence of glint noise. In *Proceedings of 21st IFAC World Congress*. (in press).

Kulikova, M.V. (2017). Square-root algorithms for maximum correntropy estimation of linear discrete-time systems in presence of non-Gaussian noise. *Systems & Control Letters*, 108, 8–15.

Kulikova, M.V. (2019). Factored-form Kalman-like implementations under maximum correntropy criterion. *Signal Processing*, 160, 328–338.

Kulikova, M.V. (2020). Sequential maximum correntropy Kalman filtering. *Asian Journal of Control*, 22(1), 25–33.

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), 5286–5298.

Liu, X., Chen, B., Xu, B., Wu, Z., and Honeine, P. (2017a). Maximum correntropy unscented filter. *International Journal of Systems Science*, 48(8), 1607–1615.

Liu, X., Qu, H., Zhao, J., and Chen, B. (2016). Extended Kalman filter under maximum correntropy criterion. In *International Joint Conference on Neural Networks (IJCNN)*, 1733–1737.

Liu, X., Qu, H., Zhao, J., and Chen, B. (2017b). State space maximum correntropy filter. *Signal Processing*, 130, 152–158.

Morf, M. and Kailath, T. (1975). Square-root algorithms for least-squares estimation. *IEEE Transactions on Automatic Control*, 20(4), 487–497.

Park, P. and Kailath, T. (1995). New square-root algorithms for Kalman filtering. *IEEE Transactions on Automatic Control*, 40(5), 895–899.

Qin, W., Wang, X., and Cui, N. (2017). Maximum correntropy sparse Gauss-Hermite quadrature filter and its application in tracking ballistic missile. *IET Radar, Sonar & Navigation*, 11(9), 1388–1396.

Simon, D. (2006). *Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches*. John Wiley & Sons.

Wang, G., Li, N., and Zhang, Y. (2017). Maximum correntropy unscented Kalman and information filters for non-Gaussian measurement noise. *Journal of the Franklin Institute*, 354(18), 8659–8677.

Wang, Y., Zheng, W., Sun, S., and Li, L. (2016). Robust information filter based on maximum correntropy criterion. *Journal of Guidance, Control, and Dynamics*, 1126–1131.

Yang, Y. and Huang, G. (2017). Map-based localization under adversarial attacks. In *International Symposium on Robotics Research (ISRR), Puerto Varas, Chile*.

Yang, Y. and Huang, G. (2018). Attack-resilient Map-based localization. In *RSS Workshop: Adversarial Robotics, Carnegie Mellon University, Pittsburgh, USA*.