2008 American Control Conference
Westin Seattle Hotel, Seattle, Washington, USA
June 11-13, 2008

Cholesky-Based Reduced-Rank Square-Root Kalman Filtering

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

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, 2], 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.

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, 4]. Alternative reduced-order variants of the classical Kalman filter have been developed for computationally demanding applications [5, 6], 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 [7].

A widely studied technique for reducing the computational requirements of the Kalman filter for large scale systems is the *reduced-rank filter* [8, 9]. 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 [10, 11].

The primary technique for truncating the error-covariance matrix is the singular value decomposition (SVD) [8–11], 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 criteria to support the optimality of approximation based on the SVD within the context of recursive state estimation.

In the present paper we begin by observing that the Kalman filter update depends on the product C_kP_k, where C_k is the measurement map and P_k is the error covariance. To develop this idea, we show that approximation of C_kP_k leads directly to truncation based on the Cholesky decomposition. 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. An additional advantage of using the Cholesky decomposition in place of the SVD for reduced-rank filtering is the fact that the Cholesky decomposition is computationally less expensive than the SVD, specifically, $O(n^3/6)$ versus $O(2n^3)$[12].

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 The University of Michigan, Ann Arbor, MI, dsbaero@umich.edu

II. The Kalman filter

Consider the discrete-time system

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

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

where $x_k \in \mathbb{R}^n$, $w_k \in \mathbb{R}^{d_w}$, $y_k \in \mathbb{R}^p$, $v_k \in \mathbb{R}^{d_v}$, 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 $Q_k \triangleq G_kG_k^{\mathrm{T}}$ and $R_k \triangleq H_kH_k^{\mathrm{T}}$ 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.

Data Assimilation Step

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

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

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

Forecast Step

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

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

The states x_k^{f} and x_k^{da} are the forecast and data assimilation estimates of the state x_k, while 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,

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

where $e_k^{\mathrm{f}} \triangleq x_k - x_k^{\mathrm{f}}$, $e_k^{\mathrm{da}} \triangleq x_k - x_k^{\mathrm{da}}$.

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

To reduce the computational requirements, we consider a filter that uses reduced-rank approximations of the error covariances. Instead of updating the error covariances, we propagate predicted error covariances $\tilde{P}_{\mathrm{s},k}^{\mathrm{da}}$ and $\tilde{P}_{\mathrm{s},k}^{\mathrm{f}}$ using reduced-rank approximations $\hat{P}_{\mathrm{s},k}^{\mathrm{da}}$ and $\hat{P}_{\mathrm{s},k}^{\mathrm{f}}$. The reduced-rank approximations are chosen so that $\mathrm{rank}(\hat{P}_{\mathrm{s},k}^{\mathrm{da}}) < n$ and $\mathrm{rank}(\hat{P}_{\mathrm{s},k}^{\mathrm{f}}) < n$, and such that $\|\tilde{P}_{\mathrm{s},k}^{\mathrm{da}} - \hat{P}_{\mathrm{s},k}^{\mathrm{da}}\|_{\mathrm{F}}$ and $\|\tilde{P}_{\mathrm{s},k}^{\mathrm{f}} - \hat{P}_k^{\mathrm{f}}\|_{\mathrm{F}}$ are minimized.

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 $u_1, \ldots, u_n \in \mathbb{R}^n$ be the corresponding orthogonal singular vectors. Next, define $U_q \in \mathbb{R}^{n\times q}$ and $\Sigma_q \in \mathbb{R}^{q\times q}$ by

$$U_q \triangleq \begin{bmatrix} u_1 & \cdots & u_q \end{bmatrix}, \quad \Sigma_q \triangleq \operatorname{diag}(\sigma_1, \cdots, \sigma_q). \quad (3.1)$$

With this notation, the singular value decomposition of P is given by $P = U_n \Sigma_n U_n^{\mathrm{T}}$, 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 a square-root of P given by $\Phi_{\mathrm{SVD}}(P, q) \triangleq U_q \Sigma_q^{1/2}$. Note that SS^{T}, where $S \triangleq \Phi_{\mathrm{SVD}}(P, q)$, is the best rank-$q$ approximation of P in the Frobenius norm. Specifically, denoting the Frobenius norm by $\|\cdot\|_{\mathrm{F}}$, we have the following result in [14].

Lemma 3.1: Let $P \in \mathbb{R}^{n\times n}$ be positive semidefinite, and let $\sigma_1 \geqslant \cdots \geqslant \sigma_n$ be the singular values of P. If $S = \Phi_{\mathrm{SVD}}(P, q)$, then

$$\min_{\operatorname{rank}(\hat{P})=q} \|P - \hat{P}\|_{\mathrm{F}} = \|P - SS^{\mathrm{T}}\|_{\mathrm{F}}^2 = \sigma_{q+1}^2 + \cdots + \sigma_n^2. \quad (3.2)$$

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

Data Assimilation step

$$K_{\mathrm{s},k} = \hat{P}_{\mathrm{s},k}^{\mathrm{f}} C_k^{\mathrm{T}} (C_k \hat{P}_{\mathrm{s},k}^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1}, \quad (3.3)$$

$$\tilde{P}_{\mathrm{s},k}^{\mathrm{da}} = \hat{P}_{\mathrm{s},k}^{\mathrm{f}} - \hat{P}_{\mathrm{s},k}^{\mathrm{f}} C_k^{\mathrm{T}} (C_k \hat{P}_{\mathrm{s},k}^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1} C_k \hat{P}_{\mathrm{s},k}^{\mathrm{f}}, \quad (3.4)$$

$$x_{\mathrm{s},k}^{\mathrm{da}} = x_{\mathrm{s},k}^{\mathrm{f}} + K_{\mathrm{s},k}(y_k - C_k x_{\mathrm{s},k}^{\mathrm{f}}), \quad (3.5)$$

Forecast step

$$x_{\mathrm{s},k+1}^{\mathrm{f}} = A_k x_{\mathrm{s},k}^{\mathrm{da}}, \quad (3.6)$$

$$\tilde{P}_{\mathrm{s},k+1}^{\mathrm{f}} = A_k \hat{P}_{\mathrm{s},k}^{\mathrm{da}} A_k^{\mathrm{T}} + Q_k, \quad (3.7)$$

where $\hat{P}_{\mathrm{s},k}^{\mathrm{f}} \triangleq \tilde{S}_{\mathrm{s},k}^{\mathrm{f}} (\tilde{S}_{\mathrm{s},k}^{\mathrm{f}})^{\mathrm{T}}$, $\hat{P}_{\mathrm{s},k}^{\mathrm{da}} \triangleq \tilde{S}_{\mathrm{s},k}^{\mathrm{da}} (\tilde{S}_{\mathrm{s},k}^{\mathrm{da}})^{\mathrm{T}}$, $\tilde{S}_{\mathrm{s},k}^{\mathrm{f}} \triangleq \Phi_{\mathrm{SVD}}(\tilde{P}_{\mathrm{s},k}^{\mathrm{f}}, q)$, $\tilde{S}_{\mathrm{s},k}^{\mathrm{da}} \triangleq \Phi_{\mathrm{SVD}}(\tilde{P}_{\mathrm{s},k}^{\mathrm{da}}, q)$, 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

$$P_{\mathrm{s},k}^{\mathrm{f}} \triangleq \mathcal{E}[(x_k - x_{\mathrm{s},k}^{\mathrm{f}})(x_k - x_{\mathrm{s},k}^{\mathrm{f}})^{\mathrm{T}}], \quad (3.8)$$

$$P_{\mathrm{s},k}^{\mathrm{da}} \triangleq \mathcal{E}[(x_k - x_{\mathrm{s},k}^{\mathrm{da}})(x_k - x_{\mathrm{s},k}^{\mathrm{da}})^{\mathrm{T}}]. \quad (3.9)$$

Using (2.1), (3.5) and (3.6), it follows that

$$P_{\mathrm{s},k}^{\mathrm{da}} = (I - K_{\mathrm{s},k} C_k) P_{\mathrm{s},k}^{\mathrm{f}} (I - K_{\mathrm{s},k} C_k)^{\mathrm{T}} + K_{\mathrm{s},k} R_k K_{\mathrm{s},k}^{\mathrm{T}}, \quad (3.10)$$

$$P_{\mathrm{s},k}^{\mathrm{f}} = A_k P_{\mathrm{s},k}^{\mathrm{da}} A_k^{\mathrm{T}} + Q_k. \quad (3.11)$$

Note that $\tilde{P}_{\mathrm{s},k}^{\mathrm{f}}$ and $\tilde{P}_{\mathrm{s},k}^{\mathrm{da}}$ are predicted error covariances and not covariances of the state error. Furthermore, since $K_{\mathrm{s},k} \neq K_k$, the SVD-based rank-$q$ square-root filter is a suboptimal filter. However, under certain conditions, the SVD-based rank-q square-root filter is equivalent to the Kalman filter.

Proposition 3.1: Assume that $\tilde{P}_{\mathrm{s},k}^{\mathrm{f}} = P_k^{\mathrm{f}}$ and $\operatorname{rank}(P_k^{\mathrm{f}}) \leqslant q$. Then, $K_{\mathrm{s},k} = K_k$, $\tilde{P}_{\mathrm{s},k}^{\mathrm{da}} = P_k^{\mathrm{da}}$, and $\tilde{P}_{\mathrm{s},k+1}^{\mathrm{f}} = P_{k+1}^{\mathrm{f}}$.

Proof. Since $\operatorname{rank}(\tilde{P}_{\mathrm{s},k}^{\mathrm{f}}) \leqslant q$, it follows from Lemma 3.1 that

$$\hat{P}_{\mathrm{s},k}^{\mathrm{f}} = \tilde{S}_{\mathrm{s},k}^{\mathrm{f}} \left(\tilde{S}_{\mathrm{s},k}^{\mathrm{f}}\right)^{\mathrm{T}} = \tilde{P}_{\mathrm{s},k}^{\mathrm{f}}. \quad (3.12)$$

Hence, it follows from (3.3) that $K_{\mathrm{s},k} = K_k$. Furthermore, it follows from (2.4), (3.4), and (3.12) that

$$\tilde{P}_{\mathrm{s},k}^{\mathrm{da}} = P_k^{\mathrm{da}}. \quad (3.13)$$

Since $\operatorname{rank}(P_k^{\mathrm{f}}) \leqslant q$, it follows from (2.4) that $\operatorname{rank}(P_k^{\mathrm{da}}) \leqslant q$ and hence (3.13) implies that $\operatorname{rank}(\tilde{P}_{\mathrm{s},k}^{\mathrm{da}}) \leqslant q$. Therefore, Lemma 3.1, (**??**) and (**??**) imply that

$$\hat{P}_{\mathrm{s},k}^{\mathrm{da}} = \tilde{S}_{\mathrm{s},k}^{\mathrm{da}} \left(\tilde{S}_{\mathrm{s},k}^{\mathrm{da}}\right)^{\mathrm{T}} = \tilde{P}_{\mathrm{s},k}^{\mathrm{da}}. \quad (3.14)$$

Hence, it follows from (3.13) and (3.14) that $\hat{P}_{\mathrm{s},k}^{\mathrm{da}} = P_k^{\mathrm{da}}$, and therefore (2.7) and (3.7) imply that $\tilde{P}_{\mathrm{s},k+1}^{\mathrm{f}} = P_{k+1}^{\mathrm{f}}$. □

Corollary 3.1: Assume that $x_{\mathrm{s},0}^{\mathrm{f}} = x_0^{\mathrm{f}}$, $\tilde{P}_{\mathrm{s},0}^{\mathrm{f}} = P_0^{\mathrm{f}}$, and $\operatorname{rank}(P_0^{\mathrm{f}}) \leqslant q$. Furthermore, assume that, for all $k \geqslant 0$, $\operatorname{rank}(A_k) + \operatorname{rank}(Q_k) \leqslant q$. Then, for all $k \geqslant 0$, $K_{\mathrm{s},k} = K_k$ and $x_{\mathrm{s},k}^{\mathrm{f}} = x_k^{\mathrm{f}}$.

Proof. Using Proposition 3.1 and induction, it can be shown that $K_{\mathrm{s},k} = K_k$ and $x_{\mathrm{s},k}^{\mathrm{f}} = x_k^{\mathrm{f}}$ for all $k \geqslant 0$. □

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

The Kalman filter gain K_k depends only on the correlation $C_k P_k^{\mathrm{f}}$ between the error in the measured states and the unmeasured states. We thus have the following observation.

Lemma 4.1: Assume that $\hat{P}_k \in \mathbb{R}^{n\times n}$ is positive semidefinite. Partition $\hat{P}_k$ and P_k^{f} as

$$\hat{P}_k = \begin{bmatrix} \hat{P}_{q,k} & (\hat{P}_{\bar{q}q,k})^{\mathrm{T}} \\ \hat{P}_{\bar{q}q,k} & \hat{P}_{\bar{q},k} \end{bmatrix}, P_k^{\mathrm{f}} = \begin{bmatrix} P_{q,k}^{\mathrm{f}} & (P_{\bar{q}q,k}^{\mathrm{f}})^{\mathrm{T}} \\ P_{\bar{q}q,k}^{\mathrm{f}} & P_{\bar{q},k}^{\mathrm{f}} \end{bmatrix}, \quad (4.1)$$

where $\hat{P}_{q,k}, P_{q,k}^{\mathrm{f}} \in \mathbb{R}^{q\times q}$ and $\hat{P}_{\bar{q},k}, P_{\bar{q},k}^{\mathrm{f}} \in \mathbb{R}^{\bar{q}\times\bar{q}}$, assume that C_k has the form

$$C_k = \begin{bmatrix} I_q & 0 \end{bmatrix}, \quad (4.2)$$

and define $\hat{K}_k$ by

$$\hat{K}_k \triangleq \hat{P}_k C_k^{\mathrm{T}} (C_k \hat{P}_k C_k^{\mathrm{T}} + R_k)^{-1}. \quad (4.3)$$

Furthermore, let $\begin{bmatrix} \hat{P}_{q,k} & \left(\hat{P}_{\bar{q}q,k}\right)^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} P_{q,k}^{\mathrm{f}} & \left(P_{\bar{q}q,k}^{\mathrm{f}}\right)^{\mathrm{T}} \end{bmatrix}$. Then, $\hat{K}_k = K_k$.

Proof. The result follows from (2.3), (4.1), (4.2) and (4.3) that $\hat{K}_k = K_k$. □

Next, we consider a filter that updates the predicted error covariances $\tilde{P}_{\mathrm{c},k}^{\mathrm{da}}$ and $\tilde{P}_{\mathrm{c},k}^{\mathrm{f}}$ using reduced-rank approximations $\hat{P}_{\mathrm{c},k}^{\mathrm{da}}$ and $\hat{P}_{\mathrm{c},k}^{\mathrm{f}}$ such that $\operatorname{rank}(\hat{P}_{\mathrm{c},k}^{\mathrm{da}}) < n$ and $\operatorname{rank}(\hat{P}_{\mathrm{c},k}^{\mathrm{f}}) < n$, and such that $\|C_k(\tilde{P}_{\mathrm{c},k}^{\mathrm{da}} - \hat{P}_{\mathrm{c},k}^{\mathrm{da}})\|_{\mathrm{F}}$ and $\|C_k(\tilde{P}_{\mathrm{c},k}^{\mathrm{f}} - \hat{P}_{\mathrm{c},k}^{\mathrm{f}})\|_{\mathrm{F}}$ are minimized.

Let $P \in \mathbb{R}^{n\times n}$ be positive definite. The Cholesky factorization yields a lower triangular Cholesky factor $L \in \mathbb{R}^{n\times n}$ that satisfies $LL^{\mathrm{T}} = P$. Partition L as $L = \begin{bmatrix} L_1 & \cdots & L_n \end{bmatrix}$, so that truncating the last $n-q$ columns of L yields the rank-q Cholesky factor

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

Lemma 4.2: Let $P \in \mathbb{R}^{n\times n}$ be positive definite, define $S \triangleq \Phi_{\mathrm{CHOL}}(P, q)$ and $\hat{P} \triangleq SS^{\mathrm{T}}$, and partition P and $\hat{P}$ as

$$P = \begin{bmatrix} P_q & P_{q\bar{q}} \\ (P_{q\bar{q}})^{\mathrm{T}} & P_{\bar{q}} \end{bmatrix}, \quad \hat{P} = \begin{bmatrix} \hat{P}_q & \hat{P}_{q\bar{q}} \\ (\hat{P}_{q\bar{q}})^{\mathrm{T}} & \hat{P}_{\bar{q}} \end{bmatrix}, \quad (4.5)$$

where $P_q, \hat{P}_q \in \mathbb{R}^{q\times q}$ and $P_{\overline{q}}, \hat{P}_{\overline{q}} \in \mathbb{R}^{\overline{q}\times\overline{q}}$. Then, $\left[\begin{array}{cc} \hat{P}_q & \hat{P}_{q\overline{q}} \end{array}\right] = \left[\begin{array}{cc} P_q & P_{q\overline{q}} \end{array}\right]$.

Proof. Let L be the Cholesky factor of P. Since L is lower triangular, $L_i L_i^{\mathrm{T}} \in \mathbb{R}^{n\times n}$ has the structure

$$L_i L_i^{\mathrm{T}} = \left[\begin{array}{cc} 0_{i-1} & 0_{(i-1)\times(n-i+1)} \\ 0_{(n-i+1)\times(i-1)} & \sharp \end{array}\right], \tag{4.6}$$

where $\sharp$ denotes an inconsequential entry, and therefore

$$\sum_{i=q+1}^{n} L_i L_i^{\mathrm{T}} = \left[\begin{array}{cc} 0_q & 0_{q\times\overline{q}} \\ 0_{\overline{q}\times q} & \sharp \end{array}\right]. \tag{4.7}$$

Since $P = \sum_{i=1}^{n} L_i L_i^{\mathrm{T}}$, it follows from (4.4) that $P = \hat{P} + \sum_{i=q+1}^{n} L_i L_i^{\mathrm{T}}$. Therefore, using (4.7) yields $\hat{P}_q = P_q$ and $\hat{P}_{q\overline{q}} = P_{q\overline{q}}$. □

Lemma 4.2 implies that, if $S = \Phi_{\mathrm{CHOL}}(P, q)$, then the first q columns and rows of SS^{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

$$K_{\mathrm{c},k} = \hat{P}^{\mathrm{f}}_{\mathrm{c},k} C_k^{\mathrm{T}} (C_k \hat{P}^{\mathrm{f}}_{\mathrm{c},k} C_k^{\mathrm{T}} + R_k)^{-1}, \tag{4.8}$$

$$\tilde{P}^{\mathrm{da}}_{\mathrm{c},k} = \hat{P}^{\mathrm{f}}_{\mathrm{c},k} - \hat{P}^{\mathrm{f}}_{\mathrm{c},k} C_k^{\mathrm{T}} (C_k \hat{P}^{\mathrm{f}}_{\mathrm{c},k} C_k^{\mathrm{T}} + R_k)^{-1} C_k \hat{P}^{\mathrm{f}}_{\mathrm{c},k}, \tag{4.9}$$

$$x^{\mathrm{da}}_{\mathrm{c},k} = x^{\mathrm{f}}_{\mathrm{c},k} + K_{\mathrm{c},k}(y_k - C_k x^{\mathrm{f}}_{\mathrm{c},k}). \tag{4.10}$$

Forecast step

$$x^{\mathrm{f}}_{\mathrm{c},k+1} = A_k x^{\mathrm{da}}_{\mathrm{c},k}, \tag{4.11}$$

$$\tilde{P}^{\mathrm{f}}_{\mathrm{c},k+1} = A_k \hat{P}^{\mathrm{da}}_{\mathrm{c},k} A_k^{\mathrm{T}} + Q_k, \tag{4.12}$$

where

$$\hat{P}^{\mathrm{f}}_{\mathrm{c},k} \triangleq \tilde{S}^{\mathrm{f}}_{\mathrm{c},k} \left(\tilde{S}^{\mathrm{f}}_{\mathrm{c},k}\right)^{\mathrm{T}}, \quad \hat{P}^{\mathrm{da}}_{\mathrm{c},k} \triangleq \tilde{S}^{\mathrm{da}}_{\mathrm{c},k} \left(\tilde{S}^{\mathrm{da}}_{\mathrm{c},k}\right)^{\mathrm{T}}, \tag{4.13}$$

$$\tilde{S}^{\mathrm{f}}_{\mathrm{c},k} \triangleq \Phi_{\mathrm{CHOL}}(\tilde{P}^{\mathrm{f}}_{\mathrm{c},k}, q), \quad \tilde{S}^{\mathrm{da}}_{\mathrm{c},k} \triangleq \Phi_{\mathrm{CHOL}}(\tilde{P}^{\mathrm{da}}_{\mathrm{c},k}, q), \tag{4.14}$$

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

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

$$P^{\mathrm{f}}_{\mathrm{c},k} \triangleq \mathcal{E}[(x_k - x^{\mathrm{f}}_{\mathrm{c},k})(x_k - x^{\mathrm{f}}_{\mathrm{c},k})^{\mathrm{T}}], \tag{4.15}$$

$$P^{\mathrm{da}}_{\mathrm{s},k} \triangleq \mathcal{E}[(x_k - x^{\mathrm{da}}_{\mathrm{c},k})(x_k - x^{\mathrm{da}}_{\mathrm{c},k})^{\mathrm{T}}], \tag{4.16}$$

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), (4.10) and (4.11), it can be shown that

$$P^{\mathrm{da}}_{\mathrm{c},k} = (I - K_{\mathrm{c},k} C_k) P^{\mathrm{f}}_{\mathrm{c},k} (I - K_{\mathrm{c},k} C_k)^{\mathrm{T}} + K_{\mathrm{c},k} R_k K_{\mathrm{c},k}^{\mathrm{T}}, \tag{4.17}$$

$$P^{\mathrm{f}}_{\mathrm{c},k} = A_k P^{\mathrm{da}}_{\mathrm{c},k} A_k^{\mathrm{T}} + Q_k. \tag{4.18}$$

Again, like the SVD-based rank-q square-root filter, the Cholesky-based rank-q square-root filter is suboptimal and generally not equivalent to the Kalman filter. However, the following result shows that, in certain cases, the Cholesky-based rank-q square-root filter is equivalent to the Kalman filter.

Proposition 4.1: Assume that $p = q$, C_k has the form

$$C_k = \left[\begin{array}{cc} I_q & 0 \end{array}\right], \tag{4.19}$$

partition P^{f}_k and $\tilde{P}^{\mathrm{f}}_{\mathrm{c},k}$ as

$$P^{\mathrm{f}}_k = \left[\begin{array}{cc} P^{\mathrm{f}}_{q,k} & (P^{\mathrm{f}}_{\overline{q}q,k})^{\mathrm{T}} \\ P^{\mathrm{f}}_{\overline{q}q,k} & P^{\mathrm{f}}_{\overline{q},k} \end{array}\right], \tilde{P}^{\mathrm{f}}_{\mathrm{c},k} = \left[\begin{array}{cc} \tilde{P}^{\mathrm{f}}_{\mathrm{c},q,k} & (\tilde{P}^{\mathrm{f}}_{\mathrm{c},\overline{q}q,k})^{\mathrm{T}} \\ \tilde{P}^{\mathrm{f}}_{\mathrm{c},\overline{q}q,k} & \tilde{P}^{\mathrm{f}}_{\mathrm{c}\overline{q},k} \end{array}\right], \tag{4.20}$$

where $P^{\mathrm{f}}_{q,k}, \tilde{P}^{\mathrm{f}}_{\mathrm{c},q,k} \in \mathbb{R}^{q\times q}$ and $P^{\mathrm{f}}_{\overline{q},k}, \tilde{P}^{\mathrm{f}}_{\mathrm{c},\overline{q},k} \in \mathbb{R}^{\overline{q}\times\overline{q}}$, and assume that $[\tilde{P}^{\mathrm{f}}_{\mathrm{c},q,k}\ \tilde{P}^{\mathrm{f}}_{\mathrm{c},\overline{q}q,k}] = [P^{\mathrm{f}}_{q,k}\ P^{\mathrm{f}}_{\overline{q}q,k}]$. Then, $K_{\mathrm{c},k} = K_k$. If, in addition, A_k has the form

$$A_k = \left[\begin{array}{cc} A_{q,k} & 0 \\ A_{\overline{q}q,k} & A_{\overline{q},k} \end{array}\right], \tag{4.21}$$

where $A_{q,k} \in \mathbb{R}^{q\times q}$ and $A_{\overline{q},k} \in \mathbb{R}^{\overline{q}\times\overline{q}}$, then $\left[\begin{array}{cc} \tilde{P}^{\mathrm{f}}_{\mathrm{c},q,k+1} & \tilde{P}^{\mathrm{f}}_{\mathrm{c},\overline{q}q,k+1} \end{array}\right] = \left[\begin{array}{cc} P^{\mathrm{f}}_{q,k+1} & P^{\mathrm{f}}_{\overline{q}q,k+1} \end{array}\right]$.

Proof. Partition $\hat{P}^{\mathrm{f}}_{\mathrm{c},k}$ as

$$\hat{P}^{\mathrm{f}}_{\mathrm{c},k} = \left[\begin{array}{cc} \hat{P}^{\mathrm{f}}_{\mathrm{c},q,k} & (\hat{P}^{\mathrm{f}}_{\mathrm{c},\overline{q}q,k})^{\mathrm{T}} \\ \hat{P}^{\mathrm{f}}_{\mathrm{c},\overline{q}q,k} & \hat{P}^{\mathrm{f}}_{\mathrm{c},\overline{q},k} \end{array}\right], \tag{4.22}$$

where $\hat{P}^{\mathrm{f}}_{q,k} \in \mathbb{R}^{q\times q}$ is positive semidefinite and $\hat{P}^{\mathrm{f}}_{\overline{q},k} \in \mathbb{R}^{\overline{q}\times\overline{q}}$. It follows from Lemma 4.2 and (4.14) that

$$\hat{P}^{\mathrm{f}}_{\mathrm{c},q,k} = \tilde{P}^{\mathrm{f}}_{\mathrm{c},q,k}, \quad \hat{P}^{\mathrm{f}}_{\mathrm{c},\overline{q}q,k} = \tilde{P}^{\mathrm{f}}_{\mathrm{c},\overline{q}q,k}. \tag{4.23}$$

Therefore, it follows from Lemma 4.1 and (4.8) that $K_{\mathrm{c},k} = K_k$.

Next, partition P^{da}_k as

$$P^{\mathrm{da}}_k = \left[\begin{array}{cc} P^{\mathrm{da}}_{q,k} & (P^{\mathrm{da}}_{\overline{q}q,k})^{\mathrm{T}} \\ P^{\mathrm{da}}_{\overline{q}q,k} & P^{\mathrm{da}}_{\overline{q},k} \end{array}\right], \tag{4.24}$$

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

$$P^{\mathrm{da}}_{q,k} = P^{\mathrm{f}}_{q,k} - P^{\mathrm{f}}_{q,k}(P^{\mathrm{f}}_{q,k} + R_k)^{-1} P^{\mathrm{f}}_{q,k}, \tag{4.25}$$

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

Now, partition $\tilde{P}^{\mathrm{da}}_{\mathrm{c},k}$ and $\hat{P}^{\mathrm{da}}_{\mathrm{c},k}$ as

$$\tilde{P}^{\mathrm{da}}_{\mathrm{c},k} = \left[\begin{array}{cc} \tilde{P}^{\mathrm{da}}_{\mathrm{c},q,k} & (\tilde{P}^{\mathrm{da}}_{\mathrm{c},\overline{q}q,k})^{\mathrm{T}} \\ \tilde{P}^{\mathrm{da}}_{\mathrm{c},\overline{q}q,k} & \tilde{P}^{\mathrm{da}}_{\mathrm{c},\overline{q},k} \end{array}\right], \hat{P}^{\mathrm{da}}_{\mathrm{c},k} = \left[\begin{array}{cc} \hat{P}^{\mathrm{da}}_{\mathrm{c},q,k} & (\hat{P}^{\mathrm{da}}_{\mathrm{c},\overline{q}q,k})^{\mathrm{T}} \\ \hat{P}^{\mathrm{da}}_{\mathrm{c},\overline{q}q,k} & \hat{P}^{\mathrm{da}}_{\mathrm{c},\overline{q},k} \end{array}\right], \tag{4.27}$$

where $\tilde{P}^{\mathrm{da}}_{q,k}, \hat{P}^{\mathrm{da}}_{q,k} \in \mathbb{R}^{q\times q}$ are positive semidefinite and $\tilde{P}^{\mathrm{da}}_{\overline{q},k}, \hat{P}^{\mathrm{da}}_{\overline{q},k} \in \mathbb{R}^{\overline{q}\times\overline{q}}$. Therefore, it follows from (4.9), (4.19), (4.22), and (4.27) that

$$\tilde{P}^{\mathrm{da}}_{\mathrm{c},q,k} = \hat{P}^{\mathrm{f}}_{\mathrm{c},q,k} - \hat{P}^{\mathrm{f}}_{\mathrm{c},q,k}(\hat{P}^{\mathrm{f}}_{\mathrm{c},q,k} + R_k)^{-1} \hat{P}^{\mathrm{f}}_{\mathrm{c},q,k}, \tag{4.28}$$

$$\tilde{P}^{\mathrm{da}}_{\mathrm{c},\overline{q}q,k} = \hat{P}^{\mathrm{f}}_{\mathrm{c},\overline{q}q,k} - \hat{P}^{\mathrm{f}}_{\mathrm{c},\overline{q}q,k}(\hat{P}^{\mathrm{f}}_{\mathrm{c},q,k} + R_k)^{-1} \hat{P}^{\mathrm{f}}_{\mathrm{c},q,k}. \tag{4.29}$$

Hence, comparing (4.25) with (4.28) and (4.26) with (4.29), and using $\left[\tilde{P}^{\mathrm{f}}_{\mathrm{c},q,k} \left(\tilde{P}^{\mathrm{f}}_{\mathrm{c},\overline{q}q,k}\right)^{\mathrm{T}}\right] = \left[P^{\mathrm{f}}_{q,k} \left(P^{\mathrm{f}}_{\overline{q}q,k}\right)^{\mathrm{T}}\right]$ and (4.23) yields

$$\tilde{P}^{\mathrm{da}}_{\mathrm{c},q,k} = P^{\mathrm{da}}_{q,k}, \quad \tilde{P}^{\mathrm{da}}_{\mathrm{c},\overline{q}q,k} = P^{\mathrm{da}}_{\overline{q}q,k}. \tag{4.30}$$

Moreover, since $S^{\mathrm{da}}_{\mathrm{c},k} = \Phi_{\mathrm{CHOL}}(\tilde{P}^{\mathrm{da}}_{\mathrm{c},k}, q)$, it follows from Lemma 4.2 that $\hat{P}^{\mathrm{da}}_{\mathrm{c},q,k} = \tilde{P}^{\mathrm{da}}_{\mathrm{c},q,k}$, $\hat{P}^{\mathrm{da}}_{\mathrm{c},\overline{q}q,k} = \tilde{P}^{\mathrm{da}}_{\mathrm{c},\overline{q}q,k}$. Therefore, (4.30) implies that $\hat{P}^{\mathrm{da}}_{\mathrm{c},q,k} = P^{\mathrm{da}}_{q,k}, \hat{P}^{\mathrm{da}}_{\mathrm{c},\overline{q}q,k} = P^{\mathrm{da}}_{\overline{q}q,k}$.

Now assume that A_k has the form (4.21). Then (2.7) implies that

$$P^{\mathrm{f}}_{q,k+1} = A_{q,k} P^{\mathrm{da}}_{q,k} A_{q,k}^{\mathrm{T}} + Q_{q,k}, \tag{4.31}$$

$$P^{\mathrm{f}}_{\overline{q}q,k+1} = A_{\overline{q},k} P^{\mathrm{da}}_{\overline{q}q,k} A_{q,k}^{\mathrm{T}} + A_{\overline{q}q,k} P^{\mathrm{da}}_{q,k} A_{q,k}^{\mathrm{T}} + Q_{\overline{q}q,k}, \tag{4.32}$$

where Q_k has entries

$$Q_k = \begin{bmatrix} Q_{q,k} & (Q_{\overline{q}q,k})^{\mathrm{T}} \\ Q_{\overline{q}q,k} & Q_{\overline{q},k} \end{bmatrix}. \tag{4.33}$$

Furthermore, it follows from (4.12), (4.21) and (4.27) that

$$\tilde{P}^{\mathrm{f}}_{\mathrm{c},q,k+1} = A_{q,k}\hat{P}^{\mathrm{da}}_{\mathrm{c},q,k}A^{\mathrm{T}}_{q,k} + Q_{q,k}, \tag{4.34}$$

$$\tilde{P}^{\mathrm{f}}_{\mathrm{c},\overline{q}q,k+1} = A_{\overline{q},k}\hat{P}^{\mathrm{da}}_{\mathrm{c},\overline{q}q,k}A^{\mathrm{T}}_{q,k} + A_{\overline{q}q,k}\hat{P}^{\mathrm{da}}_{\mathrm{c},q,k}A^{\mathrm{T}}_{q,k} + Q_{\overline{q}q,k}. \tag{4.35}$$

Therefore, (4.31), (4.32), (4.34), and (4.35) imply that $\tilde{P}^{\mathrm{f}}_{\mathrm{c},q,k+1} = P^{\mathrm{f}}_{q,k+1}$ and $\tilde{P}^{\mathrm{f}}_{\mathrm{c},\overline{q}q,k+1} = P^{\mathrm{f}}_{\overline{q}q,k+1}$. □

The previous result showed that the Cholesky-based rank-q square-root filter is equivalent to the Kalman filter for only a single time-step. The following result guarantees this equivalence for all time-steps.

Corollary 4.1: Assume that C_k and A_k are of the form (4.19) and (4.21). Let $\tilde{P}^{\mathrm{f}}_{\mathrm{c},q,0} = P^{\mathrm{f}}_{q,0}$, $\tilde{P}^{\mathrm{f}}_{\mathrm{c},\overline{q}q,0} = P^{\mathrm{f}}_{\overline{q}q,0}$, 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 4.1 yields $K_{\mathrm{c},k} = K_k$ for all $k \geqslant 0$. Hence, it follows from (2.5), (2.6), (4.10), and (4.11) that $x^{\mathrm{f}}_{\mathrm{c},k} = x^{\mathrm{f}}_k$ for all $k \geqslant 0$. □

In the following section, we present a result for time-invariant systems that guarantees that the Cholesky-based rank-q filter is equivalent to the Kalman filter for a few time-steps even when the measurement and dynamics map are not constrained according to (4.19) and (4.21)

A. Linear Time-Invariant Systems

Next, we consider linear time-invariant systems 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 $R_k = R$. Next, 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

$$\mathcal{O} \triangleq \begin{bmatrix} C^{\mathrm{T}} & (CA)^{\mathrm{T}} & \cdots & (CA^{n-1})^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \tag{4.36}$$

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

$$\mathcal{O} = \begin{bmatrix} I_n \\ 0_{(p-1)n\times n} \end{bmatrix}. \tag{4.37}$$

Therefore, (4.36) and (4.37) imply that, for every positive integer i such that $ip \leqslant n$,

$$CA^{i-1} = \begin{bmatrix} 0_{p\times p(i-1)} & I_p & 0_{p\times(n-pi)} \end{bmatrix}. \tag{4.38}$$

Next, we present a result that shows that the Cholesky-based rank-q square-root filter is equivalent to the Kalman filter for a specific number of time steps. To do this, we first present the following results.

Lemma 4.3: Let i be a positive integer, and for all $k > 0$, let $\hat{P}_k \in \mathbb{R}^{n\times n}$ satisfy

$$CA^{i-1}\hat{P}_{k+1} = CA^i\hat{P}_kA^{\mathrm{T}} + CA^{i-1}Q - CA^i\hat{P}_kC^{\mathrm{T}}(C\hat{P}_kC + R)^{-1}C\hat{P}_kA^{\mathrm{T}}. \tag{4.39}$$

Assume that $CA^i\hat{P}_k = CA^iP^{\mathrm{f}}_k$ and $C\hat{P}_k = CP^{\mathrm{f}}_k$. Then, $CA^{i-1}\hat{P}_{k+1} = CA^{i-1}P^{\mathrm{f}}_{k+1}$.

Proof. The result follows from (2.4),(2.7), and (4.39). □

Lemma 4.4: Assume that $\hat{P}_k \in \mathbb{R}^{n\times n}$ satisfies (4.39) for all $k > 0$ and $i = 1,\ldots,r$. Let $CA^{i-1}\hat{P}_0 = CA^{i-1}P^{\mathrm{f}}_0$ for $i = 1,\ldots,r$. Then, for all $k = 0,\ldots,r$, $C\hat{P}_k = CP^{\mathrm{f}}_k$.

Proof. The result follows from repeated application of Lemma 4.3. □

Proposition 4.2: Let $r > 0$ be an integer such that $0 < q = pr < n$. Furthermore, assume that $\tilde{P}^{\mathrm{f}}_{\mathrm{c},0} = P^{\mathrm{f}}_0$. Then, for all $k = 0,\ldots,r$, $K_{\mathrm{c},k} = K_k$. If, in addition, $x^{\mathrm{f}}_{\mathrm{c},0} = x^{\mathrm{f}}_0$, then for all $k = 0,\ldots,r$, $x^{\mathrm{f}}_{\mathrm{c},k} = x^{\mathrm{f}}_k$.

Proof. It follows from Lemma 4.2 and (4.38) that, for all $k \geqslant 0$ and $i = 1,\ldots,r$,

$$CA^{i-1}\hat{P}^{\mathrm{f}}_{\mathrm{c},k} = CA^{i-1}\tilde{P}^{\mathrm{f}}_{\mathrm{c},k}, CA^{i-1}\hat{P}^{\mathrm{da}}_{\mathrm{c},k} = CA^{i-1}\tilde{P}^{\mathrm{da}}_{\mathrm{c},k}. \tag{4.40}$$

Note that $\tilde{P}^{\mathrm{f}}_{\mathrm{c},k+1} = A\hat{P}^{\mathrm{da}}_{\mathrm{c},k}A^{\mathrm{T}} + Q$, and therefore

$$CA^{i-1}\tilde{P}^{\mathrm{f}}_{\mathrm{c},k+1} = CA^i\hat{P}^{\mathrm{da}}_{\mathrm{c},k}A^{\mathrm{T}} + CA^{i-1}Q. \tag{4.41}$$

Substituting (4.40) into (4.41) yields

$$CA^{i-1}\hat{P}^{\mathrm{f}}_{\mathrm{c},k+1} = CA^i\tilde{P}^{\mathrm{da}}_{\mathrm{c},k}A^{\mathrm{T}} + CA^{i-1}Q, \tag{4.42}$$

for $i = 1,\ldots,r$. Using (4.9) in (4.42) yields

$$CA^{i-1}\hat{P}^{\mathrm{f}}_{\mathrm{c},k+1} = CA^i\Big[\hat{P}^{\mathrm{f}}_{\mathrm{c},k} + CA^{i-1}Q - \hat{P}^{\mathrm{f}}_{\mathrm{c},k}C^{\mathrm{T}}(C\hat{P}^{\mathrm{f}}_{\mathrm{c},k}C^{\mathrm{T}} + R)^{-1}C\hat{P}^{\mathrm{f}}_{\mathrm{c},k}\Big]A^{\mathrm{T}}, \tag{4.43}$$

for all $k \geqslant 0$ and $i = 1,\ldots,r$. Since $\tilde{P}^{\mathrm{f}}_{\mathrm{c},0} = P^{\mathrm{f}}_0$, it follows from Lemma 4.2 that, for $i = 1,\ldots,r$, $CA^{i-1}\hat{P}^{\mathrm{f}}_{\mathrm{c},0} = CA^{i-1}P^{\mathrm{f}}_0$. Hence, it follows from (4.43) and Lemma 4.4 that, for $k = 0,\ldots,r$, $C\hat{P}^{\mathrm{f}}_k = CP^{\mathrm{f}}_k$. Finally, (2.3) and (4.8) imply that, for $k = 0,\ldots,r$, $K_{\mathrm{c},k} = K_k$, and hence, for all $k = 0,\ldots,r$, $x^{\mathrm{f}}_{\mathrm{c},k} = x^{\mathrm{f}}_k$. □

Hence, the Cholesky-based rank-q square-root filter is equivalent to the Kalman filter for a fixed number of time steps that depend on the rank q of the approximations of the predicted error covariances, as well as the dimension p of the output. However, in general $\tilde{P}^{\mathrm{f}}_{\mathrm{c},k}$ and P^{f}_k are not equal for all $k = 0,\ldots,r$ even though Proposition 4.2 implies that $K_{\mathrm{c},k}$ and K_k are equal. Moreover, $K_{\mathrm{c},k}$ and K_k are generally not equal for $k > r$.

V. Examples

We compare the performance of the SVD-based rank-q square-root filter and the Cholesky-based rank-q square-root filter with the Kalman filter for two linear time-invariant systems.

A. Compartmental Model

Consider n compartments or subsystems exchanging energy through mutual interaction [13]. Applying conservation of energy yields, for $i = 1,\ldots,n$,

$$E_{i,k+1} = (1-\beta)E_{i,k} - \alpha\,(x_{i+1,k} - x_{i,k}) - \alpha\,(x_{i,k} - x_{i-1,k}) + P_{i,k}, \tag{5.1}$$

where $E_{i,k}$ is the energy in the i-th compartment, $P_{i,k}$ is the external disturbance affecting the i-th compartment, $0 < \beta < 1$ is the loss coefficient, and $0 < \alpha < 1$ is the flow coefficient. It follows from (5.1) that

$$x_{k+1} = Ax_k + Gw_k, \tag{5.2}$$

where $P_{i,k} = g_i w_{i,k}$, $x_k \triangleq [E_{1,k} \cdots E_{n,k}]^{\mathrm{T}}$, $w_k \triangleq [w_{1,k} \cdots w_{n,k}]^{\mathrm{T}}$,

$A \in \mathbb{R}^{n\times n}$, and $G \in \mathbb{R}^{n\times n}$ is defined by $G \triangleq \mathrm{diag}(g_1, \ldots, g_n)$. Let $n = 20$, $\alpha = 0.35$ and $\beta = 0.5$. We assume that the disturbance w_k affects all of the compartments so that $g_i \neq 0$ for $i = 1, \ldots, n$. The external disturbance w_k is modeled as a white-noise process with unit covariance. Finally, we use measurements of the energy in the 10th and 11th compartments to estimate the energy in all of the compartments, that is, $y_k = \left[\begin{array}{cc} x_{10,k} & x_{11,k} \end{array} \right]^{\mathrm{T}} + v_k$.

To evaluate the performance of the SVD-based and Cholesky-based reduced-rank square-root filters, we compare the costs J_k, $J_{\mathrm{s},k}$ and $J_{\mathrm{c},k}$, where $J_k \triangleq \mathrm{tr}(P_k^{\mathrm{f}})$, $J_{\mathrm{s},k} = \mathrm{tr}(P_{\mathrm{s},k}^{\mathrm{f}})$, $J_{\mathrm{c},k} = \mathrm{tr}(P_{\mathrm{c},k}^{\mathrm{f}})$. In all cases, we initialize the three filters with $x_0^{\mathrm{f}} = x_{\mathrm{c},0}^{\mathrm{f}} = x_{\mathrm{s},0}^{\mathrm{f}} = 0$ and $P_0^{\mathrm{f}} = \tilde{P}_{\mathrm{c},0}^{\mathrm{f}} = \tilde{P}_{\mathrm{s},0}^{\mathrm{f}} = I_{20}$.

We compare the performance of the SVD-based and Cholesky-based filters for $q = 2, 5, 10$. The steady-state performance of the SVD-based rank-q square-root filter and the Cholesky-based rank-q square-root filter, respectively, is shown in Figure 2. Figure 3 shows the performance of the SVD-based reduced-rank square-root filter $J_{\mathrm{s},k}$ and the Cholesky-based reduced-rank square-root filter $J_{\mathrm{c},k}$, when $q = 2$ in both cases. Finally, we plot $J_{\mathrm{c},k}/J_k$ and $J_{\mathrm{s},k}/J_k$ when $q = 10$. Note that $p = 2$, and hence, $r = 5$ satisfies $q = pr$. Therefore, it follows from Proposition 4.2 that the Cholesky-based rank-q square-root filter is equivalent to the Kalman filter for $k = 0, \ldots, 5$, as confirmed by Figure 4.

B. N-mass system

Next, we consider the mass-spring-damper model shown in Figure 1. For $i = 1, \ldots, 10$, $m_i = 1$ kg, while $k_j = 1$ N/m and $c_j = 0.2$ N-s/m for $j = 1, \ldots, 11$. We assume that an external force $w_{i,k}$ acts on the mass m_i, where $w_{i,k}$ is a white-noise process with unit covariance so that

$$x_{k+1} = Ax_k + w_k, \tag{5.3}$$

where $x \triangleq [q_1 \, \dot{q}_1 \, \cdots \, q_{10} \, \dot{q}_{10}]^{\mathrm{T}}$, $w \triangleq [w_1 \cdots w_{10}]^{\mathrm{T}}$, and $A \in \mathbb{R}^{20\times 20}$ is obtained using a zero-order-hold discretization of the continuous-time dynamics. We assume that the displacement of the 5th mass is measured so that, $y_k = q_{5,k} + v_k$, where v_k is white-noise process with unit covariance. Again, we initialize the Kalman filter and the reduced-rank square-root filters with $x_0^{\mathrm{f}} = x_{\mathrm{c},0}^{\mathrm{f}} = x_{\mathrm{s},0}^{\mathrm{f}} = 0$ and $P_0^{\mathrm{f}} = \tilde{P}_{\mathrm{c},0}^{\mathrm{f}} = \tilde{P}_{\mathrm{s},0}^{\mathrm{f}} = I_{20}$.

We compare the performance of the reduced-rank square-root filters for $q = 4$ and $q = 8$. The mean-squared error (MSE) in the estimates of the position of the masses is shown in Figure 5. It can be seen that, for a specific choice of q, the performance of the Cholesky-based rank-q square-root filter is better than the performance of the SVD-based rank-q square-root filter. The MSE in the estimates of the velocities of the masses is shown in Figure 6.

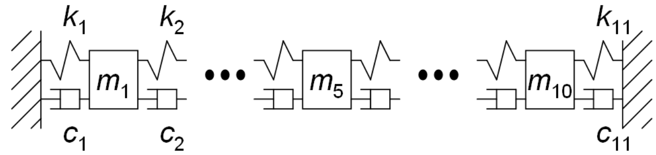

Fig. 1. Mass-spring-dashpot system.

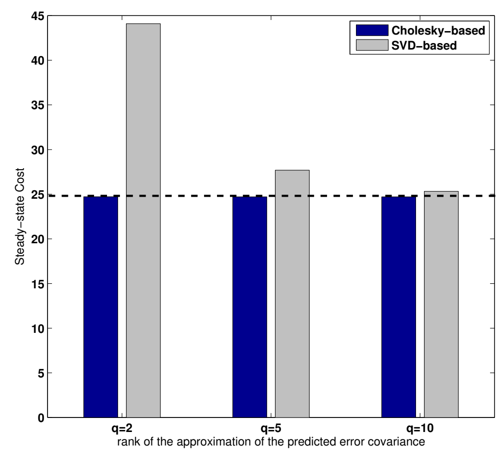

Fig. 2. Steady-state performance of the SVD-based and Cholesky-based reduced-rank filters for $q = 2, 5, 10$. Note that $n = 20$ and even when $q = 2$, the performance of the Cholesky-based reduced-rank square-root filter is similar to that of the Kalman filter. The steady-state performance of the Kalman filter is shown as the dashed line for comparison.

VI. Conclusions

We developed a reduced-rank square-root Kalman filter based on the Cholesky factorization. We presented conditions under which the SVD-based reduced-rank square-root Kalman filter and the Cholesky-based reduced-rank square-root Kalman filter are equivalent to the Kalman filter. In general, neither the Cholesky-based nor SVD-based reduced-rank square-root filter consistently outperforms the other. However, in this paper, we showed two examples where the Cholesky-based reduced-rank square-root filter performs better than the SVD-based reduced-rank square-root filter. Since the Cholesky factorization is a computationally efficient algorithm compared to the singular value decomposition, the Cholesky-based reduced-rank square-root filter provides a computationally efficient alternative method for reduced-rank square-root filtering.

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 and D. C. Hyland, "The Optimal Projection Equations for Reduced-Order State Estimation," *IEEE Trans. Autom. Contr.*, vol. AC-30, pp. 583–585, 1985.

[4] 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.

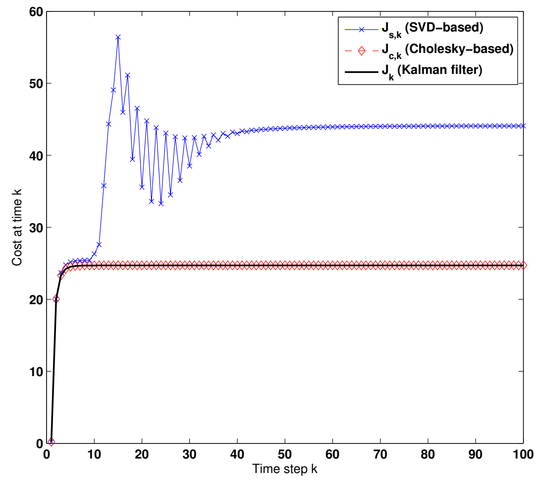

Fig. 3. The costs $J_{\mathrm{s},k}$ and $J_{\mathrm{c},k}$ of the SVD-based and Cholesky-based reduced-rank square-root filters, respectively, with $q = 2$. The performance of the Cholesky-based rank-q square-root filter is close to that of the Kalman filter. However, the performance of the SVD-based filter is poor.

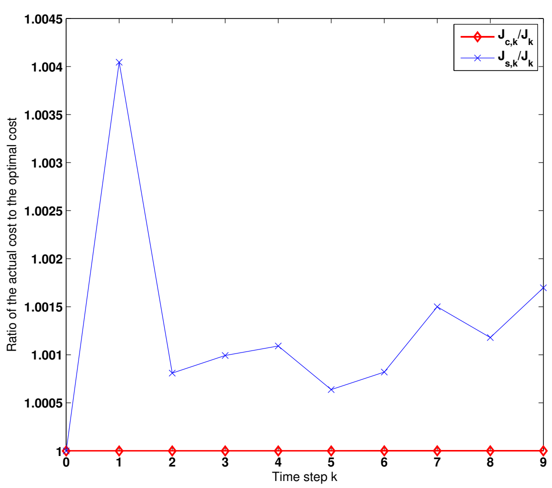

Fig. 4. Ratio of the costs $J_{\mathrm{s},k}$ and $J_{\mathrm{c},k}$ of the reduced-rank filters with $q = 10$ and the Kalman filter. The Cholesky-based rank-q square-root filter is equivalent to the Kalman filter for $k = 0, \ldots, r = 5$. In fact, the performance of the Cholesky-based rank-q square-root filter is close to the performance of the Kalman filter for $k = 0, \ldots, 9$.

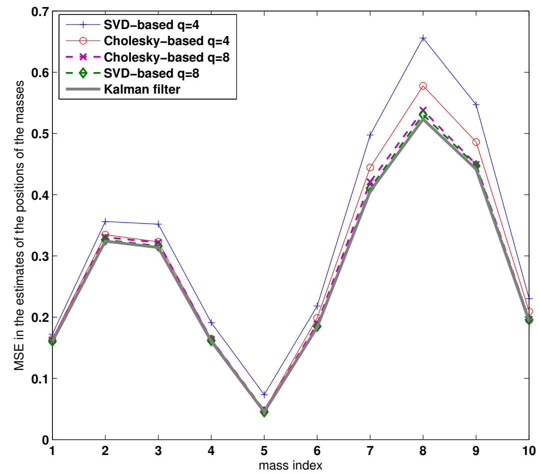

Fig. 5. Steady-state MSE in the estimates of the positions of the masses $m_1, \ldots, m_{10}$ using the Cholesky-based and SVD-based reduced-rank square-root filters for $q = 4$ and $q = 8$. The performance of the reduced-rank square-root filters improves as q increases.

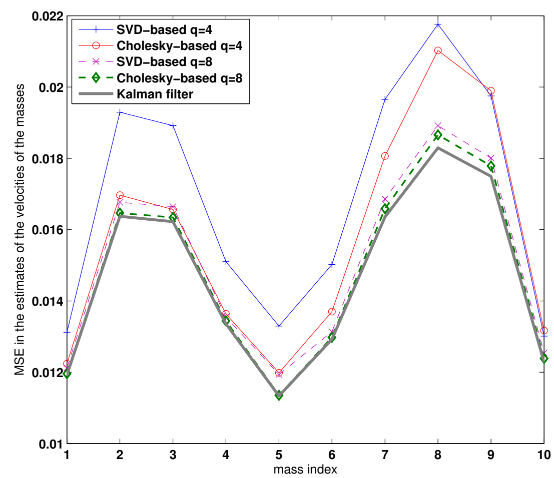

Fig. 6. Steady-state MSE in the estimates of the velocities of the masses $m_1, \ldots, m_{10}$ using the Cholesky-based and SVD-based reduced-rank square-root filters with $q = 4$ and $q = 8$.

[5] 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.

[6] S. F. Farrell and P. J. Ioannou, "State Estimation Using a Reduced-Order Kalman Filter," *J. Atmos. Sci.*, vol. 58, pp. 3666–3680, 2001.

[7] 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," *Proc. Amer. Contr. Conf.*, New York, June 2007.

[8] 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.

[9] 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.

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

[11] 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.

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

[13] D. S. Bernstein and D. C. Hyland, "Compartmental Modeling and Second-Moment Analysis of State Space Systems," *SIAM J. Matrix Anal. Appl.*, vol. 14, no. 3, pp. 880–901, 1993.

[14] D. S. Bernstein, *Matrix Mathematics*, Princeton University Press, 2005.