Cholesky-Based Reduced-Rank Square-Root Kalman Filtering

J. Chandrasekar, I. S. Kim, D. S. Bernstein, and A. Ridley

***Abstract*— We consider a reduced-rank square-root Kalman filter based on the Cholesky decomposition of the state-error covariance. We compare the performance of this filter with the reduced-rank square-root filter based on the singular value decomposition.**

I. INTRODUCTION

The problem of state estimation for large-scale systems has gained increasing attention due to computationally intensive applications such as weather forecasting [1], where state estimation is commonly referred to as data assimilation. For these problems, there is a need for algorithms that are computationally tractable despite the enormous dimension of the state. These problems also typically entail nonlinear dynamics and model uncertainty [2], although these issues are outside the scope of this paper.

One approach to obtaining more tractable algorithms is to consider reduced-order Kalman filters. These reduced-complexity filters provide state estimates that are suboptimal relative to the classical Kalman filter [3–7]. Alternative reduced-order variants of the classical Kalman filter have been developed for computationally demanding applications [8–11], where the classical Kalman filter gain and covariance are modified so as to reduce the computational requirements. A comparison of several techniques is given in [12].

A widely studied technique for reducing the computational requirements of the Kalman filter for large scale systems is the *reduced-rank filter* [13–16]. In this method, the error-covariance matrix is factored to obtain a square root, whose rank is then reduced through truncation. This factorization-and-truncation method has direct application to the problem of generating a reduced ensemble for use in particle filter methods [17, 18].

Reduced-rank filters are closely related to the classical factorization techniques [19, 20], which provide numerical stability and computational efficiency, as well as a starting point for reduced-rank approximation.

The primary technique for truncating the error-covariance matrix is the singular value decomposition (SVD) [13–18], wherein the singular values provide guidance as to which components of the error covariance are most relevant to the accuracy of the state estimates. Approximation based on the SVD is largely motivated by the fact that error-covariance truncation is optimal with respect to approximation in unitarily invariant norms, such as the Frobenius norm. Despite this theoretical grounding, there appear to be no theoretical criteria to support the optimality of approximation based on the SVD within the context of recursive state estimation. The difficult is due to the fact that optimal approximation depends on the dynamics and measurement maps in addition to the components of the error covariance.

In the present paper we begin by observing that the Kalman filter update depends on the combination of terms C_kP_k, where C_k is the measurement map and P_k is the error covariance. This observation suggests that approximation of C_kP_k may be more suitable than approximation based on P_k alone.

To develop this idea, we show that approximation of C_kP_k leads directly to truncation based on the Cholesky decomposition. Unlike the SVD, however, the Cholesky decomposition does not possess a natural measure of magnitude that is analogous to the singular values arising in the SVD. Nevertheless, filter reduction based on the Cholesky decomposition provides state-estimation accuracy that is competitive with, and in many cases superior to, that of the SVD. In particular, we show that, in special cases, the accuracy of the Cholesky-decomposition-based reduced-rank filter is equal to the accuracy of the full-rank filter, and we demonstrate examples for which the Cholesky-decomposition-based reduced-rank filter provides acceptable accuracy, whereas the SVD-based reduced-rank filter provides arbitrarily poor accuracy.

A fortuitous advantage of using the Cholesky decomposition in place of the SVD is the fact that the Cholesky decomposition is computationally less expensive than the SVD, specifically, $O(n^3/6)$ [21], and thus an asymptotic computational advantage over SVD by a factor of 12. An additional advantage is that the entire matrix need not be factored; instead, by arranging the states so that those states that contribute directly to the measurement correspond to the initial columns of the lower triangular square root, then only the leading submatrix of the error covariance must be factored, yielding yet further savings over the SVD. Once the factorization is performed, the algorithm effectively retains only the initial "tall" columns of the full Cholesky factorization and truncates the "short" columns.

II. THE KALMAN FILTER

Consider the time-varying discrete-time system

$$x_{k+1} = A_kx_k + G_kw_k, \tag{2.1}$$

$$y_k = C_kx_k + H_kv_k, \tag{2.2}$$

This research was supported by the National Science Foundation, under grants CNS-0539053 and ATM-0325332 to the University of Michigan, Ann Arbor, USA.

J. Chandrasekar, I. S. Kim, and D. S. Bernstein are with Department of Aerospace Engineering, and A. Ridley is with Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, Ann Arbor, MI, dsbaero@umich.edu

where $x_k \in \mathbb{R}^{n_k}$, $w_k \in \mathbb{R}^{d_w}$, $y_k \in \mathbb{R}^{p_k}$, $v_k \in \mathbb{R}^{d_v}$, and A_k, G_k, C_k, and H_k are known real matrices of appropriate sizes. We assume that w_k and v_k are zero-mean white processes with unit covariances. Define $Q_k \triangleq G_k G_k^{\mathrm{T}}$ and $R_k \triangleq H_k H_k^{\mathrm{T}}$, and assume that R_k is positive definite for all $k \geqslant 0$. Furthermore, we assume that w_k and v_k are uncorrelated for all $k \geqslant 0$. The objective is to obtain an estimate of the state x_k using the measurements y_k.

The Kalman filter provides the optimal minimum-variance estimate of the state x_k. The Kalman filter can be expressed in two steps, namely, the *data assimilation step*, where the measurements are used to update the states, and the *forecast step*, which uses the model. These steps can be summarized as follows:

Data Assimilation Step

$$K_k = P_k^{\mathrm{f}} C_k^{\mathrm{T}} (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1}, \tag{2.3}$$

$$P_k^{\mathrm{da}} = P_k^{\mathrm{f}} - P_k^{\mathrm{f}} C_k^{\mathrm{T}} (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1} C_k P_k^{\mathrm{f}}, \tag{2.4}$$

$$x_k^{\mathrm{da}} = x_k^{\mathrm{f}} + K_k (y_k - C_k x_k^{\mathrm{f}}). \tag{2.5}$$

Forecast Step

$$x_{k+1}^{\mathrm{f}} = A_k x_k^{\mathrm{da}}, \tag{2.6}$$

$$P_{k+1}^{\mathrm{f}} = A_k P_k^{\mathrm{da}} A_k^{\mathrm{T}} + Q_k. \tag{2.7}$$

The matrices $P_k^{\mathrm{f}} \in \mathbb{R}^{n\times n}$ and $P_k^{\mathrm{da}} \in \mathbb{R}^{n\times n}$ are the state-error covariances, that is,

$$P_k^{\mathrm{f}} = \mathcal{E}[e_k^{\mathrm{f}} (e_k^{\mathrm{f}})^{\mathrm{T}}], \quad P_k^{\mathrm{da}} = \mathcal{E}[e_k^{\mathrm{da}} (e_k^{\mathrm{da}})^{\mathrm{T}}], \tag{2.8}$$

where

$$e_k^{\mathrm{f}} \triangleq x_k - x_k^{\mathrm{f}}, \quad e_k^{\mathrm{da}} \triangleq x_k - x_k^{\mathrm{da}}. \tag{2.9}$$

In the following sections, we consider reduced-rank square-root filters that propagate approximations of a square-root of the error covariance instead of the actual error covariance.

III. SVD-Based Reduced-Rank Square-Root Filter

Note that the Kalman filter uses the error covariances P_k^{da} and P_k^{f}, which are updated using (2.4) and (2.7). For computational efficiency, we construct a suboptimal filter that uses reduced-rank approximations of the error covariances P_k^{da} and P_k^{f}. Specifically, we consider reduced-rank approximations $\hat{P}_k^{\mathrm{da}}$ and $\hat{P}_k^{\mathrm{f}}$ of the error covariances P_k^{da} and P_k^{f} such that $\|P_k^{\mathrm{da}} - \hat{P}_k^{\mathrm{da}}\|_{\mathrm{F}}$ and $\|P_k^{\mathrm{f}} - \hat{P}_k^{\mathrm{f}}\|_{\mathrm{F}}$ are minimized, where $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm. To achieve this approximation, we compute singular value decompositions of the error covariances at each time step.

Let $P \in \mathbb{R}^{n\times n}$ be positive semidefinite, let $\sigma_1 \geqslant \cdots \geqslant \sigma_n$ be the singular values of P, and let $u_1, \ldots, u_n \in \mathbb{R}^n$ be corresponding orthonormal eigenvectors. Next, define $U_q \in \mathbb{R}^{n\times q}$ and $\Sigma_q \in \mathbb{R}^{q\times q}$ by

$$U_q \triangleq \begin{bmatrix} u_1 & \cdots & u_q \end{bmatrix}, \quad \Sigma_q \triangleq \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_q \end{bmatrix}. \tag{3.1}$$

With this notation, the singular value decomposition of P is given by

$$P = U_n \Sigma_n U_n^{\mathrm{T}}, \tag{3.2}$$

where U_n is orthogonal. For $q \leqslant n$, let $\Phi_{\mathrm{SVD}}(P,q) \in \mathbb{R}^{n\times q}$ denote the SVD-based rank-q approximation of the square root $\Sigma_q^{1/2}$ of P given by

$$\Phi_{\mathrm{SVD}}(P,q) \triangleq U_q \Sigma_q^{1/2}. \tag{3.3}$$

The following standard result shows that SS^{T}, where $S \triangleq \Phi_{\mathrm{SVD}}(P,q)$, is the best rank-$q$ approximation of P in the Frobenius norm.

Lemma III.1. *Let $P \in \mathbb{R}^{n\times n}$ be positive semidefinite, and let $\sigma_1 \cdots \geqslant \sigma_n$ be the singular values of P. If $S = \Phi_{\mathrm{SVD}}(P,q)$, then*

$$\min_{\mathrm{rank}(\hat{P})=q} \|P - \hat{P}\|_{\mathrm{F}}^2 = \|P - SS^{\mathrm{T}}\|_{\mathrm{F}}^2 = \sigma_{q+1}^2 + \cdots + \sigma_n^2. \tag{3.4}$$

The data assimilation and forecast steps of the SVD-based rank-q square-root filter are given by the following steps:

Data Assimilation step

$$K_{\mathrm{s},k} = \hat{P}_{\mathrm{s},k}^{\mathrm{f}} C_k^{\mathrm{T}} \left(C_k \hat{P}_{\mathrm{s},k}^{\mathrm{f}} C_k^{\mathrm{T}} + R_k \right)^{-1}, \tag{3.5}$$

$$\tilde{P}_{\mathrm{s},k}^{\mathrm{da}} = \hat{P}_k^{\mathrm{f}} - \hat{P}_k^{\mathrm{f}} C_k^{\mathrm{T}} \left(C_k \hat{P}_{\mathrm{s},k}^{\mathrm{f}} C_k^{\mathrm{T}} + R_k \right)^{-1} C_k \hat{P}_{\mathrm{s},k}^{\mathrm{f}}, \tag{3.6}$$

$$x_{\mathrm{s},k}^{\mathrm{da}} = x_{\mathrm{s},k}^{\mathrm{f}} + K_{\mathrm{s},k} (y_k - C_k x_{\mathrm{s},k}^{\mathrm{f}}), \tag{3.7}$$

where

$$\tilde{S}_{\mathrm{s},k}^{\mathrm{f}} \triangleq \Phi_{\mathrm{SVD}}(\tilde{P}_{\mathrm{s},k}^{\mathrm{f}}, q), \tag{3.8}$$

$$\hat{P}_{\mathrm{s},k}^{\mathrm{f}} \triangleq \tilde{S}_{\mathrm{s},k}^{\mathrm{f}} (\tilde{S}_{\mathrm{s},k}^{\mathrm{f}})^{\mathrm{T}}. \tag{3.9}$$

Forecast step

$$x_{\mathrm{s},k+1}^{\mathrm{f}} = A_k x_{\mathrm{s},k}^{\mathrm{da}}, \tag{3.10}$$

$$\tilde{P}_{\mathrm{s},k+1}^{\mathrm{f}} = A_k \hat{P}_{\mathrm{s},k}^{\mathrm{da}} A_k^{\mathrm{T}} + Q_k, \tag{3.11}$$

where

$$\tilde{S}_{\mathrm{s},k}^{\mathrm{da}} \triangleq \Phi_{\mathrm{SVD}}(\tilde{P}_{\mathrm{s},k}^{\mathrm{da}}, q), \tag{3.12}$$

$$\hat{P}_{\mathrm{s},k}^{\mathrm{da}} \triangleq \tilde{S}_{\mathrm{s},k}^{\mathrm{da}} (\tilde{S}_{\mathrm{s},k}^{\mathrm{da}})^{\mathrm{T}}, \tag{3.13}$$

and $\tilde{P}_{\mathrm{s},0}^{\mathrm{f}}$ is positive semidefinite.

Next, define the forecast and data assimilation error covariances $P_{\mathrm{s},k}^{\mathrm{f}}$ and $P_{\mathrm{s},k}^{\mathrm{da}}$ of the SVD-based rank-q square-root filter by

$$P_{\mathrm{s},k}^{\mathrm{f}} \triangleq \mathcal{E}\left[(x_k - x_{\mathrm{s},k}^{\mathrm{f}})(x_k - x_{\mathrm{s},k}^{\mathrm{f}})^{\mathrm{T}} \right], \tag{3.14}$$

$$P_{\mathrm{s},k}^{\mathrm{da}} \triangleq \mathcal{E}\left[(x_k - x_{\mathrm{s},k}^{\mathrm{da}})(x_k - x_{\mathrm{s},k}^{\mathrm{da}})^{\mathrm{T}} \right]. \tag{3.15}$$

Using (2.1), (3.7) and (3.10), it can be shown that

$$P_{\mathrm{s},k}^{\mathrm{da}} = (I - K_{\mathrm{s},k} C) P_{\mathrm{s},k}^{\mathrm{f}} (I - K_{\mathrm{s},k} C)^{\mathrm{T}} + K_{\mathrm{s},k} R_k K_{\mathrm{s},k}^{\mathrm{T}}, \tag{3.16}$$

$$P_{\mathrm{s},k+1}^{\mathrm{f}} = A_k P_{\mathrm{s},k}^{\mathrm{da}} A_k^{\mathrm{T}} + Q_k. \tag{3.17}$$

Note that $\tilde{S}_{\mathrm{s},k}^{\mathrm{f}} \left(\tilde{S}_{\mathrm{s},k}^{\mathrm{f}} \right)^{\mathrm{T}} \leqslant \tilde{P}_{\mathrm{s},k}^{\mathrm{f}}$ and $\tilde{S}_{\mathrm{s},k}^{\mathrm{da}} \left(\tilde{S}_{\mathrm{s},k}^{\mathrm{da}} \right)^{\mathrm{T}} \leqslant \tilde{P}_{\mathrm{s},k}^{\mathrm{da}}$. Hence, even if $\tilde{P}_{\mathrm{s},0}^{\mathrm{f}} = P_{\mathrm{s},0}^{\mathrm{f}}$, it does not necesarily follow that $\tilde{P}_{\mathrm{s},k}^{\mathrm{f}} = P_{\mathrm{s},k}^{\mathrm{f}}$ and $\tilde{P}_{\mathrm{s},k}^{\mathrm{da}} = P_{\mathrm{s},k}^{\mathrm{da}}$ for all $k > 0$. Therefore, since $K_{\mathrm{s},k}$ does not use the true error covariance $P_{\mathrm{s},k}^{\mathrm{f}}$, the SVD-based rank-$q$ square-root filter is generally not equivalent to the Kalman filter. However, under certain conditions, the SVD-based rank-q square-root filter is equivalent to the Kalman filter. Specifically, we have the following result.

Proposition III.1. *Assume that $\tilde{P}^{\rm f}_{{\rm s},k} = P^{\rm f}_{{\rm s},k}$ and $\operatorname{rank}(P^{\rm f}_{{\rm s},k}) \leqslant q$. Then, $\tilde{P}^{\rm da}_{{\rm s},k} = P^{\rm da}_{{\rm s},k}$ and $\tilde{P}^{\rm f}_{{\rm s},k+1} = P^{\rm f}_{{\rm s},k+1}$. If, in addition, $P^{\rm f}_{{\rm s},k} = P^{\rm f}_k$, then $K_{{\rm s},k} = K_k$ and $P^{\rm f}_{{\rm s},k+1} = P^{\rm f}_{k+1}$.*

Proof. Since $\operatorname{rank}(\tilde{P}^{\rm f}_k) \leqslant q$, it follows from Lemma III.1 that

$$\hat{P}^{\rm f}_{{\rm s},k} = \tilde{S}^{\rm f}_{{\rm s},k}\left(\tilde{S}^{\rm f}_{{\rm s},k}\right)^{\rm T} = \tilde{P}^{\rm f}_{{\rm s},k}. \tag{3.18}$$

Hence, it follows from (3.5) that

$$K_{{\rm s},k} = P^{\rm f}_{{\rm s},k} C_k^{\rm T}\left(C_k P^{\rm f}_{{\rm s},k} C_k^{\rm T} + R_k\right)^{-1}, \tag{3.19}$$

while substituting (3.18) into (3.6) yields

$$\tilde{P}^{\rm da}_{{\rm s},k} = \tilde{P}^{\rm f}_{{\rm s},k} - \tilde{P}^{\rm f}_{{\rm s},k} C_k^{\rm T}(C_k \tilde{P}^{\rm f}_{{\rm s},k} C_k^{\rm T} + R_k)^{-1} C_k \tilde{P}^{\rm f}_{{\rm s},k}. \tag{3.20}$$

Next, substituting (3.19) into (3.16) and using $\tilde{P}^{\rm f}_{{\rm s},k} = P^{\rm f}_{{\rm s},k}$ in the resulting expression yields

$$\tilde{P}^{\rm da}_{{\rm s},k} = P^{\rm da}_{{\rm s},k}. \tag{3.21}$$

Since $\operatorname{rank}(\tilde{P}^{\rm f}_{{\rm s},k}) \leqslant q$, it follows from (3.20) that $\operatorname{rank}(\tilde{P}^{\rm da}_{{\rm s},k}) \leqslant q$, and therefore Lemma III.1 implies that

$$\hat{P}^{\rm da}_{{\rm s},k} = \tilde{S}^{\rm da}_{{\rm s},k}\left(\tilde{S}^{\rm da}_{{\rm s},k}\right)^{\rm T} = \tilde{P}^{\rm da}_{{\rm s},k}. \tag{3.22}$$

Hence, it follows from (3.11), (3.17), and (3.21) that

$$\tilde{P}^{\rm f}_{{\rm s},k+1} = P^{\rm f}_{{\rm s},k+1}. \tag{3.23}$$

Finally, it follows from (2.3) and (3.19) that

$$K_{{\rm s},k} = K_k. \tag{3.24}$$

Therefore, (2.4) and (3.16) imply that

$$P^{\rm da}_{{\rm s},k} = P^{\rm da}_k. \tag{3.25}$$

Hence, it follows from (2.7), (3.17) and (3.25) that

$$P^{\rm f}_{{\rm s},k+1} = P^{\rm f}_{k+1}.$$ □

Corollary III.1. *Assume that $x^{\rm f}_{{\rm s},0} = x^{\rm f}_0$, $\tilde{P}^{\rm f}_{{\rm s},0} = P^{\rm f}_{{\rm s},0}$, and $\operatorname{rank}(P^{\rm f}_0) \leqslant q$. Furthermore, assume that, for all $k \geqslant 0$, $\operatorname{rank}(A_k) + \operatorname{rank}(Q_k) \leqslant q$. Then, for all $k \geqslant 0$, $K_{{\rm s},k} = K_k$ and $x^{\rm f}_{{\rm s},k} = x^{\rm f}_k$.*

Proof. Since $x^{\rm f}_{{\rm s},0} = x^{\rm f}_0$, (2.8) and (3.14) imply that $P^{\rm f}_{{\rm s},0} = P^{\rm f}_0$. It follows from (3.16) that if $\operatorname{rank}(P^{\rm f}_{{\rm s},k}) \leqslant q$, then $\operatorname{rank}(P^{\rm da}_{{\rm s},k}) \leqslant q$, and hence (3.17) implies that $\operatorname{rank}(P^{\rm f}_{{\rm s},k+1}) \leqslant q$. Therefore, using Proposition III.1 and induction, it follows that $K_{{\rm s},k} = K_k$ for all $k \geqslant 0$. Therefore, (2.5), (2.6), (3.7) and (3.10) imply that $x^{\rm f}_{{\rm s},k} = x^{\rm f}_k$ for all $k \geqslant 0$. □

IV. Cholesky-Decomposition-Based Reduced-Rank Square-Root Filter

The Kalman filter gain K_k depends on a particular subspace of the range of the error covariance. Specifically, K_k depends only on $C_k P^{\rm f}_k$ and not on the entire error covariance. We thus consider a filter that uses reduced-rank approximations $\hat{P}^{\rm da}_k$ and $\hat{P}^{\rm f}_k$ of the error covariances $P^{\rm da}_k$ and $P^{\rm f}_k$ such that $\|C_k(P^{\rm da}_k - \hat{P}^{\rm da}_k)\|_{\rm F}$ and $\|C_k(P^{\rm f}_k - \hat{P}^{\rm f}_k)\|_{\rm F}$ are minimized. To achieve this minimization, we compute a Cholesky decomposition of both error covariances at each time step.

Since $P \in \mathbb{R}^{n\times n}$ is positive semidefinite, the Cholesky decomposition yields a lower triangular Cholesky factor $L \in \mathbb{R}^{n\times n}$ of P that satisfies

$$LL^{\rm T} = P. \tag{4.1}$$

Partition L as

$$L = \begin{bmatrix} L_1 & \cdots & L_n \end{bmatrix}, \tag{4.2}$$

where, for $i = 1, \ldots, n$, $L_i \in \mathbb{R}^n$ has real entries

$$L_i = \begin{bmatrix} 0_{1\times(i-1)} & L_{i,1} & \cdots & L_{i,n-i+1} \end{bmatrix}^{\rm T}. \tag{4.3}$$

Truncating the last $n-q$ columns of L yields the reduced-rank Cholesky factor

$$\Phi_{\rm CHOL}(P,q) \triangleq \begin{bmatrix} L_1 & \cdots & L_q \end{bmatrix} \in \mathbb{R}^{n\times q}. \tag{4.4}$$

Lemma IV.1. *Let $P \in \mathbb{R}^{n\times n}$ be positive definite, define $S \triangleq \Phi_{\rm CHOL}(P,q)$ and $\hat{P} \triangleq SS^{\rm T}$, and partition P and $\hat{P}$ as*

$$P = \begin{bmatrix} P_1 & P_{12} \\ (P_{12})^{\rm T} & P_2 \end{bmatrix},\quad \hat{P} = \begin{bmatrix} \hat{P}_1 & \hat{P}_{12} \\ (\hat{P}_{12}) & \hat{P}_2 \end{bmatrix}, \tag{4.5}$$

where $P_1, \hat{P}_1 \in \mathbb{R}^{q\times q}$ and $P_2, \hat{P}_2 \in \mathbb{R}^{r\times r}$. Then, $\hat{P}_1 = P_1$ and $\hat{P}_{12} = P_{12}$.

Proof. Let L be the Cholesky factor of P. It follows from (4.3) that $L_i L_i^{\rm T} \in \mathbb{R}^n$ has the structure

$$L_i L_i^{\rm T} = \begin{bmatrix} 0_{i-1} & 0_{(i-1)\times(n-i+1)} \\ 0_{(n-i+1)\times(i-1)} & X_i \end{bmatrix}, \tag{4.6}$$

where $X_i \in \mathbb{R}^{(n-i+1)\times(n-i+1)}$. Therefore,

$$\sum_{i=q+1}^{n} L_i L_i^{\rm T} = \begin{bmatrix} 0_{q\times q} & 0_{q\times r} \\ 0_{r\times q} & Y_r \end{bmatrix}, \tag{4.7}$$

where $Y_r \in \mathbb{R}^{r\times r}$. Since

$$P = \sum_{i=1}^{n} L_i L_i^{\rm T}, \tag{4.8}$$

it follows from (4.4) that

$$P = \hat{P} + \sum_{i=q+1}^{n} L_i L_i^{\rm T}. \tag{4.9}$$

Substituting (4.7) into (4.9) yields $\hat{P}_1 = P_1$ and $\hat{P}_{12} = P_{12}$. □

Lemma IV.1 implies that, if $S = \Phi_{\rm CHOL}(P,q)$, then the first q rows and columns of $SS^{\rm T}$ and P are equal.

The data assimilation and forecast steps of the Cholesky-based rank-q square-root filter are given by the following steps:

Data Assimilation step

$$K_{{\rm c},k} = \hat{P}^{\rm f}_{{\rm c},k} C_k^{\rm T}\left(C_k \hat{P}^{\rm f}_{{\rm c},k} C_k^{\rm T} + R_k\right)^{-1}, \tag{4.10}$$

$$\tilde{P}^{\rm da}_{{\rm c},k} = \hat{P}^{\rm f}_{{\rm c},k} - \hat{P}^{\rm f}_{{\rm c},k} C_k^{\rm T}\left(C_k \hat{P}^{\rm f}_{{\rm c},k} C_k^{\rm T} + R_k\right)^{-1} C_k \hat{P}^{\rm f}_{{\rm c},k}, \tag{4.11}$$

$$x^{\rm da}_{{\rm c},k} = x^{\rm f}_{{\rm c},k} + K_{{\rm c},k}(y_k - C_k x^{\rm f}_{{\rm c},k}), \tag{4.12}$$

where

$$\tilde{S}^{\rm f}_{{\rm c},k} \triangleq \Phi_{\rm SVD}(\tilde{P}^{\rm f}_{{\rm c},k}, q), \tag{4.13}$$

$$\hat{P}^{\rm f}_{{\rm c},k} \triangleq \tilde{S}^{\rm f}_{{\rm c},k}(\tilde{S}^{\rm f}_{{\rm s},k})^{\rm T}. \tag{4.14}$$

Forecast step

$$x^{\mathrm{f}}_{\mathrm{c},k+1} = A_k x^{\mathrm{da}}_{\mathrm{c},k}, \tag{4.15}$$

$$\tilde{P}^{\mathrm{f}}_{\mathrm{c},k+1} = A_k \hat{P}^{\mathrm{da}}_{\mathrm{c},k} A_k^{\mathrm{T}} + Q_k, \tag{4.16}$$

where

$$\tilde{S}^{\mathrm{da}}_{\mathrm{c},k} \triangleq \Phi_{\mathrm{SVD}}(\tilde{P}^{\mathrm{da}}_{\mathrm{c},k}, q), \tag{4.17}$$

$$\hat{P}^{\mathrm{da}}_{\mathrm{c},k} \triangleq \tilde{S}^{\mathrm{da}}_{\mathrm{c},k}(\tilde{S}^{\mathrm{da}}_{\mathrm{c},k})^{\mathrm{T}}, \tag{4.18}$$

and $\tilde{P}^{\mathrm{f}}_{\mathrm{c},0}$ is positive semidefinite.

Next, define the forecast and data assimilation error covariances $P^{\mathrm{f}}_{\mathrm{c},k}$ and $P^{\mathrm{da}}_{\mathrm{c},k}$, respectively, of the Cholesky-based rank-q square-root filter by

$$P^{\mathrm{f}}_{\mathrm{c},k} \triangleq \mathcal{E}\left[(x_k - x^{\mathrm{f}}_{\mathrm{c},k})(x_k - x^{\mathrm{f}}_{\mathrm{c},k})^{\mathrm{T}}\right], \tag{4.19}$$

$$P^{\mathrm{da}}_{\mathrm{s},k} \triangleq \mathcal{E}\left[(x_k - x^{\mathrm{da}}_{\mathrm{c},k})(x_k - x^{\mathrm{da}}_{\mathrm{c},k})^{\mathrm{T}}\right], \tag{4.20}$$

that is, $P^{\mathrm{f}}_{\mathrm{c},k}$ and $P^{\mathrm{da}}_{\mathrm{c},k}$ are the error covariances when the Cholesky-based rank-q square-root filter is used. Using (2.1), (4.12) and (4.15), it can be shown that

$$P^{\mathrm{da}}_{\mathrm{c},k} = (I - K_{\mathrm{c},k}C)P^{\mathrm{f}}_{\mathrm{c},k}(I - K_{\mathrm{c},k}C)^{\mathrm{T}} + K_{\mathrm{c},k}R_k K^{\mathrm{T}}_{\mathrm{c},k}, \tag{4.21}$$

$$P^{\mathrm{f}}_{\mathrm{c},k} = A_k P^{\mathrm{da}}_{\mathrm{c},k} A_k^{\mathrm{T}} + Q_k. \tag{4.22}$$

Note that $\tilde{S}^{\mathrm{f}}_{\mathrm{c},k}\left(\tilde{S}^{\mathrm{f}}_{\mathrm{c},k}\right)^{\mathrm{T}} \leqslant \tilde{P}^{\mathrm{f}}_{k}$ and $\tilde{S}^{\mathrm{da}}_{\mathrm{c},k}\left(\tilde{S}^{\mathrm{da}}_{\mathrm{c},k}\right)^{\mathrm{T}} \leqslant \tilde{P}^{\mathrm{da}}_{k}$. Hence, even if $\tilde{P}^{\mathrm{f}}_{\mathrm{c},0} = P^{\mathrm{f}}_{\mathrm{c},0}$, the Cholesky-based rank-q square-root filter is generally not equivalent to the Kalman filter. However, in certain cases, the Cholesky-based rank-q square root filter is equivalent to the Kalman filter.

Proposition IV.1. *Let A_k and C_k have the structure*

$$A_k = \begin{bmatrix} A_{1,k} & 0 \\ A_{21,k} & A_{2,k} \end{bmatrix}, \quad C_k = \begin{bmatrix} I_p & 0 \end{bmatrix}, \tag{4.23}$$

where $A_{1,k} \in \mathbb{R}^{p\times p}$ and $A_{2,k} \in \mathbb{R}^{r\times r}$, partition P^{f}_k and $\tilde{P}^{\mathrm{f}}_{\mathrm{c},k}$ as

$$P^{\mathrm{f}}_k = \begin{bmatrix} P^{\mathrm{f}}_{1,k} & (P^{\mathrm{f}}_{21,k})^{\mathrm{T}} \\ P^{\mathrm{f}}_{21,k} & P^{\mathrm{f}}_{2,k} \end{bmatrix}, \quad \tilde{P}^{\mathrm{f}}_{\mathrm{c},k} = \begin{bmatrix} \tilde{P}^{\mathrm{f}}_{\mathrm{c},1,k} & (\tilde{P}^{\mathrm{f}}_{\mathrm{c},21,k})^{\mathrm{T}} \\ \tilde{P}^{\mathrm{f}}_{\mathrm{c},21,k} & \tilde{P}^{\mathrm{f}}_{\mathrm{c},2,k} \end{bmatrix}, \tag{4.24}$$

where $P^{\mathrm{f}}_{1,k}, \tilde{P}^{\mathrm{f}}_{\mathrm{c},1,k} \in \mathbb{R}^{p\times p}$ and $P^{\mathrm{f}}_{2,k}, \tilde{P}^{\mathrm{f}}_{\mathrm{c},2,k} \in \mathbb{R}^{r\times r}$, and assume that $q = p$, $\tilde{P}^{\mathrm{f}}_{\mathrm{c},1,k} = P^{\mathrm{f}}_{1,k}$, and $\tilde{P}^{\mathrm{f}}_{\mathrm{c},21,k} = P^{\mathrm{f}}_{21,k}$. Then, $K_{\mathrm{c},k} = K_k$, $\tilde{P}^{\mathrm{f}}_{\mathrm{c},1,k+1} = P^{\mathrm{f}}_{1,k+1}$, and $\tilde{P}^{\mathrm{f}}_{\mathrm{c},21,k+1} = P^{\mathrm{f}}_{21,k+1}$.

Proof. Let P^{da}_k have entries

$$P^{\mathrm{da}}_k = \begin{bmatrix} P^{\mathrm{da}}_{1,k} & (P^{\mathrm{da}}_{21,k})^{\mathrm{T}} \\ P^{\mathrm{da}}_{21,k} & P^{\mathrm{da}}_{2,k} \end{bmatrix}, \tag{4.25}$$

where $P^{\mathrm{da}}_{1,k} \in \mathbb{R}^{p\times p}$ is positive semidefinite and $P^{\mathrm{da}}_{2,k} \in \mathbb{R}^{r\times r}$. It follows from (2.4) that

$$P^{\mathrm{da}}_{1,k} = P^{\mathrm{f}}_{1,k} - P^{\mathrm{f}}_{1,k}(P^{\mathrm{f}}_{1,k} + R_k)^{-1}P^{\mathrm{f}}_{1,k}, \tag{4.26}$$

$$P^{\mathrm{da}}_{21,k} = P^{\mathrm{f}}_{21,k} - P^{\mathrm{f}}_{21,k}(P^{\mathrm{f}}_{1,k} + R_k)^{-1}P^{\mathrm{f}}_{1,k}. \tag{4.27}$$

Substituting (4.23) into (2.3) yields

$$K_k = \begin{bmatrix} P^{\mathrm{f}}_{1,k} \\ P^{\mathrm{f}}_{21,k} \end{bmatrix}(P^{\mathrm{f}}_{1,k} + R_k)^{-1}. \tag{4.28}$$

Furthermore, (2.7) and (4.23) imply that

$$P^{\mathrm{f}}_{1,k+1} = A_{1,k}P^{\mathrm{da}}_{1,k}A^{\mathrm{T}}_{1,k} + Q_{1,k}, \tag{4.29}$$

$$P^{\mathrm{f}}_{21,k+1} = A_{2,k}P^{\mathrm{da}}_{21,k}A^{\mathrm{T}}_{1,k} + A_{21,k}P^{\mathrm{da}}_{1,k}A^{\mathrm{T}}_{1,k} + Q_{21,k}, \tag{4.30}$$

where Q_k has entries

$$Q_k = \begin{bmatrix} Q_{1,k} & (Q_{21,k})^{\mathrm{T}} \\ Q_{21,k} & Q_{2,k} \end{bmatrix}. \tag{4.31}$$

Define $\hat{P}^{\mathrm{f}}_{\mathrm{c},k}$ and $\hat{P}^{\mathrm{da}}_{\mathrm{c},k}$ by (4.14) and (4.18), and let $\hat{P}^{\mathrm{f}}_{\mathrm{c},k}$ and $\hat{P}^{\mathrm{da}}_{\mathrm{c},k}$ have entries

$$\hat{P}^{\mathrm{f}}_{\mathrm{c},k} = \begin{bmatrix} \hat{P}^{\mathrm{f}}_{\mathrm{c},1,k} & (\hat{P}^{\mathrm{f}}_{\mathrm{c},21,k})^{\mathrm{T}} \\ \hat{P}^{\mathrm{f}}_{\mathrm{c},21,k} & \hat{P}^{\mathrm{f}}_{\mathrm{c},2,k} \end{bmatrix}, \hat{P}^{\mathrm{da}}_{\mathrm{c},k} = \begin{bmatrix} \hat{P}^{\mathrm{da}}_{\mathrm{c},1,k} & (\hat{P}^{\mathrm{da}}_{\mathrm{c},21,k})^{\mathrm{T}} \\ \hat{P}^{\mathrm{da}}_{\mathrm{c},21,k} & \hat{P}^{\mathrm{da}}_{\mathrm{c},2,k} \end{bmatrix}, \tag{4.32}$$

where $\hat{P}^{\mathrm{da}}_{1,k}, \hat{P}^{\mathrm{f}}_{1,k} \in \mathbb{R}^{p\times p}$ are positive semidefinite and $\hat{P}^{\mathrm{da}}_{2,k}, \hat{P}^{\mathrm{f}}_{2,k} \in \mathbb{R}^{r\times r}$. Substituting (4.32) into (4.10) yields

$$K_{\mathrm{c},k} = \begin{bmatrix} \hat{P}^{\mathrm{f}}_{1,k} \\ \hat{P}^{\mathrm{f}}_{21,k} \end{bmatrix}(\hat{P}^{\mathrm{f}}_{1,k} + R_k)^{-1}. \tag{4.33}$$

Since $S^{\mathrm{f}}_{\mathrm{c},k} = \Phi_{\mathrm{CHOL}}(\tilde{P}^{\mathrm{f}}_{\mathrm{c},k}, q)$ and we assume that $q = p$, $\tilde{P}^{\mathrm{f}}_{\mathrm{c},1,k} = P^{\mathrm{f}}_{1,k}$, and $\tilde{P}^{\mathrm{f}}_{\mathrm{c},21,k} = P^{\mathrm{f}}_{21,k}$, it follows from Lemma IV.1 that

$$\hat{P}^{\mathrm{f}}_{\mathrm{c},1,k} = P^{\mathrm{f}}_{\mathrm{c},1,k}, \quad \hat{P}^{\mathrm{f}}_{\mathrm{c},21,k} = P^{\mathrm{f}}_{\mathrm{c},21,k}. \tag{4.34}$$

Hence, $K_{\mathrm{c},k} = K_k$.

Next, substituting (4.14) into (4.11) yields

$$\tilde{P}^{\mathrm{da}}_{\mathrm{c},k} = \hat{P}^{\mathrm{f}}_{\mathrm{c},k} - \hat{P}^{\mathrm{f}}_{\mathrm{c},k}C_k^{\mathrm{T}}(C_k\hat{P}^{\mathrm{f}}_{\mathrm{c},k}C_k^{\mathrm{T}} + R_k)^{-1}C_k\hat{P}^{\mathrm{f}}_{\mathrm{c},k}. \tag{4.35}$$

Let $\tilde{P}^{\mathrm{da}}_{\mathrm{c},k}$ have entries

$$\tilde{P}^{\mathrm{da}}_{\mathrm{c},k} = \begin{bmatrix} \tilde{P}^{\mathrm{da}}_{\mathrm{c},1,k} & (\tilde{P}^{\mathrm{da}}_{\mathrm{c},21,k})^{\mathrm{T}} \\ \tilde{P}^{\mathrm{da}}_{\mathrm{c},21,k} & \tilde{P}^{\mathrm{da}}_{\mathrm{c},2,k} \end{bmatrix}, \tag{4.36}$$

where $\tilde{P}^{\mathrm{da}}_{\mathrm{c},1,k} \in \mathbb{R}^{p\times p}$ is positive semidefinite and $\tilde{P}^{\mathrm{da}}_{\mathrm{c},2,k} \in \mathbb{R}^{r\times r}$. Therefore, it follows from (4.23) and (4.32) that

$$\tilde{P}^{\mathrm{da}}_{\mathrm{c},1,k} = \hat{P}^{\mathrm{f}}_{\mathrm{c},1,k} - \hat{P}^{\mathrm{f}}_{\mathrm{c},1,k}(\hat{P}^{\mathrm{f}}_{\mathrm{c},1,k} + R_k)^{-1}\hat{P}^{\mathrm{f}}_{\mathrm{c},1,k}, \tag{4.37}$$

$$\tilde{P}^{\mathrm{da}}_{\mathrm{c},21,k} = \hat{P}^{\mathrm{f}}_{\mathrm{c},21,k} - \hat{P}^{\mathrm{f}}_{\mathrm{c},21,k}(\hat{P}^{\mathrm{f}}_{\mathrm{c},1,k} + R_k)^{-1}\hat{P}^{\mathrm{f}}_{\mathrm{c},1,k}. \tag{4.38}$$

Substituting (4.34) into (4.37) and using (4.25) and (4.26) yields

$$\tilde{P}^{\mathrm{da}}_{\mathrm{c},1,k} = P^{\mathrm{da}}_{1,k}, \quad \tilde{P}^{\mathrm{da}}_{\mathrm{c},21,k} = P^{\mathrm{da}}_{21,k}. \tag{4.39}$$

Moreover, since $\tilde{S}^{\mathrm{da}}_{\mathrm{c},k} = \Phi_{\mathrm{CHOL}}(\tilde{P}^{\mathrm{da}}_{\mathrm{c},k}, q)$, it follows from Lemma IV.1 that

$$\hat{P}^{\mathrm{da}}_{\mathrm{c},1,k} = \tilde{P}^{\mathrm{da}}_{\mathrm{c},1,k}, \quad \hat{P}^{\mathrm{da}}_{\mathrm{c},21,k} = \tilde{P}^{\mathrm{da}}_{\mathrm{c},21,k}. \tag{4.40}$$

It follows from (4.16) and (4.23) that

$$\tilde{P}^{\mathrm{f}}_{\mathrm{c},1,k+1} = A_{1,k}\hat{P}^{\mathrm{da}}_{1,k}A^{\mathrm{T}}_{1,k} + Q_{1,k}, \tag{4.41}$$

$$\tilde{P}^{\mathrm{f}}_{\mathrm{c},21,k+1} = A_{2,k}\hat{P}^{\mathrm{da}}_{21,k}A^{\mathrm{T}}_{1,k} + A_{21,k}\hat{P}^{\mathrm{da}}_{1,k}A^{\mathrm{T}}_{1,k} + Q_{21,k}. \tag{4.42}$$

Therefore, (4.29), (4.39), and (4.40) imply that

$$\tilde{P}^{\mathrm{f}}_{\mathrm{c},1,k+1} = P^{\mathrm{f}}_{1,k+1}, \quad \tilde{P}^{\mathrm{f}}_{\mathrm{c},21,k+1} = P^{\mathrm{f}}_{21,k+1}. \qquad \square$$

Corollary IV.1. *Assume that A_k and C_k have the structure in* (4.23)*. Let $q = p$, $\tilde{P}^{\mathrm{f}}_{\mathrm{c},1,0} = P^{\mathrm{f}}_{1,0}$, $\tilde{P}^{\mathrm{f}}_{\mathrm{c},21,0} = P^{\mathrm{f}}_{21,0}$, and $x^{\mathrm{f}}_{\mathrm{c},0} = x^{\mathrm{f}}_0$. Then for all $k \geqslant 0$, $K_{\mathrm{c},k} = K_k$, and hence $x^{\mathrm{f}}_{\mathrm{c},k} = x^{\mathrm{f}}_k$.*

Proof. Using induction and Proposition IV.1 yields $K_{\mathrm{c},k} = K_k$ for all $k \geqslant 0$. Hence, it follows from (2.5), (2.6), (4.12), and (4.15) that $x^{\mathrm{f}}_{\mathrm{c},k} = x^{\mathrm{f}}_k$ for all $k \geqslant 0$. $\square$

V. Linear Time-Invariant Systems

In this section, we consider a linear time-invariant system and hence assume that, for all $k \geqslant 0$, $A_k = A$, $C_k = C$, $G_k = G$, $H_k = H$, $Q_k = Q$, and $R_k = R$. Furthermore, we assume that $p < n$ and (A, C) is observable so that the observability matrix $\mathcal{O} \in \mathbb{R}^{pn\times n}$ defined by

$$\mathcal{O} \triangleq \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \tag{5.1}$$

has full column rank. Next, without loss of generality we consider a basis such that

$$\mathcal{O} = \begin{bmatrix} I_n \\ 0_{(p-1)n\times n} \end{bmatrix}. \tag{5.2}$$

Therefore, (5.1) and (5.2) imply that, for $i \in \{1,\ldots,n\}$ such that $ip \leqslant n$,

$$CA^{i-1} = \begin{bmatrix} 0_{p\times p(i-1)} & I_p & 0_{p\times(n-pi)} \end{bmatrix}. \tag{5.3}$$

The following result shows that the Cholesky decomposition based rank-q square-root filter is equivalent to the optimal Kalman filter for a specific number of time steps.

Proposition V.1. *Let $r > 0$ be an integer such that $0 < q = pr < n$. Then, for all $k = k_0, \ldots, k_0 + r - 1$, $K_{\mathrm{c,k}} = K_k$.*

Proof. Substituting (2.4) into (2.7) yields

$$P^{\mathrm{f}}_{k+1} = AP^{\mathrm{f}}_k A^{\mathrm{T}} - AP^{\mathrm{f}}_k C^{\mathrm{T}}(CP^{\mathrm{f}}_k C^{\mathrm{T}} + R)^{-1}CP^{\mathrm{f}}_k A^{\mathrm{T}} + Q. \tag{5.4}$$

Hence, for $i = 1, \ldots, r-1$,

$$P^{\mathrm{f}}_{k+1}(A^{i-1})^{\mathrm{T}}C^{\mathrm{T}} = \psi\left(P^{\mathrm{f}}_k(A^i)^{\mathrm{T}}C^{\mathrm{T}}, P^{\mathrm{f}}_k C^{\mathrm{T}}, A, Q, C, R, i\right), \tag{5.5}$$

where

$$\begin{aligned}&\psi(X, Y, A, Q, C, R, i) \\ &\quad = AX - AY(CY + R)^{-1}CX + Q(A^{i-1})^{\mathrm{T}}C^{\mathrm{T}}.\end{aligned} \tag{5.6}$$

Therefore,

$$\begin{aligned}&P^{\mathrm{f}}{}_{k+i}C^{\mathrm{T}} \\ &\quad = \varphi(P^{\mathrm{f}}_k(A^i)^{\mathrm{T}}C^{\mathrm{T}}, P^{\mathrm{f}}_k(A^{i-1})^{\mathrm{T}}C^{\mathrm{T}}, \ldots, P^{\mathrm{f}}_k C^{\mathrm{T}}, A, Q, C, i).\end{aligned} \tag{5.7}$$

Define $\hat{P}^{\mathrm{f}}_{\mathrm{c},k}$ and $\hat{P}^{\mathrm{da}}_{\mathrm{c},k}$ by

$$\hat{P}^{\mathrm{f}}_{\mathrm{c},k} \triangleq \tilde{S}^{\mathrm{f}}_{\mathrm{c},k}(\tilde{S}^{\mathrm{f}}_{\mathrm{c},k})^{\mathrm{T}}, \quad \hat{P}^{\mathrm{da}}_{\mathrm{c},k} \triangleq \tilde{S}^{\mathrm{da}}_{\mathrm{c},k}(\tilde{S}^{\mathrm{da}}_{\mathrm{c},k})^{\mathrm{T}}. \tag{5.8}$$

It follows from Lemma IV.1 and (5.3) that for all $k \geqslant 0$ and $i = 1, \ldots, r$,

$$CA^{i-1}\hat{P}^{\mathrm{f}}_{\mathrm{c},k} = CA^{i-1}\tilde{P}^{\mathrm{f}}_{\mathrm{c},k}, \quad CA^{i-1}\hat{P}^{\mathrm{da}}_{\mathrm{c},k} = CA^{i-1}\tilde{P}^{\mathrm{da}}_{\mathrm{c},k}. \tag{5.9}$$

Note that

$$\tilde{P}^{\mathrm{f}}_{\mathrm{c},k+1} = A\hat{P}^{\mathrm{da}}_{\mathrm{c},k}A^{\mathrm{T}} + Q. \tag{5.10}$$

Multiplying (5.10) by $(A^{i-1})^{\mathrm{T}}C^{\mathrm{T}}$ yields

$$\tilde{P}^{\mathrm{f}}_{\mathrm{c},k+1}(A^{i-1})^{\mathrm{T}}C^{\mathrm{T}} = A\hat{P}^{\mathrm{da}}_{\mathrm{c},k}(A^i)^{\mathrm{T}}C^{\mathrm{T}} + Q(A^{i-1})^{\mathrm{T}}C^{\mathrm{T}}. \tag{5.11}$$

Substituting (4.40) into (5.11) yields

$$\hat{P}^{\mathrm{f}}_{\mathrm{c},k+1}(A^{i-1})^{\mathrm{T}}C^{\mathrm{T}} = A\tilde{P}^{\mathrm{da}}_{\mathrm{c},k}(A^i)^{\mathrm{T}}C^{\mathrm{T}} + Q(A^{i-1})^{\mathrm{T}}C^{\mathrm{T}}, \tag{5.12}$$

for $i = 1, \ldots, r-1$. Using (4.11) in (5.12) yields

$$\begin{aligned}&\hat{P}^{\mathrm{f}}_{\mathrm{c},k+1}(A^{i-1})^{\mathrm{T}}C^{\mathrm{T}} \\ &\quad = A\left[\hat{P}^{\mathrm{f}}_{\mathrm{c},k} - \hat{P}^{\mathrm{f}}_{\mathrm{c},k}C^{\mathrm{T}}(C\hat{P}^{\mathrm{f}}_{\mathrm{c},k}C^{\mathrm{T}} + R)^{-1}C\hat{P}^{\mathrm{f}}_{\mathrm{c},k}\right](A^i)^{\mathrm{T}}C^{\mathrm{T}} \\ &\qquad + Q(A^{i-1})^{\mathrm{T}}C^{\mathrm{T}},\end{aligned} \tag{5.13}$$

Therefore, (5.6) implies that, for all $i = 1, \ldots, r-1$,

$$\begin{aligned}&\hat{P}^{\mathrm{f}}{}_{\mathrm{c},k+1}(A^{i-1})^{\mathrm{T}}C^{\mathrm{T}} \\ &\quad = \psi\left(\hat{P}^{\mathrm{f}}_{\mathrm{c},k}(A^i)^{\mathrm{T}}C^{\mathrm{T}}, \hat{P}^{\mathrm{f}}_{\mathrm{c},k}C^{\mathrm{T}}, A, Q, C, R, i\right).\end{aligned} \tag{5.14}$$

Hence, it follows from (5.5)-(5.7) that for $i = 1, \ldots, r-1$,

$$\begin{aligned}&\hat{P}^{\mathrm{f}}_{k+i}C^{\mathrm{T}} \\ &\quad = \varphi(\hat{P}^{\mathrm{f}}_{\mathrm{c},k}(A^i)^{\mathrm{T}}C^{\mathrm{T}}, \hat{P}^{\mathrm{f}}_{\mathrm{c},k}(A^{i-1})^{\mathrm{T}}C^{\mathrm{T}}, \ldots, \hat{P}^{\mathrm{f}}_{\mathrm{c},k}C^{\mathrm{T}}, A, Q, C, i).\end{aligned} \tag{5.15}$$

Since $\tilde{P}^{\mathrm{f}}_{\mathrm{c},k_0} = P^{\mathrm{f}}_{k_0}$, it follows from Lemma IV.1 that for $i = 1, \ldots, r$,

$$CA^{i-1}\hat{P}^{\mathrm{f}}_{\mathrm{c},k_0} = CA^{i-1}P^{\mathrm{f}}_{k_0}. \tag{5.16}$$

Hence, it follows from (5.7) and (5.15) that for $i = 1, \ldots, r-1$,

$$\hat{P}^{\mathrm{f}}_{k_0+i}C^{\mathrm{T}} = P^{\mathrm{f}}_{k_0+i}C^{\mathrm{T}}. \tag{5.17}$$

Finally, (2.3) and (4.10) imply that for $i = 0, \ldots, r-1$

$$K_{\mathrm{c},k_0+i} = K_{k_0+i}.$$

□

However, in general $P^{\mathrm{f}}_{\mathrm{c},k} \neq P^{\mathrm{f}}_k$ for $k = k_0, \ldots, k_0 + r - 1$, $K_{\mathrm{c},k}$.

VI. Examples

We compare performance of the SVD-based and Cholesky-based reduced-rank square-root Kalman filters using a compartmental model [22] and 10-DOF mass-spring-damper system.

A schematic diagram of the compartmental model is shown in Fig 1. The number n of cells is 20 with one state per cell. All η_{ii} and all η_{ij} $(i \neq j)$ are set to 0.1. We assume that the state of the 9th cell is measured, and disturbances enter all cells so that the disturbance covariance Q has full rank.

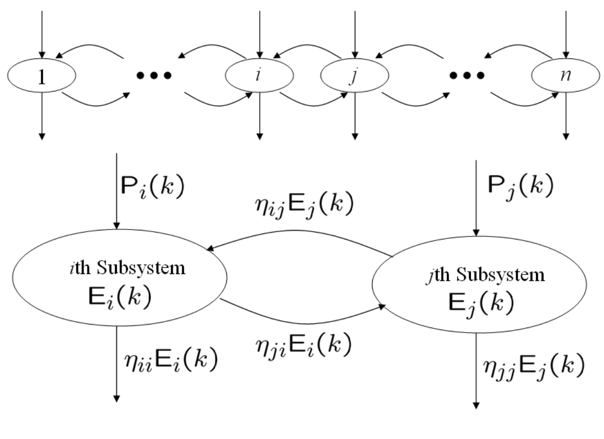


Fig. 1. Compartmental model involving interconnected subsystems.

We simulate three cases of disturbance covariance and compare costs $J \triangleq \mathcal{E}(e_k^{\mathrm{T}} e_k)$ in Figure 3 where (a) shows the cost when $Q = I$, ((b) shows costs when Q is diagonal with entries

$$Q_{i,i} = \begin{cases} 4.0, & \text{if } i = 9, \\ 1.0, & \text{else}, \end{cases} \tag{6.1}$$

and (c) shows costs when Q is diagonal with entries

$$Q_{i,i} = \begin{cases} 0.25, & \text{if } i = 9, \\ 1.0, & \text{else}, \end{cases} \tag{6.2}$$

We reduce the rank of the square-root covariance to 2 in all three cases. In (a) and (b) of Figure 3, the Cholesky-based reduced-rank square-root Kalman filter exhibits almost the same performance as the optimal filter whereas the SVD-based reduced-rank square-root Kalman filter shows degraded performance in (a). Meanwhile, in (c), the SVD-based has a large transient and large steady-state offset from the optimal, whereas the Cholesky-based reduced-rank square-root Kalman filter behaves close to the full-order filter.

The mass-spring-damper model is shown in Figure 2. The total number of masses is 10 with two states (displacement and velocity) per mass. For $i = 1, \ldots, 10$, $m_i = 1$ kg, and $k_j = 1$ N/m, $c_j = 0.01$ Ns/m, $j = 1, \ldots, 11$. We assume that the displacement of the 5th mass is measured and disturbances enter all states so that disturbance covariance Q has full rank.

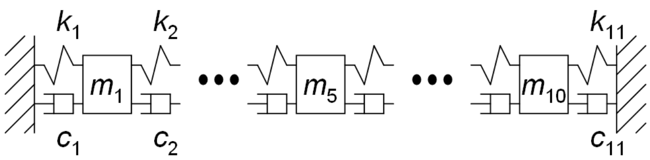


Fig. 2. Mass-spring-damper system.

We simulate three cases of disturbance covariance and compare costs J in (a),(b) and (c) of Figure 4. (a) shows costs when $Q = I$, (b) shows costs when Q is given by (6.1), and (c) shows costs when Q is given by (6.2). We reduce the rank of square-root covariance to 2 in all three cases. In (a) and (b), the costs for the Cholesky-based and SVD-based reduced-rank square-root Kalman filters are close to each other but larger than the optimal. Meanwhile, in (c), the SVD-based shows unstable behavior, while the Cholesky-based filter remains stable and nearly optimal.

VII. Conclusions

We developed a Cholesky decomposition method to obtain reduced-rank square-root Kalman filters. We showed that the SVD-based reduced-rank square-root Kalman filter is equivalent to the optimal filter when the reduced-rank is equal to or greater than the rank of error-covariance, while the Cholesky-based is equivalent to the optimal when the system is lower triangular block-diagonal according to the observation matrix C which has the form $[I_{p\times p} \;\; 0]$ and p is equal to the reduced-rank q. Furthermore, the Cholesky-based rank-q square root filter is equivalent to the optimal Kalman filter for a specific number of time steps. In general cases, the Cholesky-based does not always perform better that the SVD-based and vice versa. Finally, using simulation examples, we showed that the Cholesky-based exhibits more stable performance than the SVD-based filter, which can become unstable when the strong disturbances enter the system states that are not measured.

References

[1] M. Lewis, S. Lakshmivarahan, and S. Dhall, *Dynamic Data Assimilation: A Least Squares Approach*. Cambridge, 2006.

[2] G. Evensen, *Data Assimilation: The Ensemble Kalman Filter*. Springer, 2006.

[3] D. S. Bernstein, L. D. Davis, and D. C. Hyland, "The optimal projection equations for reduced-order, discrete-time modelling, estimation and control," *AIAA J. Guid. Contr. Dyn.*, vol. 9, pp. 288–293, 1986.

[4] D. S. Bernstein and D. C. Hyland, "The optimal projection equations for reduced-order state estimation," *IEEE Trans. Autom. Contr.*, vol. AC-30, pp. 583–585, 1985.

[5] W. M. Haddad and D. S. Bernstein, "Optimal reduced-order observer-estimators," *AIAA J. Guid. Contr. Dyn.*, vol. 13, pp. 1126–1135, 1990.

[6] P. Hippe and C. Wurmthaler, "Optimal reduced-order estimators in the frequency domain: The discrete-time case," *Int. J. Contr.*, vol. 52, pp. 1051–1064, 1990.

[7] C.-S. Hsieh, "The Unified Structure of Unbiased Minimum-Variance Reduced-Order Filters," in *Proc. Contr. Dec. Conf.*, Maui, HI, December 2003, pp. 4871–4876.

[8] J. Ballabrera-Poy, A. J. Busalacchi, and R. Murtugudde, "Application of a reduced-order kalman filter to initialize a coupled atmosphere-ocean model: Impact on the prediction of el nino," *J. Climate*, vol. 14, pp.1720–1737, 2001.

[9] S. F. Farrell and P. J. Ioannou, "State estimation using a reduced-order Kalman filter," *J. Atmos. Sci.*, vol. 58, pp. 3666–3680, 2001.

[10] P. Fieguth, D. Menemenlis, and I. Fukumori, "Mapping and Pseudo-Inverse Algorithms for Data Assimilation," in *Proc. Int. Geoscience Remote Sensing Symp.*, 2002, pp. 3221–3223.

[11] L. Scherliess, R. W. Schunk, J. J. Sojka, and D. C. Thompson, "Development of a physics-based reduced state kalman filter for the ionosphere," *Radio Science*, vol. 39-RS1S04, 2004.

[12] I. S. Kim, J. Chandrasekar, H. J. Palanthandalam-Madapusi, A. Ridley, and D. S. Bernstein, "State Estimation for Large-Scale Systems Based on Reduced-Order Error-Covariance Propagation," in *Proc. Amer. Contr. Conf.*, New York, June 2007.

[13] S. Gillijns, D. S. Bernstein, and B. D. Moor, "The Reduced Rank Transform Square Root Filter for Data Assimilation," *Proc. of the 14th IFAC Symposium on System Identification (SYSID2006),Newcastle, Australia*, vol. 11, pp. 349–368, 2006.

[14] A. W. Heemink, M. Verlaan, and A. J. Segers, "Variance Reduced Ensemble Kalman Filtering," *Monthly Weather Review*, vol. 129, pp. 1718–1728, 2001.

[15] D. Treebushny and H. Madsen, "A New Reduced Rank Square Root Kalman Filter for Data Assimilation in Mathematical Models," *Lecture Notes in Computer Science*, vol. 2657, pp. 482–491, 2003.

[16] M. Verlaan and A. W. Heemink, "Tidal flow forecasting using reduced rank square root filters," *Stochastics Hydrology and Hydraulics*, vol. 11, pp. 349–368, 1997.

[17] J. L. Anderson, "An Ensemble Adjustment Kalman Filter for Data Assimilation," *Monthly Weather Review*, vol. 129, pp. 2884–2903, 2001.

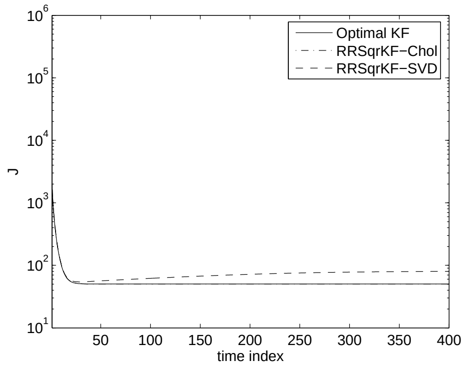
[18] M. K. Tippett, J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, *Monthly Weather Review*, vol. 131, pp. 1485–1490, 2003.

[19] G. J. Bierman, *Factorization Methods for Discrete Sequential Estimation*. reprinted by Dover, 2006.

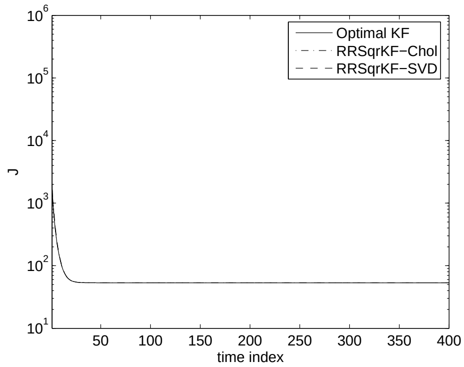
[20] M. Morf and T. Kailath, "Square-root algorithms for least-squares estimation," *IEEE Trans. Autom. Contr.*, vol. AC-20, pp. 487–497, 1975.

[21] G. W. Stewart, "Matrix Algorithms Volume 1: Basic Decompositions," *SIAM*, 1998.

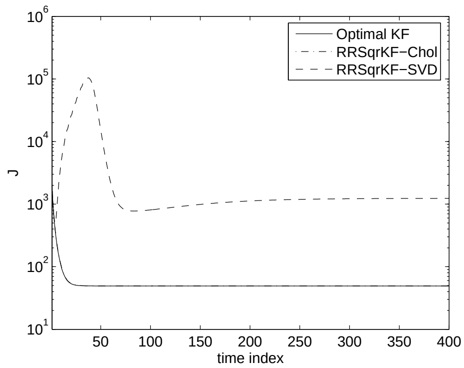
[22] D. S. Bernstein and D. C. Hyland. Compartmental modeling and second-moment analysis of state space systems. *SIAM J. Matrix Anal. Appl.*, 14(3):880–901, 1993.



(a)

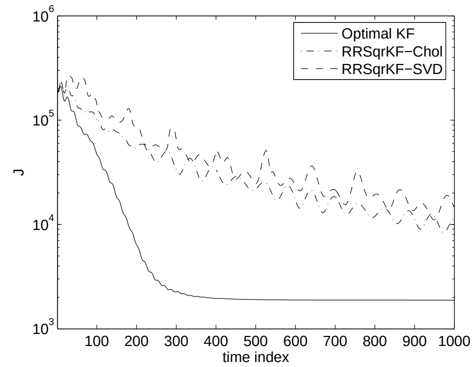


(b)

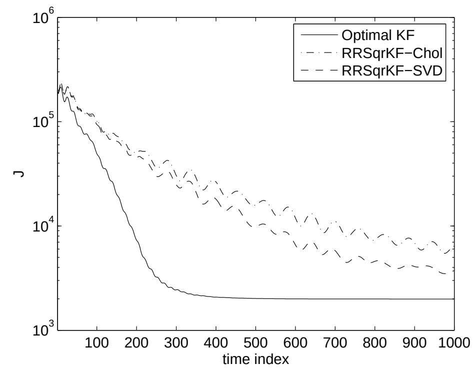


(c)

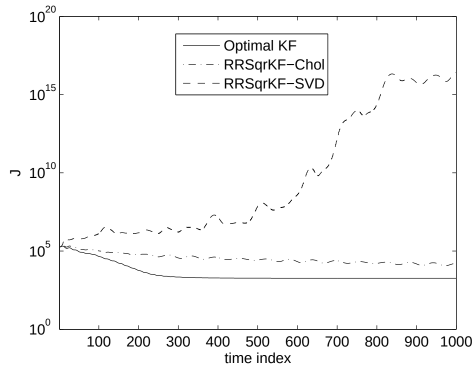
Fig. 3. Time evolutions of the cost $J \triangleq \mathcal{E}(e_k^{\mathrm{T}} e_k)$ for the compartmental system. For this example, $R = 10^{-4}$, $P_0 = 10^2 I$, the rank of reduced rank filters is fixed at 2 and energy measurement is taken at compartment 9. Disturbances enter compartments 1,2,. . .,20. In (a), the cost J is when disturbance covariance is $Q = I_{20\times 20}$, (b) is for the case when $Q_{(9,9)}$ is changed to 4.0 and (c) is for the case when $Q_{(9,9)}$ is changed to 0.25.



(a)



(b)



(c)

Fig. 4. Time evolutions of the cost $J = \mathcal{E}(e_k^{\mathrm{T}} e_k)$ of the mass-spring-damper system. Here $R = 10^{-4}$, $P_0 = 100^2 I$, the rank of reduced rank filters is fixed at 2, and a displacement measurement of m_5 is available. Velocity and force disturbances enter mass 1,2,. . .,10. (a) shows the cost J when disturbance covariance is $Q = I_{20\times 20}$ while (b) is for the case when $Q_{(9,9)}$ is changed to 4.0 and (c) is for the case when $Q_{(9,9)}$ is changed to 0.25.