

One-Step Condensed Forms for Square-Root Maximum Correntropy Criterion Kalman Filtering

Maria V. Kulikova

Center for Computational and Stochastic Mathematics

Instituto Superior Técnico, Universidade de Lisboa

Lisbon, Portugal

maria.kulikova at ist.utl.pt

Abstract—This paper suggests a few novel Cholesky-based square-root algorithms for the maximum correntropy criterion Kalman filtering. In contrast to the previously obtained results, new algorithms are developed in the so-called *condensed* form that corresponds to the *a priori* filtering. Square-root filter implementations are known to possess a better conditioning and improved numerical robustness when solving ill-conditioned estimation problems. Additionally, the new algorithms permit easier propagation of the state estimate and do not require a back-substitution for computing the estimate. Performance of novel filtering methods is examined by using a fourth order benchmark navigation system example.

Keywords—maximum correntropy, Kalman filter, Cholesky decomposition, one-step filtering, condensed form

I. INTRODUCTION

The Kalman filtering (KF) like estimators developed under the maximum correntropy criterion (MCC) methodology are shown to outperform the classical KF for estimation quality in case of non-Gaussian and impulsive noise scenario in state-space models, both for linear and nonlinear systems [1]–[10]. Among nonlinear MCC filtering strategies, we may mention the MCC extended KF in [11], [12], the MCC unscented KF in [13], [14], the MCC-based Gauss-Hermite quadrature filter in [15] and the so-called accurate continuous-discrete extended KF [16], [17] explored under the MCC in [18]. Here, we focus on the MCC-KF estimators suggested in [6], [9].

The numerical stability issues of the cited MCC-KF estimators are of special interest in this paper. Recently, their reliable square-root implementation methods have been designed in [19]. Following the approaches existed in the KF community, the new MCC-KF implementations belong to three main *factored-form* classes: (i) the Cholesky-based algorithms; e.g., [20]–[22], (ii) the UDU^T factorization-based methods, e.g., [23], [24], and (iii) singular value decomposition solution, e.g., [25], [26]. All mentioned square-root strategies have been utilized for the robust MCC-KF filtering in [19]. In the cited paper, the implementation methods are developed in terms of covariance quantities and in the *a posteriori* form that yields the two-stage implementations. In contrast to the previously

obtained results, the goal of this paper is to suggest square-root MCC-KF algorithms in the *a priori* form that results to the *condensed* implementations. We stress that the stated problem has never been solved in engineering literature, before.

More precisely, we develop the *a priori* covariance filtering approach for the original MCC-KF proposed in [6] and for its improved variant (IMCC-KF) published in [9]. Additionally, we discuss the square-root problem for the examined estimators and suggest a few condensed implementations for the robust filtering in terms of *lower* triangular Cholesky factors. As all square-root implementations, the new algorithms possess a better conditioning, improved numerical robustness (with respect to roundoff errors) and reliability when solving ill-conditioned estimation problems. Additionally, the new algorithms permit easier propagation of the state estimate and do not require a back-substitution for computing the estimate. It is achieved by utilizing orthogonal rotations for propagating the involved Cholesky factors as far as possible. Finally, we stress that derivation of new implementation methods is important for its own sake because it provides practitioners a diversity of methods with a fair possibility to choose any of them depending on practical application, complexity and accuracy requirements. Performance of novel filtering algorithms is examined by using a fourth order benchmark navigation system example.

The paper is organized as follows. Section II contains the problem statement and presents the derivation of the *a priori* MCC-KF recursions. The new square-root condensed implementation methods can be found in Section III together with their derivation. Section IV contains the results of numerical simulation study and Section V concludes the paper.

II. PROBLEM STATEMENT

Consider a linear discrete-time stochastic system

$$x_{k+1} = F_k x_k + G_k w_k, \quad (1)$$

$$y_k = H_k x_k + v_k, \quad k \geq 0 \quad (2)$$

where system matrices $F_k \in \mathbb{R}^{n \times n}$, $G_k \in \mathbb{R}^{n \times q}$ and $H_k \in \mathbb{R}^{m \times n}$ are known and constant over time. The vectors $x_k \in \mathbb{R}^n$ and $y_k \in \mathbb{R}^m$ are unknown dynamic state and

The author acknowledges the financial support of the Portuguese FCT — *Fundação para a Ciência e a Tecnologia*, through the project UID/Multi/04621/2019 of CEMAT/IST-ID, Center for Computational and Stochastic Mathematics, Instituto Superior Técnico, University of Lisbon.

available measurement vector, respectively. We assume that random variables x_0 , w_k and v_k satisfy

$$\begin{aligned} \mathbf{E}\{x_0\} &= \bar{x}_0, & \mathbf{E}\{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^\top\} &= \Pi_0, \\ \mathbf{E}\{w_k\} &= \mathbf{E}\{v_k\} = 0, & \mathbf{E}\{w_k w_0^\top\} &= \mathbf{E}\{v_k x_0^\top\} = 0, \\ \mathbf{E}\{w_k v_k^\top\} &= 0, & \mathbf{E}\{w_k w_j^\top\} &= Q_k \delta_{kj}, \\ & & \mathbf{E}\{v_k v_j^\top\} &= R_k \delta_{kj} \end{aligned}$$

where covariance matrices $Q_k \in \mathbb{R}^{q \times q}$ and $R_k \in \mathbb{R}^{m \times m}$ are known. The symbol δ_{kj} is the Kronecker delta function.

The classical KF produces the minimum *linear* expected mean square error (MSE) estimate of the unknown state vector x_k . It can be formulated in the *a priori* form as shown in [27, Theorem 9.2.1]; see also the derivation in [28, Section 5.3]. More precisely, the one-step ahead predicted estimate $\hat{x}_{k+1|k}$ (*a priori* estimate) is calculated by

$$\hat{x}_{k+1|k} = F_k \hat{x}_{k|k-1} + K_{p,k}(y_k - H_k \hat{x}_{k|k-1}), \quad (3)$$

$$R_{e,k} = R_k + H_k P_{k|k-1} H_k^\top, \quad (4)$$

$$K_{p,k} = F_k P_{k|k-1} H_k^\top R_{e,k}^{-1} \quad (5)$$

where $e_k = y_k - H_k \hat{x}_{k|k-1}$ are innovations (residuals) of discrete-time KF with covariance $R_{e,k} = \mathbf{E}\{e_k e_k^\top\}$ and $K_{p,k} = \mathbf{E}\{\hat{x}_{k+1|k} e_k^\top\} = F_k P_{k|k-1} H_k^\top$. The matrix $P_{k|k-1}$ is one-step ahead predicted error covariance $P_{k|k-1} = \mathbf{E}\{(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^\top\}$ computed via the Riccati difference recursion:

$$P_{k+1|k} = F_k P_{k|k-1} F_k^\top + G_k Q_k G_k^\top - K_{p,k} R_{e,k} K_{p,k}^\top \quad (6)$$

with initial values $P_{0|-1} = \Pi_0 > 0$ and $\hat{x}_{0|-1} = \bar{x}_0$.

Being a *linear* estimator, the classical KF exhibits only sub-optimal behavior in non-Gaussian settings. To enhance its estimation quality and robustness with respect to outliers (impulsive noise), the maximum correntropy criterion has been utilized together with the KF estimation cost function in [5], [6]. In general, the *correntropy* is a similarity measure between two random variables [1]. Following [29, Chapter 5], it can be used in the estimation problem as follows: an estimator of unknown state $X \in \mathbb{R}$ is defined as a function of measurements $Y \in \mathbb{R}^m$, i.e., $\hat{X} = g(Y)$ where g is solved by maximizing the correntropy between X and \hat{X} that is [30]

$$\arg \max_{g \in G} V(X, \hat{X}) = \arg \max_{g \in G} \mathbf{E}\left\{k_\sigma(X - g(Y))\right\} \quad (7)$$

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. The Gaussian kernel is the most widely used kernel function and it is given as follows:

$$k_\sigma(X - \hat{X}) = \exp\left\{-(X - \hat{X})^2/(2\sigma^2)\right\}. \quad (8)$$

It is not difficult to see that the MCC cost (7) with Gaussian kernel (8) reaches its maximum if and only if $X = \hat{X}$.

In [6], the MCC-KF is developed by solving the following estimation problem with the Gaussian kernel:

$$\arg \max \left\{ k_\sigma(\|\hat{x}_{k|k} - F_{k-1} \hat{x}_{k-1|k-1}\|) + k_\sigma(\|y_k - H_k \hat{x}_{k|k}\|) \right\}$$

where weighted norms are additionally utilized.

Next, the resulted nonlinear equation arisen in the optimization problem above is solved by a fixed point rule that yields the following recursion [6, p. 503]:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k^\lambda (y_k - H_k \hat{x}_{k|k-1}) \quad (9)$$

where $\hat{x}_{k|k-1} := F_{k-1} \hat{x}_{k-1|k-1}$ and the gain matrix is defined as $K_k^\lambda = \lambda_k (P_{k|k-1}^{-1} + \lambda_k H_k^\top R_k^{-1} H_k)^{-1} H_k^\top R_k^{-1}$ with a scalar adjusting weight λ_k given as follows:

$$\lambda_k = \frac{k_\sigma(\|y_k - H_k \hat{x}_{k|k-1}\|_{R_k^{-1}})}{k_\sigma(\|\hat{x}_{k|k-1} - F_{k-1} \hat{x}_{k-1|k-1}\|_{P_{k|k-1}^{-1}})}. \quad (10)$$

Finally, the resulted MCC-KF estimator in [6] utilizes the recursion for state estimate in (9) together with a symmetric Joseph stabilized equation for calculating the error covariance matrix $P_{k|k}$ that is known to improve the numerical robustness of KF-like implementations because of ensured symmetric form of $P_{k|k}$ in the presence of roundoff errors [28], [31]. In summary, the MCC-KF estimator is given as follows [6].

MCC-KF estimator. TIME UPDATE ($k = \overline{1, K}$).

$$\hat{x}_{k|k-1} = F_{k-1} \hat{x}_{k-1|k-1}, \quad (11)$$

$$P_{k|k-1} = F_{k-1} P_{k-1|k-1} F_{k-1}^\top + G_{k-1} Q_{k-1} G_{k-1}^\top. \quad (12)$$

MEASUREMENT UPDATE. Find *a posteriori* estimate:

$$K_k^\lambda = \lambda_k \left(P_{k|k-1}^{-1} + \lambda_k H_k^\top R_k^{-1} H_k \right)^{-1} H_k^\top R_k^{-1}, \quad (13)$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k^\lambda e_k, \quad e_k = y_k - H_k \hat{x}_{k|k-1}, \quad (14)$$

$$P_{k|k} = (I - K_k^\lambda H_k) P_{k|k-1} (I - K_k^\lambda H_k)^\top + K_k^\lambda R_k [K_k^\lambda]^\top \quad (15)$$

where the scaling (inflation) parameter λ_k is computed by (10).

In [9], the estimation quality of original MCC-KF estimator presented above has been boosted by deriving mathematically equivalent formulas for the gain and error covariance, similar to the KF equations in [28, pp. 128-129]. It yields the improved MCC-KF (IMCC-KF) estimator summarized below.

IMCC-KF estimator. TIME UPDATE ($k = \overline{1, K}$). Calculate *a priori* estimate $\hat{x}_{k|k-1}$ and $P_{k|k-1}$ by equations (11), (12).

MEASUREMENT UPDATE. Find *a posteriori* estimate $\hat{x}_{k|k}$ by formula (14) with the scaling parameter λ_k from (10) and

$$K_k^\lambda = \lambda_k P_{k|k-1} H_k^\top [R_{e,k}^\lambda]^{-1}, \quad (16)$$

$$R_{e,k}^\lambda = \lambda_k H_k P_{k|k-1} H_k^\top + R_k, \quad (17)$$

$$P_{k|k} = (I - K_k^\lambda H_k) P_{k|k-1}. \quad (18)$$

As can be seen, the original MCC-KF and IMCC-KF equations recursively updates *a posteriori* estimate $\hat{x}_{k|k}$ given the initial value $\hat{x}_{0|0} := \bar{x}_0$. The discussed estimators can be re-formulated in the so-called *a priori* form similar to the classical KF equations (3) – (6). This yields one-step KF-like equations and allows for convenient *condensed* form of the resulted square-root implementation methods. Next, we derive one-step equations for both the MCC-KF and IMCC-KF estimators and explore the possibility of designing their Cholesky factorization-based implementations.

A. The MCC-KF estimator in a priori form

Following [28, Section 5.3], we derive *a priori* MCC-KF recursion related to the original MCC-KF equations (11)–(15). Having substituted (14) into (11), we get the recursion for one-step ahead predicted (*a priori*) estimate $\hat{x}_{k+1|k}$ as follows:

$$\begin{aligned}\hat{x}_{k+1|k} &= F_k \hat{x}_{k|k} = F_k (\hat{x}_{k|k-1} + K_k^\lambda e_k), \\ &= F_k \hat{x}_{k|k-1} + F_k K_k^\lambda e_k = F_k \hat{x}_{k|k-1} + K_{p,k}^\lambda e_k\end{aligned}\quad (19)$$

where we introduced the notation $K_{p,k}^\lambda = F_k K_k^\lambda$.

It is proved in [9] that equations (13) and (16) for gain calculation K_k^λ are equivalent. Hence, we have

$$K_{p,k}^\lambda = \lambda_k F_k \left(P_{k|k-1}^{-1} + \lambda_k H_k^\top R_k^{-1} H_k \right)^{-1} H_k^\top R_k^{-1} \quad (20)$$

$$= \lambda_k F_k P_{k|k-1} H_k^\top [R_{e,k}^\lambda]^{-1}. \quad (21)$$

Having substituted formula (15) into (12) and taking into account formula (21) and $K_{p,k}^\lambda = F_k K_k^\lambda$, we obtain

$$\begin{aligned}P_{k+1|k} &= F_k P_{k|k} F_k^\top + G_k Q_k G_k^\top \\ &= F_k (I - K_k^\lambda H_k) P_{k|k-1} (I - K_k^\lambda H_k)^\top F_k^\top \\ &\quad + F_k K_k^\lambda R_k [K_k^\lambda]^\top F_k^\top + G_k Q_k G_k^\top \\ &= (F_k - K_{p,k}^\lambda H_k) P_{k|k-1} (F_k - K_{p,k}^\lambda H_k)^\top \\ &\quad + K_{p,k}^\lambda R_k [K_{p,k}^\lambda]^\top + G_k Q_k G_k^\top \\ &= F_k P_{k|k-1} F_k^\top + G_k Q_k G_k^\top - F_k P_{k|k-1} H_k^\top [K_{p,k}^\lambda]^\top \\ &\quad - K_{p,k}^\lambda H_k P_{k|k-1} F_k^\top \\ &\quad + K_{p,k}^\lambda (H_k P_{k|k-1} H_k^\top + R_k) [K_{p,k}^\lambda]^\top.\end{aligned}$$

Taking into account that $H_k P_{k|k-1} H_k^\top + R_k = R_{e,k}$, formula (21) and symmetric form of $R_{e,k}$ and $R_{e,k}^\lambda$, we get

$$F_k P_{k|k-1} H_k^\top = \lambda_k^{-1} K_{p,k}^\lambda R_{e,k}^\lambda,$$

we obtain the recursion for *a priori* error covariance $P_{k+1|k}$:

$$\begin{aligned}P_{k+1|k} &= F_k P_{k|k-1} F_k^\top + G_k Q_k G_k^\top - 2\lambda_k^{-1} K_{p,k}^\lambda R_{e,k}^\lambda [K_{p,k}^\lambda]^\top \\ &\quad + K_{p,k}^\lambda R_{e,k} [K_{p,k}^\lambda]^\top = F_k P_{k|k-1} F_k^\top + G_k Q_k G_k^\top \\ &\quad - K_{p,k}^\lambda [2\lambda_k^{-1} R_{e,k}^\lambda - R_{e,k}] [K_{p,k}^\lambda]^\top \\ &= F_k P_{k|k-1} F_k^\top + G_k Q_k G_k^\top \\ &\quad - K_{p,k}^\lambda \left(H_k P_{k|k-1} H_k^\top + (2\lambda_k^{-1} - 1) R_k \right) [K_{p,k}^\lambda]^\top.\end{aligned}$$

As can be seen, the MCC-KF equation for matrix $P_{k+1|k}$ computation does not allow the related square-root form because the term $2\lambda_k^{-1} - 1$ might be negative, although the adjusting weight λ_k defined in (10) is a nonnegative value. Meanwhile, for the improved MCC-KF counterpart the efficient square-root implementations do exist as shown below.

B. The IMCC-KF estimator in a priori form

Similarly, *a priori* recursion for the IMCC-KF equations (11), (12), (14), (16)–(18) has been recently derived in [32]. For readers' convenience, we briefly discuss it here. First, one-step ahead predicted estimate $\hat{x}_{k+1|k}$ is calculated by equation (19) as in the MCC-KF examined above.

Next, having substituted formula (18) into (12) and taking into account formula (21) and symmetric form of any covariance matrix, we obtain

$$\begin{aligned}P_{k+1|k} &= F_k P_{k|k} F_k^\top + G_k Q_k G_k^\top \\ &= F_k [(I - K_k^\lambda H_k) P_{k|k-1}] F_k^\top + G_k Q_k G_k^\top \\ &= F_k P_{k|k-1} F_k^\top + G_k Q_k G_k^\top - F_k K_k^\lambda H_k P_{k|k-1} F_k^\top \\ &= F_k P_{k|k-1} F_k^\top + G_k Q_k G_k^\top - K_{p,k}^\lambda (F_k P_{k|k-1} H_k^\top)^\top \\ &= F_k P_{k|k-1} F_k^\top + G_k Q_k G_k^\top - \frac{1}{\lambda_k} K_{p,k}^\lambda R_{e,k}^\lambda [K_{p,k}^\lambda]^\top\end{aligned}$$

or, alternatively, we can avoid a scalar division by introducing

$$K_{p,k} = \frac{1}{\lambda_k} F_k K_k^\lambda = F_k P_{k|k-1} H_k^\top [R_{e,k}^\lambda]^{-1}, \quad (22)$$

i.e., $K_{p,k}^\lambda = \lambda_k K_{p,k}$ defined in (21). Thus, the equation for error covariance computation is re-formulated as follows:

$$P_{k+1|k} = F_k P_{k|k-1} F_k^\top + G_k Q_k G_k^\top - \lambda_k K_{p,k} R_{e,k}^\lambda K_{p,k}^\top \quad (23)$$

$$\text{where } K_{p,k} = F_k P_{k|k-1} H_k^\top [R_{e,k}^\lambda]^{-1}, \quad (24)$$

$$\hat{x}_{k+1|k} = F_k \hat{x}_{k|k-1} + \lambda_k K_{p,k} e_k. \quad (25)$$

Formulas (23)–(25) are the *a priori* form of the IMCC-KF estimator, which we are looking for. It is important that the adjusting weight λ_k defined in (10) is a nonnegative value and, hence, a square root exists (for real nonnegative numbers).

III. MAIN RESULTS: CONDENSED FORMS

The resulted one-step IMCC-KF equations (23) – (25) allow for convenient *condensed* form of stable square-root implementation methods that we derive in this section. Our square-root solution is based on the Cholesky decomposition 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 elements.

The condensed form implies that one-step IMCC-KF equations (23) – (25) are summarized into unique pre-array and, next, a stable orthogonal rotation is applied for obtaining the resulted post-array with the propagated filter quantities. Such implementation methods are additionally suitable for parallel implementation. Thus, we formulate the first condensed implementation method for the IMCC-KF estimator.

Algorithm 1. SR IMCC-KF (*a priori*, *one-step form*)

INITIALIZATION: ($k = 0$)

- 1 Apply Cholesky factorization: $\Pi_0 = \Pi_0^{1/2} \Pi_0^{\top/2}$;
 - 2 Set initial values: $\hat{x}_{0|-1} = \bar{x}_0$, $P_{0|-1}^{1/2} = \Pi_0^{1/2}$;
- FILTER RECURSION: ($k = \overline{0, K}$)
- 3 Compute λ_k and residual e_k by formulas (10), (14);
 - 4 Build pre-array and block-triangularize it

$$\underbrace{\begin{bmatrix} R_k^{1/2} & \lambda_k^{1/2} H_k P_{k|k-1}^{1/2} & 0 \\ 0 & F_k P_{k|k-1}^{1/2} & G_k Q_k^{1/2} \end{bmatrix}}_{\text{Pre-array } \mathbb{A}_k} \mathbb{Q}_k = \underbrace{\begin{bmatrix} [R_{e,k}^\lambda]^{1/2} & 0 & 0 \\ \bar{K}_{p,k}^\lambda & P_{k+1|k}^{1/2} & 0 \end{bmatrix}}_{\text{Post-array } \mathbb{R}_k};$$
 - 5 Read-off from post-array: $P_{k+1|k}^{1/2}$, $[R_{e,k}^\lambda]^{1/2}$, $\bar{K}_{p,k}^\lambda$;
 - 6 Find $\hat{x}_{k+1|k} = F_k \hat{x}_{k|k-1} + \lambda_k^{1/2} [\bar{K}_{p,k}^\lambda] [R_{e,k}^\lambda]^{-1/2} e_k$.

Following [21], it is not difficult to prove that the transformation in line 4 of Algorithm 1 implies the calculation by formulas (23) and (24) of the conventional one-step implementation. Let's consider the pre-array in line 4 of Algorithm 1

$$\begin{bmatrix} R_k^{1/2} & \lambda_k^{1/2} H_k P_{k|k-1}^{1/2} & 0 \\ 0 & F_k P_{k|k-1}^{1/2} & G_k Q_k^{1/2} \end{bmatrix}$$

and block-triangularize it by any orthogonal rotation, i.e.,

$$\mathbb{A}_k \mathbb{Q}_k = \begin{bmatrix} X & 0 & 0 \\ Y & Z & 0 \end{bmatrix}.$$

It is worth noting here that the orthogonal transformation \mathbb{Q}_k sets up a conformal (i.e., a norm- and angle-preserving) mapping between the (block) rows of the pre-array \mathbb{A}_k and post-array. Thus, we have

$$\begin{aligned} & \langle [R_k^{1/2}, \lambda_k^{1/2} H_k P_{k|k-1}^{1/2}, 0], [R_k^{1/2}, \lambda_k^{1/2} H_k P_{k|k-1}^{1/2}, 0] \rangle \\ & \quad = \langle [X, 0, 0], [X, 0, 0] \rangle, \\ & \langle [R_k^{1/2}, \lambda_k^{1/2} H_k P_{k|k-1}^{1/2}, 0], [0, F_k P_{k|k-1}^{1/2}, G_k Q_k^{1/2}] \rangle \\ & \quad = \langle [X, 0, 0], [Y, Z, 0] \rangle, \\ & \langle [0, F_k P_{k|k-1}^{1/2}, G_k Q_k^{1/2}], [0, F_k P_{k|k-1}^{1/2}, G_k Q_k^{1/2}] \rangle \\ & \quad = \langle [Y, Z, 0], [Y, Z, 0] \rangle. \end{aligned}$$

From the first equality above, we obtain

$$X X^\top = \lambda_k H_k (P_{k|k-1}^{1/2} P_{k|k-1}^{\top/2}) H_k^\top + R_k.$$

Having compared with equation (17), we conclude that $X X^\top = R_{e,k}^\lambda$ and, hence, $X := [R_{e,k}^\lambda]^{1/2}$.

At the same way, from the second equality we get

$$X Y^\top = \lambda_k^{1/2} H_k (P_{k|k-1}^{1/2} P_{k|k-1}^{\top/2}) F_k^\top$$

and, hence, we define $Y = \lambda_k^{1/2} F_k P_{k|k-1} H_k^\top [R_{e,k}^\lambda]^{-\top/2} := \bar{K}_{p,k}^\lambda$. Having compared the normalized quantity $\bar{K}_{p,k}^\lambda$ naturally appeared in square-root Algorithm 1 with the value $K_{p,k}$ defined by equation (24) in the conventional one-step method, we obtain the relationship

$$\bar{K}_{p,k}^\lambda = \lambda_k^{1/2} K_{p,k} [R_{e,k}^\lambda]^{1/2} \text{ or } K_{p,k} = \lambda_k^{-1/2} \bar{K}_{p,k}^\lambda [R_{e,k}^\lambda]^{-1/2}$$

Taking into account the formula above, we express equation (25) of the conventional method in terms of values available in the square-root Algorithm 1, i.e., in terms of $[\bar{K}_{p,k}^\lambda]$ and $[R_{e,k}^\lambda]^{1/2}$ that are directly read-off from the post-array

$$\begin{aligned} \hat{x}_{k+1|k} &= F_k \hat{x}_{k|k-1} + \lambda_k K_{p,k} e_k \\ &= F_k \hat{x}_{k|k-1} + \lambda_k^{1/2} [\bar{K}_{p,k}^\lambda] [R_{e,k}^\lambda]^{-1/2} e_k \end{aligned}$$

that is exactly the formula in line 6 of Algorithm 1.

Finally, we consider the last equality, i.e.,

$$F_k (P_{k|k-1}^{1/2} P_{k|k-1}^{\top/2}) F_k^\top + G_k Q_k G_k^\top = Y Y^\top + Z Z^\top.$$

Having substituted $Y = \bar{K}_{p,k}^\lambda$ and $\bar{K}_{p,k}^\lambda = \lambda_k^{1/2} K_{p,k} [R_{e,k}^\lambda]^{1/2}$, we get

$$\begin{aligned} Z Z^\top &= F_k P_{k|k-1} F_k^\top + G_k Q_k G_k^\top - \bar{K}_{p,k}^\lambda [\bar{K}_{p,k}^\lambda]^\top \\ &= F_k P_{k|k-1} F_k^\top + G_k Q_k G_k^\top - \lambda_k K_{p,k} R_{e,k}^\lambda K_{p,k}^\top \end{aligned}$$

i.e., from (23) we conclude $Z := P_{k+1|k}^{1/2}$. Thus, the square-root Algorithm 1 is justified.

The newly-derived square-root condensed form in Algorithm 1 is equivalent to conventional one-step formulas (23) – (25), but it is convenient for practical application and more suited to parallel computations. Besides, the square-root formulation makes the new Algorithm 1 is inherently more stable (with respect to roundoff errors) than any conventional one-step implementation; e.g., the one-step formulas (23) – (25).

We further suggest one more stable square-root implementation developed in the condensed form with the following additional benefit. As can be seen, Algorithms 1 allows for simple propagation of the filter quantities $[\bar{K}_{p,k}^\lambda]$ and $[R_{e,k}^\lambda]^{1/2}$ at each iteration step. Indeed, these values are simply read-off from the resulted post-array obtained from block-triangularization of the pre-array. Although stable orthogonal rotations are used for $[\bar{K}_{p,k}^\lambda]$ and $[R_{e,k}^\lambda]^{1/2}$ calculation, the state estimate computation still requires the $m \times m$ matrix inversion that is the inverse of $[R_{e,k}^\lambda]^{1/2}$; see line 6 of Algorithm 1. We stress that $[R_{e,k}^\lambda]^{1/2}$ is a triangular matrix and, hence, the inversion can be performed in efficient way (i.e., with the reduced computational cost) by solving the related linear equations through backward substitution. However, we can avoid the matrix inversion at all by the following trick. We augment the pre-array by a specially designed “data” column and apply the same orthogonal rotation to get the augmented post-array. Such algorithms are called the *extended* implementations and they skip the matrix inversion for computing the state estimate.

Following [21], consider the extended pre-array

$$\begin{bmatrix} R_k^{1/2} & \lambda_k^{1/2} H_k P_{k|k-1}^{1/2} & 0 \\ 0 & F_k P_{k|k-1}^{1/2} & G_k Q_k^{1/2} \\ -\lambda_k^{1/2} y_k^\top R_k^{-1/2} & \hat{x}_{k|k-1}^\top P_{k|k-1}^{-1/2} & 0 \end{bmatrix}$$

and lower-triangularize the first two (block) rows

$$\tilde{\mathbb{A}}_k \mathbb{Q}_k = \begin{bmatrix} [R_{e,k}^\lambda]^{1/2} & 0 & 0 \\ \bar{K}_{p,k}^\lambda & P_{k+1|k}^{1/2} & 0 \\ \alpha & \beta & \gamma \end{bmatrix}.$$

Similar to above presented derivation of Algorithm 1, we need to find α , β and γ . Taking into account properties of orthogonal matrices, from the inner product of the first and the last (block) rows of the pre-array, we get

$$\begin{aligned} & -\lambda_k^{1/2} R_k^{1/2} R_k^{-1/2} y_k + \lambda_k^{1/2} H_k P_{k|k-1}^{1/2} P_{k|k-1}^{-1/2} \hat{x}_{k|k-1} \\ & \quad = [R_{e,k}^\lambda]^{1/2} \alpha^\top = -\lambda_k^{1/2} (y_k - H_k \hat{x}_{k|k-1}) = -\lambda_k^{1/2} e_k, \end{aligned}$$

i.e., $\alpha^\top = -\lambda_k^{1/2} [R_{e,k}^\lambda]^{-1/2} e_k$. We introduce notation for the normalized residuals

$$\bar{e}_k = [R_{e,k}^\lambda]^{-1/2} e_k \text{ and } \bar{e}_k^\lambda = \lambda_k^{1/2} \bar{e}_k = \lambda_k^{1/2} [R_{e,k}^\lambda]^{-1/2} e_k.$$

Hence, we conclude that $\alpha^\top = -\bar{e}_k^\lambda$. Next, equation (25) of conventional one-step method is expressed in terms of available $[\bar{K}_{p,k}^\lambda]$ and the normalized residuals as follows:

$$\hat{x}_{k+1|k} = F_k \hat{x}_{k|k-1} + \lambda_k^{1/2} [\bar{K}_{p,k}^\lambda] \bar{e}_k = F_k \hat{x}_{k|k-1} + [\bar{K}_{p,k}^\lambda] \bar{e}_k^\lambda$$

where no matrix inversion is required for calculating the state, in contrast to the formula in line 6 of Algorithm 1.

It is worth noting here that the state vector can be calculated in alternative way. For that, we consider the inner product of the second and the last (block) rows of the pre-array in order to define β value

$$\begin{aligned}\bar{K}_{p,k}^\lambda \alpha^\top + P_{k+1|k}^{1/2} \beta^\top &= -\bar{K}_{p,k}^\lambda \bar{e}_k^\lambda + P_{k+1|k}^{1/2} \beta^\top \\ &= F_k P_{k|k-1}^{1/2} P_{k|k-1}^{-1/2} \hat{x}_{k|k-1} = F_k \hat{x}_{k|k-1},\end{aligned}$$

and, next, we conclude

$$\beta^\top = P_{k+1|k}^{-1/2} (F_k \hat{x}_{k|k-1} + \bar{K}_{p,k}^\lambda \bar{e}_k^\lambda) = P_{k+1|k}^{-1/2} \hat{x}_{k+1|k}.$$

Thus, an alternative state computation way is to read-off values $[P_{k+1|k}^{-1/2} \hat{x}_{k+1|k}]$ and $[P_{k+1|k}^{1/2}]$ from the extended post-array, and simply multiply these blocks to obtain

$$\hat{x}_{k+1|k} = [P_{k+1|k}^{1/2}] [P_{k+1|k}^{-1/2} \hat{x}_{k+1|k}].$$

In summary, the extended square-root condensed form implementation is presented in Algorithm 2.

Algorithm 2. ESR IMCC-KF (*a priori*, one-step, extended)

- INITIALIZATION: ($k = 0$)
- 1 Apply Cholesky factorization: $\Pi_0 = \Pi_0^{1/2} \Pi_0^{\top/2}$;
 - 2 Set initials: $P_{0|-1}^{1/2} = \Pi_0^{1/2}$, $P_{0|-1}^{-1/2} \hat{x}_{0|-1} = \Pi_0^{-1/2} \bar{x}_0$;
- FILTER RECURSION: ($k = 0, K$)
- 3 Compute λ_k by equation (10);
 - 4 Build and triangularize the first two (block) rows
- $$\underbrace{\begin{bmatrix} R_k^{1/2} & \lambda_k^{1/2} H_k P_{k|k-1}^{1/2} & 0 \\ 0 & F_k P_{k|k-1}^{1/2} & G_k Q_k^{1/2} \\ -\lambda_k^{1/2} y_k^\top R_k^{-\top/2} & [P_{k|k-1}^{-1/2} \hat{x}_{k|k-1}]^\top & 0 \end{bmatrix}}_{\text{Extended pre-array } \tilde{\mathbb{A}}_k} \mathbb{Q}_k$$
- $$= \underbrace{\begin{bmatrix} [R_{e,k}^\lambda]^{1/2} & 0 & 0 \\ \bar{K}_{p,k}^\lambda & P_{k+1|k}^{1/2} & 0 \\ -[\bar{e}_k^\lambda]^\top & [P_{k+1|k}^{-1/2} \hat{x}_{k+1|k}]^\top & (*) \end{bmatrix}}_{\text{Extended post-array } \tilde{\mathbb{B}}_k};$$
- 5 Read-off from post-array: $[P_{k+1|k}^{1/2}]$, $[P_{k+1|k}^{-1/2} \hat{x}_{k+1|k}]$;
 - 6 Compute $\hat{x}_{k+1|k} = [P_{k+1|k}^{1/2}] [P_{k+1|k}^{-1/2} \hat{x}_{k+1|k}]$.

Finally, we note that value γ is of no interest for the filtering algorithm above. Hence, we use notation $(*)$ for the block that is not used in the filter. However, it is worth noting here that this quantity turns to be utilized in the square-root Rauch-Tung-Striebel formulas for the smoothed estimate $\hat{x}_{k|K}$, if one derives the smoother under the MCC methodology. In this case, the suggested Algorithm 2 yields a simple computation of this value.

IV. NUMERICAL EXPERIMENTS

The goal of this section is to substantiate the theoretical derivation and correctness of the suggested condensed square-root Algorithms 1 and 2 on practical example.

Example 1 (see [6]). Consider a benchmark navigation problem where the vehicle dynamics is given as follows:

$$x_k = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} x_{k-1} + \begin{bmatrix} w_{k-1}^1 \\ w_{k-1}^2 \\ w_{k-1}^3 \\ w_{k-1}^4 \end{bmatrix},$$

where $T = 0.01$ is the sampling period, 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. A position-measuring device provides a noisy measurement of the north and east positions as follows:

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

The MCC-KF methods are tested at the presence of impulsive noise, i.e., in the presence of outliers. More precisely, the process and measurement noises are generated as follows:

$$\begin{aligned}w_k &\sim \mathcal{N}(0, Q) + \text{Shot noise}, & Q &= 0.1 I_4 \\ v_k &\sim \mathcal{N}(0, R) + \text{Shot noise}, & R &= 0.1 I_2\end{aligned}$$

where $\bar{x}_0 = [1, 1, 0, 0]^T$ and $\Pi_0 = \text{diag}([4, 4, 3, 3])$.

To simulate the impulsive noise (the shot noise), we follow the approach suggested in [6]. The Matlab routine `Shot_noise` recently published in [33, Appendix] can be used for generating the process and measurement noise in this set of experiments with 20% of outliers distributed randomly in the time interval $[21, K - 1]$ (where $K = 300$) and with the randomly chosen outliers' magnitude from $[0, 5]$.

In our numerical simulations, we test all square-root IMCC-KF implementations existed nowadays [9], [19]. As mentioned in Introduction, they are developed in two-step *a posteriori* form, only. The new one-step condensed Algorithms 1 and 2 are examined against *a priori* one-step IMCC-KF conventional equations (23) – (25). All filtering algorithms under examination utilize the same initial filtering conditions, the same measurements and the same noise covariances. The experiment is repeated for $M = 100$ Monte Carlo runs and the root mean square error (RMSE) is calculated for *a posteriori* and *a priori* estimates. The results of numerical experiments are summarized in Table I.

We observe that all examined implementation methods produce the same RMSEs both in *a priori* and *a posteriori* forms, respectively. In other words, they work with the same estimation accuracy. For *a priori* one-step methods derived in this paper, this means that the results of practical example substantiate the correctness of their theoretical derivation presented in Section III. This also justifies an algebraic equivalence of new square-root condensed algorithms and the conventional way of calculations by equations (23) – (25).

V. CONCLUSION

This paper proposes one-step filtering strategy for the maximum correntropy criterion Kalman filters that corresponds to *a priori* filtering. It has been shown that the resulted *a*

TABLE I
THE RMSE ERRORS FOR IMCC-KF IMPLEMENTATIONS IN EXAMPLE 1, $M = 100$ MONTE CARLO SIMULATIONS.

Estimator Implementation method	Covariance factorization	RMSE _{x_1}	RMSE _{x_2}	RMSE _{x_3}	RMSE _{x_4}	$\ RMSE\ _2$
The errors are computed for <i>a priori</i> (predicted) estimates						
IMCC-KF: conventional eqs. (23)–(25)	—	6.7524	6.3572	38.7909	38.5804	55.4905
SR IMCC-KF: new Algorithm 1	Cholesky	6.7524	6.3572	38.7909	38.5804	55.4905
eSR IMCC-KF: new Algorithm 2	Cholesky	6.7524	6.3572	38.7909	38.5804	55.4905
The errors are computed for <i>a posteriori</i> (filtered) estimates						
IMCC-KF: eqs. (11), (12), (14), (16)–(18)	—	6.6162	6.2099	38.7753	38.5790	55.4455
SR IMCC-KF: Algorithm 2 in [9]	Cholesky	6.6162	6.2099	38.7753	38.5790	55.4455
eSR IMCC-KF: Algorithm 3 in [9]	Cholesky	6.6162	6.2099	38.7753	38.5790	55.4455
UD IMCC-KF: Algorithm 2b in [19]	modified Cholesky	6.6162	6.2099	38.7753	38.5790	55.4455
SVD IMCC-KF: Algorithm 2c in [19]	SVD	6.6162	6.2099	38.7753	38.5790	55.4455

priori recursion for the original MCC-KF does not allow for robust square-root implementations because of the possibly negative coefficient involved. In contrast, the improved MCC-KF estimator (IMCC-KF) easily permits stable square-root implementations for its *a priori* recursion in the *condensed* form. Two Cholesky-based methods of this kind are derived in this paper. Finally, the square-root solution might be derived via UDU^T factorization as well. This problem is still open.

REFERENCES

- [1] W. Liu, P. P. Pokharel, and J. C. Principe, “Correntropy: properties and applications in non-Gaussian signal processing,” *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5286–5298, 2007.
- [2] B. Chen, L. Xing, J. Liang, N. Zheng, and J. C. Principe, “Steady-state mean-square error analysis for adaptive filtering under the maximum correntropy criterion,” *IEEE Signal Processing Letters*, vol. 21, no. 7, pp. 880–884, 2014.
- [3] B. Chen, J. Wang, H. Zhao, N. Zheng, and J. C. Principe, “Convergence of a fixed-point algorithm under maximum correntropy criterion,” *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1723–1727, 2015.
- [4] G. T. Cinar and J. C. Principe, “Adaptive background estimation using an information theoretic cost for hidden state estimation,” in *The 2011 International Joint Conference on Neural Networks (IJCNN)*, July 2011, pp. 489–494.
- [5] —, “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)*, June 2012, pp. 1–6.
- [6] R. Izanloo, S. A. Fakoorian, H. S. Yazdi, and D. Simon, “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)*, 2016, pp. 500–505.
- [7] B. Chen, X. Liu, H. Zhao, and J. C. Principe, “Maximum Correntropy Kalman Filter,” *Automatica*, vol. 76, pp. 70–77, 2017.
- [8] X. Liu, H. Qu, J. Zhao, and B. Chen, “State space maximum correntropy filter,” *Signal Processing*, vol. 130, pp. 152–158, 2017.
- [9] M. V. Kulikova, “Square-root algorithms for maximum correntropy estimation of linear discrete-time systems in presence of non-Gaussian noise,” *Systems & Control Letters*, vol. 108, pp. 8–15, 2017.
- [10] Y. Yang and G. Huang, “Map-based localization under adversarial attacks,” in *International Symposium on Robotics Research (ISRR)*, Puerto Varas, Chile, Dec. 2017.
- [11] X. Liu, H. Qu, J. Zhao, and B. Chen, “Extended Kalman filter under maximum correntropy criterion,” in *International Joint Conference on Neural Networks (IJCNN)*, 2016, pp. 1733–1737.
- [12] Y. Wang, W. Zheng, S. Sun, and L. Li, “Robust information filter based on maximum correntropy criterion,” *Journal of Guidance, Control, and Dynamics*, pp. 1126–1131, 2016.
- [13] X. Liu, B. Chen, B. Xu, Z. Wu, and P. Honeine, “Maximum correntropy unscented filter,” *International Journal of Systems Science*, vol. 48, no. 8, pp. 1607–1615, 2017.
- [14] G. Wang, N. Li, and Y. Zhang, “Maximum correntropy unscented Kalman and information filters for non-Gaussian measurement noise,” *Journal of the Franklin Institute*, vol. 354, no. 18, pp. 8659–8677, 2017.
- [15] W. Qin, X. Wang, and N. Cui, “Maximum correntropy sparse Gauss-Hermite quadrature filter and its application in tracking ballistic missile,” *IET Radar, Sonar & Navigation*, vol. 11, no. 9, pp. 1388–1396, 2017.
- [16] G. Yu. Kulikov and M. V. Kulikova, “Estimating the state in stiff continuous-time stochastic systems within extended Kalman filtering,” *SIAM Journal on Scientific Computing*, vol. 38, no. 6, pp. A3565–A3588, 2016.
- [17] —, “Accurate cubature and extended Kalman filtering methods for estimating continuous-time nonlinear stochastic systems with discrete measurements,” *Applied Numerical Mathematics*, vol. 111, pp. 260–275, 2017.
- [18] —, “Estimation of maneuvering target in the presence of non-Gaussian noise: A coordinated turn case study,” *Signal Process.*, vol. 145, pp. 241–257, 2018.
- [19] M. V. Kulikova, “Factored-form Kalman-like implementations under maximum correntropy criterion,” *Signal Processing*, vol. 160, pp. 328–338, 2019.
- [20] J. Potter and R. Stern, “Statistical filtering of space navigation measurements,” in *Guidance and Control Conference*. Reston, Virginia: American Institute of Aeronautics and Astronautics, Aug. 1963.
- [21] P. Park and T. Kailath, “New square-root algorithms for Kalman filtering,” *IEEE Transactions on Automatic Control*, vol. 40, no. 5, pp. 895–899, May 1995.
- [22] G. Yu. Kulikov and M. V. Kulikova, “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*, vol. 12, no. 16, pp. 2205–2212, 2018.
- [23] C. L. Thornton, “Triangular covariance factorizations for Kalman filtering,” Ph.D. dissertation, University of California at Los Angeles, 1976.
- [24] G. J. Bierman and C. L. Thornton, “Numerical comparison of Kalman filter algorithms: Orbit determination case study,” *Automatica*, vol. 13, no. 1, pp. 23–35, 1977.
- [25] M. V. Kulikova and J. V. Tsyganova, “Improved discrete-time Kalman filtering within singular value decomposition,” *IET Control Theory & Applications*, vol. 11, no. 15, pp. 2412–2418, 2017.
- [26] M. V. Kulikova, “Numerically robust SVD-based Kalman filter implementations,” in *Proceedings of the 22nd International Conference on System Theory, Control and Computing (ICSTCC)*, Oct. 2018, pp. 170–175.
- [27] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. New Jersey: Prentice Hall, 2000.
- [28] D. Simon, *Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches*. John Wiley & Sons, 2006.
- [29] D. Communiello and J. C. Principe, *Adaptive Learning Methods for Nonlinear System Modeling*, 1st ed. Elsevier, 2018.
- [30] B. Chen and J. C. Principe, “Maximum correntropy estimation is a smoothed MAP estimation,” *IEEE Signal Processing Letters*, vol. 19, no. 8, pp. 491–494, 2012.
- [31] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice using MATLAB*, 4th ed. New Jersey: John Wiley & Sons, 2015.
- [32] M. V. Kulikova, “Chandrasekhar-based maximum correntropy Kalman filtering with the adaptive kernel size selection,” *IEEE Transactions on Automatic Control*, in press., DOI: 10.1109/TAC.2019.2919341.
- [33] —, “Sequential maximum correntropy Kalman filtering,” *Asian Journal of Control*, vol. 21, no. 6, pp. 1–9, 2019.