On the stability bounds of Kalman filters for linear deterministic discrete-time systems

Mark Haring, and Tor Arne Johansen, *Senior Member, IEEE*

***Abstract*—In this note, we prove input-to-state stability of the estimation error of the discrete-time Kalman filter under suitable assumptions. Input-to-state stability is an important prerequisite for the use of many contemporary analysis tools for cascaded and interconnected systems. In this way, this note provides a missing link for the rigorous analysis of systems, of which the Kalman filter is a subsystem, using these tools.**

***Index Terms*—cascaded systems, input-to-state stability, interconnected systems, Kalman filter, Riccati difference equation.**

I. Introduction

We begin this note by repeating the main point in [21], [22] that the Kalman filter, being a stochastically optimal linear filter, can perfectly well be derived using an entirely deterministic least-squares approach, yielding identical solutions. Using a stochastic formulation, the Kalman filter has been shown to produce a minimum-variance state estimate assuming that the system is linear and that all disturbances are Gaussian, white and have known variances [14]. These assumptions are restrictive and unlikely to hold for real-life systems. They are almost farcical if a nonlinear system is transformed into a state-affine time-varying system by a linearizing transformation, as in [4], for example. One of the advantages of a deterministic least-squares formulation is that almost all of these model assumptions can be avoided [21]. As we point out in this note, without unnecessarily restrictive assumptions, modern analysis tools based on input-to-state stability (see [13] for a definition) can be applied to study cascaded and interconnected systems of which the Kalman filter is a subsystem, even if these systems are nonlinear or time-delayed. Examples of such analysis tools are the Lyapunov-based theorems in [18] and the small-gain theorems in [7].

The reason for writing this note is that, so far, input-to-state stability of the estimation error of the Kalman filter has not been proved for discrete-time systems. Being an essential prerequisite, we cannot apply the above-mentioned analysis tools without each subsystem being input-to-state stable. Although the stability of the discrete-time Kalman filter is studied extensively in [8], [12], [16][1], we cannot conclude input-to-state stability from the stochastic analyses in these works. There are two main reasons for this. First, instead of input-to-state stability, these works focus on asymptotic stability of the estimation error by neglecting the influence of the disturbances. Second, the derived bounds on the error covariance, that are essential to the analyses, are based on the argument that the Kalman filter produces a minimum-variance estimate[2]. This argument does not have a direct deterministic equivalent. In [15], different bounds on the error covariance are obtained using deterministic arguments. However, the class of considered systems is somewhat smaller (see Remark 4).

The main contribution of this note is a rigorous proof that the estimation error of the Kalman filter is exponentially input-to-state stable under appropriate assumptions. In order to prove this, we derive deterministic bounds on the error covariance, which may be of separate interest. For example, these bounds may be used to analyse the extended Kalman filter [3], [12] and the arrival cost of the moving-horizon estimator [19], [20], which exploit a similar Riccati difference equation.

We use the following notations in this note. The sets of real numbers, positive real numbers, natural numbers (nonnegative integers) and positive integers are denoted by $\mathbb{R}$, $\mathbb{R}_{>0}$, $\mathbb{N}$ and $\mathbb{N}_{>0}$, respectively. The identity matrix and the zero matrix are written as $\mathbf{I}$ and $\mathbf{0}$, respectively. $\mathbf{M}^T$ denotes the transpose of the matrix $\mathbf{M}$. The Euclidean norm is denoted by $\|\cdot\|$.

II. Deterministic Kalman filter formulation

Consider the following linear discrete-time system:

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}_k\mathbf{x}_k + \mathbf{b}_k + \mathbf{G}_k\mathbf{w}_k \\ \mathbf{y}_k &= \mathbf{C}_k\mathbf{x}_k + \mathbf{d}_k + \mathbf{v}_k, \end{aligned} \tag{1}$$

with state $\mathbf{x}_k \in \mathbb{R}^{n_\mathbf{x}}$, output $\mathbf{y}_k \in \mathbb{R}^{n_\mathbf{y}}$, known vector signals $\mathbf{b}_k \in \mathbb{R}^{n_\mathbf{x}}$ and $\mathbf{d}_k \in \mathbb{R}^{n_\mathbf{y}}$, and unknown deterministic disturbances $\mathbf{w}_k \in \mathbb{R}^{n_\mathbf{w}}$ and $\mathbf{v}_k \in \mathbb{R}^{n_\mathbf{y}}$, where $n_\mathbf{x}, n_\mathbf{y}, n_\mathbf{w} \in \mathbb{N}_{>0}$ are the corresponding dimensions and $k \in \mathbb{N}$ is the time index. We note that the disturbances $\mathbf{w}_k$ and $\mathbf{v}_k$ may represent anything from model mismatch to unknown external influences. Consider the quadratic objective function:

$$\sum_{i=0}^{N-1} \bar{\mathbf{w}}_i^T\mathbf{Q}_i^{-1}\bar{\mathbf{w}}_i + \sum_{j=1}^{N} \bar{\mathbf{v}}_j^T\mathbf{R}_j^{-1}\bar{\mathbf{v}}_j + (\bar{\mathbf{x}}_0-\hat{\mathbf{x}}_0)^T\mathbf{P}_0^{-1}(\bar{\mathbf{x}}_0-\hat{\mathbf{x}}_0), \tag{2}$$

and the constraints:

$$\begin{aligned} \bar{\mathbf{x}}_{i+1} &= \mathbf{A}_i\bar{\mathbf{x}}_i + \mathbf{b}_i + \mathbf{G}_i\bar{\mathbf{w}}_i, & \forall i \in \{0,1,\ldots,N-1\}, \\ \mathbf{y}_j &= \mathbf{C}_j\bar{\mathbf{x}}_j + \mathbf{d}_j + \bar{\mathbf{v}}_j, & \forall j \in \{1,2,\ldots,N\}, \end{aligned} \tag{3}$$

for some $N \in \mathbb{N}$. Here, $\bar{\mathbf{x}}_k$ for all $k \in \{0,1,\ldots,N\}$, $\bar{\mathbf{w}}_i$ for all $i \in \{0,1,\ldots,N-1\}$ and $\bar{\mathbf{v}}_j$ for all $j \in \{1,2,\ldots,N\}$ are optimization variables. We denote these variables by $\{\bar{\mathbf{x}}_k\}_{k=0}^N$, $\{\bar{\mathbf{w}}_i\}_{j=0}^{N-1}$ and $\{\bar{\mathbf{v}}_j\}_{j=1}^N$ for short. The matrices $\mathbf{Q}_i$ and $\mathbf{R}_j$ for $i \in \{0,1,\ldots,N-1\}$ and $j \in \{1,2,\ldots,N\}$ are chosen to be symmetric and positive definite. They serve as weighting matrices for the constraints (3). Note that the constraints (3) are the same as the system equations (1). The last term of the objective function (2) is known as the arrival cost [20]. It consists of the a priori estimate $\hat{\mathbf{x}}_0$ of the initial state of the system and the symmetric, positive definite weighting matrix $\mathbf{P}_0$. The optimization problem

$$\min_{\{\bar{\mathbf{x}}_k\}_{k=0}^N, \{\bar{\mathbf{w}}_i\}_{j=0}^{N-1}, \{\bar{\mathbf{v}}_j\}_{j=1}^N} \quad (2) \text{ subject to } (3) \tag{4}$$

is a weighted least-squares problem. Let the vector $\bar{\mathbf{x}}_N$ that corresponds to the solution of the optimization problem (4) be denoted by $\hat{\mathbf{x}}_N$. The vector $\hat{\mathbf{x}}_N$ can be regarded as an estimate for the state $\mathbf{x}_N$ of system (1). With the help of dynamic programming [6], Pontryagin's

This work is sponsored by the Research Council of Norway with project number 250725, and through the Centres of Excellence funding scheme, project number 223254 - Centre for Autonomous Marine Operations and Systems (NTNU-AMOS).

Mark Haring was with the Centre for Autonomous Marine Operations and Systems, Department of Engineering Cybernetics, NTNU, Norwegian University of Science and Technology, 7491 Trondheim, Norway. He is now with Mathematics and Cybernetics, SINTEF Digital, Trondheim, Norway (e-mail: mark.haring@sintef.no).

Tor Arne Johansen is with the Centre for Autonomous Marine Operations and Systems, Department of Engineering Cybernetics, NTNU, Norwegian University of Science and Technology, 7491 Trondheim, Norway (e-mail: tor.arne.johansen@ntnu.no).

[1]As pointed out in [11], the analyses in [8], [12] contain an error. This error has been corrected in [16].

[2]Noting that the Kalman filter produces a minimum-variance estimate, any upper bound on the error covariance of a suboptimal filter (different from the Kalman filter) is an upper bound on the error covariance of the Kalman filter.

maximum principle [2] or feedback invariance [10], it can be shown that $\hat{\mathbf{x}}_N$ is equal to the output of the recursion

$$\begin{aligned}\hat{\mathbf{x}}_{k+1} = \mathbf{A}_k\hat{\mathbf{x}}_k + \mathbf{b}_k + \mathbf{K}_{k+1}(&\mathbf{y}_{k+1} \\ &- \mathbf{C}_{k+1}(\mathbf{A}_k\hat{\mathbf{x}}_k + \mathbf{b}_k) - \mathbf{d}_{k+1})\end{aligned} \tag{5}$$

for $k = 0, 1, \ldots, N-1$, with

$$\begin{aligned}\mathbf{M}_{k+1} &= \mathbf{A}_k\mathbf{P}_k\mathbf{A}_k^T + \mathbf{G}_k\mathbf{Q}_k\mathbf{G}_k^T, \\ \mathbf{K}_{k+1} &= \mathbf{M}_{k+1}\mathbf{C}_{k+1}^T(\mathbf{C}_{k+1}\mathbf{M}_{k+1}\mathbf{C}_{k+1}^T + \mathbf{R}_{k+1})^{-1}, \\ \mathbf{P}_{k+1} &= (\mathbf{I} - \mathbf{K}_{k+1}\mathbf{C}_{k+1})\mathbf{M}_{k+1}(\mathbf{I} - \mathbf{K}_{k+1}\mathbf{C}_{k+1})^T \\ &\quad + \mathbf{K}_{k+1}\mathbf{R}_{k+1}\mathbf{K}_{k+1}^T,^{3}\end{aligned} \tag{6}$$

and initial conditions $\hat{\mathbf{x}}_0$ and $\mathbf{P}_0$. In concurrence with [21], [22], we note that the recursive formulas in (5) and (6) are the equations of the Kalman filter, which is often derived using a stochastic formulation (see [14] for example) instead of the deterministic formulation in this note.

III. Sufficient conditions for stability

The remainder of this note is dedicated to showing that the estimation error of the Kalman filter is exponentially input-to-state stable under suitable assumptions (see Definition 5 for a definition of exponential input-to-state stability). The following assumptions are suitable in the sense that they provide sufficient conditions for stability.

Assumption 1: There exist constants $c_{\mathbf{A}1}, c_{\mathbf{A}2}, c_{\mathbf{G}}, c_{\mathbf{C}} \in \mathbb{R}_{>0}$ such that

$$c_{\mathbf{A}1}\mathbf{I} \preceq \mathbf{A}_k\mathbf{A}_k^T \preceq c_{\mathbf{A}2}\mathbf{I}, \quad \mathbf{G}_k\mathbf{G}_k^T \preceq c_{\mathbf{G}}\mathbf{I}, \quad \mathbf{C}_k^T\mathbf{C}_k \preceq c_{\mathbf{C}}\mathbf{I} \tag{7}$$

for all $k \geq 0$.

Let the state transition matrix of the system be given by

$$\mathbf{\Phi}(k, k_0) = \mathbf{A}_{k-1}\mathbf{A}_{k-2}\cdots\mathbf{A}_{k_0}, \quad \forall k > k_0, \quad \mathbf{\Phi}(k_0, k_0) = \mathbf{I} \tag{8}$$

for all $k \geq k_0 \geq 0$. Consider the controllability Gramian

$$\mathbf{W}_c(k, k_0) = \sum_{i=k_0}^{k-1} \mathbf{\Phi}(k, i+1)\mathbf{G}_i\mathbf{G}_i^T\mathbf{\Phi}^T(k, i+1) \tag{9}$$

and the observability Gramian

$$\mathbf{W}_o(k, k_0) = \sum_{i=k_0+1}^{k} \mathbf{\Phi}^T(i, k_0)\mathbf{C}_i^T\mathbf{C}_i\mathbf{\Phi}(i, k_0) \tag{10}$$

for $k \geq k_0 \geq 0$. We make the following assumptions with respect to the controllability and observability of the system.

Assumption 2: System (1) is **uniformly controllable**. That is, there exist constants $c_c \in \mathbb{R}_{>0}$ and $N_c \in \mathbb{N}_{>0}$ such that

$$\mathbf{W}_c(k + N_c, k) \succeq c_c\mathbf{I} \tag{11}$$

for all $k \geq 0$.

Assumption 3: System (1) is **uniformly observable**. That is, there exist constants $c_o \in \mathbb{R}_{>0}$ and $N_o \in \mathbb{N}_{>0}$ such that

$$\mathbf{W}_o(k + N_o, k) \succeq c_o\mathbf{\Phi}^T(k + N_o, k)\mathbf{\Phi}(k + N_o, k) \tag{12}$$

for all $k \geq 0$.

The presented assumptions are similar to those in [8], [12], [16]. We note that these assumptions may possibly be weakened by allowing the state transition matrix to be singular [17], or by assuming uniform detectability and uniform stabilizability instead of uniform observability and uniform controllability [1], [9].

Remark 4: In [15], linear systems are considered for which $\mathbf{G}_k = \mathbf{I}$ for all $k \geq 0$. We note that any system (1) for which $\mathbf{G}_k = \mathbf{I}$ for all $k \geq 0$ is immediately uniformly controllably. In this work, we consider a slightly larger class of systems by letting $\mathbf{G}_k$ differ from the identity matrix.

In addition to the previous assumptions, without loss of generality, we assume that the weighting matrices $\mathbf{Q}_k$ and $\mathbf{R}_{k+1}$ in (6) (and (2)) are chosen such that

$$\begin{aligned}c_{\mathbf{Q}1}\mathbf{I} &\preceq \mathbf{Q}_k \preceq c_{\mathbf{Q}_2}\mathbf{I}, \\ c_{\mathbf{R}1}\mathbf{I} &\preceq \mathbf{R}_{k+1} \preceq c_{\mathbf{R}_2}\mathbf{I}\end{aligned} \tag{13}$$

for all $k \geq 0$ and some constants $c_{\mathbf{Q}1}, c_{\mathbf{Q}2}, c_{\mathbf{R}1}, c_{\mathbf{R}2} \in \mathbb{R}_{>0}$.

IV. Input-to-state stability of the estimation error

Let the estimation error of the deterministic Kalman filter be given by

$$\mathbf{e}_k = \hat{\mathbf{x}}_k - \mathbf{x}_k. \tag{14}$$

From (1) and (5), it follows that the dynamics of the estimation error are governed by

$$\mathbf{e}_{k+1} = (\mathbf{I} - \mathbf{K}_{k+1}\mathbf{C}_{k+1})(\mathbf{A}_k\mathbf{e}_k - \mathbf{G}_k\mathbf{w}_k) + \mathbf{K}_{k+1}\mathbf{v}_{k+1}. \tag{15}$$

In line with the definition of input-to-state stability in [13], we give the following definition of exponential input-to-state stability.

Definition 5: The estimation error $\mathbf{e}_k$ in (14) is said to be **exponentially input-to-state stable** if there exist constants $c_{\mathbf{e}}, \rho_{\mathbf{e}}, c_{\mathbf{w}}, c_{\mathbf{v}} \in \mathbb{R}_{>0}$, with $\rho_{\mathbf{e}} < 1$, such that

$$\begin{aligned}\|\mathbf{e}_k\| \leq \max\Big\{&c_{\mathbf{e}}\rho_{\mathbf{e}}^{k-k_0}\|\mathbf{e}_{k_0}\|,\ c_{\mathbf{w}}\max_{k_0\leq i\leq k-1}\|\mathbf{w}_i\|, \\ &c_{\mathbf{v}}\max_{k_0\leq j\leq k-1}\|\mathbf{v}_{j+1}\|\Big\}\end{aligned} \tag{16}$$

for all $k \geq k_0 \geq 0$.

We will use a Lyapunov approach to prove exponential input-to-state stability of the estimation error using

$$V_k = \mathbf{e}_k^T\mathbf{P}_k^{-1}\mathbf{e}_k \tag{17}$$

as a candidate function.

A. Bounds on the error covariance matrix

For V_k to be a Lyapunov function, we require that there exist lower and upper bounds on the matrix $\mathbf{P}_k$. We note that $\mathbf{P}_k$ is the error covariance matrix in a stochastic setting; see [14]. Bounds on $\mathbf{P}_k$ are provided in Lemmas 6 and 7. To prove the lower and upper bounds on $\mathbf{P}_k$ in Lemmas 6 and 7, respectively, we frequently use the equation

$$\mathbf{P}_{k+1}^{-1} = \mathbf{M}_{k+1}^{-1} + \mathbf{C}_{k+1}^T\mathbf{R}_{k+1}^{-1}\mathbf{C}_{k+1}, \tag{18}$$

which follows directly from the last equation in (6) and the matrix inversion lemma.[4]

Lemma 6: Under Assumptions 1 and 2, there exists a constant $c_{\mathbf{P}1} \in \mathbb{R}_{>0}$ such that

$$\mathbf{P}_k \succeq c_{\mathbf{P}1}\mathbf{I} \tag{19}$$

for all $k \geq N_c$.

[3]The formula for $\mathbf{P}_{k+1}$ in (6) is equivalent to the more compact formula $\mathbf{P}_{k+1} = (\mathbf{I} - \mathbf{K}_{k+1}\mathbf{C}_{k+1})\mathbf{M}_{k+1}$. Compared to this compact form, the formula in (6) often leads to a more robust implementation in case of computational errors; see [5] for example.

[4]The invertibility of $\mathbf{P}_k$ and $\mathbf{M}_{k+1}$ for any $k \geq 0$ follows from (6), the bounds in (13) and Assumption 1, and from the positive definiteness of the initial condition $\mathbf{P}_0$.

Proof: Let $\mathbf{L}_{k+1}$ be an arbitrary matrix. From (18), it follows that

$$\begin{aligned}&\left(\mathbf{P}_{k+1}+\mathbf{L}_{k+1}\mathbf{L}_{k+1}^T\right)^{-1}\\&=\left(\left(\mathbf{M}_{k+1}^{-1}+\mathbf{C}_{k+1}^T\mathbf{R}_{k+1}^{-1}\mathbf{C}_{k+1}\right)^{-1}+\mathbf{L}_{k+1}\mathbf{L}_{k+1}^T\right)^{-1}.\end{aligned} \tag{20}$$

By applying the bounds in (13) and Assumption 1, we get the first inequality in (21), with $\beta_c = \frac{c_{\mathbf{C}}}{c_{\mathbf{R}1}}$. The three equalities in (21) are obtained using the matrix inversion lemma (thrice). The last inequality in (21) follows from Young's inequality.[5] It follows from (21) that

$$\begin{aligned}&\left(\mathbf{P}_{k+1}+\mathbf{L}_{k+1}\mathbf{L}_{k+1}^T\right)^{-1}\\&\preceq 2\left(\mathbf{M}_{k+1}+\alpha_{c,k+1}\mathbf{L}_{k+1}\mathbf{L}_{k+1}^T\right)^{-1}+2\beta_c\mathbf{I},^{6}\end{aligned} \tag{22}$$

with $\alpha_{c,k+1} = \|\mathbf{I}+\beta_c\mathbf{L}_{k+1}^T\mathbf{L}_{k+1}\|^{-1}$. Using the expression for $\mathbf{M}_{k+1}$ in (6), we obtain from (22) that

$$\begin{aligned}\left(\mathbf{P}_{k+1}+\mathbf{L}_{k+1}\mathbf{L}_{k+1}^T\right)^{-1}&\preceq 2\beta_c\mathbf{I}+2\mathbf{A}_k^{-T}\big(\mathbf{P}_k+\mathbf{A}_k^{-1}\\&\times\left(\mathbf{G}_k\mathbf{Q}_k\mathbf{G}_k^T+\alpha_{c,k+1}\mathbf{L}_{k+1}\mathbf{L}_{k+1}^T\right)\mathbf{A}_k^{-T}\big)^{-1}\mathbf{A}_k^{-1}.\end{aligned} \tag{23}$$

Now, consider any $k \geq 0$. Define in a recursive manner

$$\mathbf{L}_j\mathbf{L}_j^T = \mathbf{A}_j^{-1}\left(\mathbf{G}_j\mathbf{Q}_j\mathbf{G}_j^T+\alpha_{c,j+1}\mathbf{L}_{j+1}\mathbf{L}_{j+1}^T\right)\mathbf{A}_j^{-T} \tag{24}$$

for $j \in \{k, k+1, \dots, k+N_c-1\}$, with $\mathbf{L}_{k+N_c}\mathbf{L}_{k+N_c}^T = \mathbf{0}$. By substituting (24) in (23) and recursively applying the resulting inequality, we obtain

$$\begin{aligned}\mathbf{P}_{k+N_c}^{-1}&\preceq 2^{N_c}\mathbf{\Phi}^{-T}(k+N_c,k)\left(\mathbf{P}_k+\mathbf{L}_k\mathbf{L}_k^T\right)^{-1}\\&\times\mathbf{\Phi}^{-1}(k+N_c,k)+2\beta_c\sum_{i=0}^{N_c-1}2^i\\&\times\mathbf{\Phi}^{-T}(k+N_c,k+N_c-i)\mathbf{\Phi}^{-1}(k+N_c,k+N_c-i).\end{aligned} \tag{25}$$

From (13), (9) and (24), it follows that

$$\mathbf{L}_k\mathbf{L}_k^T \succeq \gamma_c\mathbf{\Phi}^{-1}(k+N_c,k)\mathbf{W}_c(k+N_c,k)\mathbf{\Phi}^{-T}(k+N_c,k), \tag{26}$$

with $\gamma_c = c_{\mathbf{Q}1}\prod_{j=k}^{k+N_c-1}\alpha_{c,j+1}$. Combining (25), (26) and the bounds in Assumptions 1 and 2 gives

$$\mathbf{P}_{k+N_c} \succeq \left(\frac{2^{N_c}}{\gamma_c c_c}+2\beta_c\sum_{i=0}^{N_c-1}\left(\frac{2}{c_{\mathbf{A}1}}\right)^i\right)^{-1}\mathbf{I}, \tag{27}$$

which completes the proof of the lemma. ∎

Lemma 7: Under Assumptions 1 and 3, there exists a constant $c_{\mathbf{P}2} \in \mathbb{R}_{>0}$ such that

$$\mathbf{P}_k \preceq c_{\mathbf{P}2}\mathbf{I} \tag{28}$$

for all $k \geq N_o$.

[5]Note that, for any positive semidefinite matrices $\mathbf{W}$ and $\mathbf{V}$, the inequality $\mathbf{W} \preceq \mathbf{V}$ implies $\mathbf{s}^T\mathbf{W}\mathbf{s} \leq \mathbf{s}^T\mathbf{V}\mathbf{s}$ for all real vectors $\mathbf{s}$. Define $\mathbf{p}(\mathbf{s}) = (\mathbf{M}_{k+1}+\mathbf{L}_{k+1}(\mathbf{I}+\beta_c\mathbf{L}_{k+1}^T\mathbf{L}_{k+1})^{-1}\mathbf{L}_{k+1}^T)^{-\frac{1}{2}}\mathbf{s}$ and $\mathbf{q}(\mathbf{s}) = -\beta_c(\mathbf{M}_{k+1}+\mathbf{L}_{k+1}(\mathbf{I}+\beta_c\mathbf{L}_{k+1}^T\mathbf{L}_{k+1})^{-1}\mathbf{L}_{k+1}^T)^{-\frac{1}{2}}\mathbf{L}_{k+1}(\mathbf{I}+\beta_c\mathbf{L}_{k+1}^T\mathbf{L}_{k+1})^{-1}\mathbf{L}_{k+1}^T\mathbf{s}$. Noting that, for all real vectors $\mathbf{s}$, we have $(\mathbf{p}(\mathbf{s})+\mathbf{q}(\mathbf{s}))^T(\mathbf{p}(\mathbf{s})+\mathbf{q}(\mathbf{s})) \leq 2\mathbf{p}^T(\mathbf{s})\mathbf{p}(\mathbf{s})+2\mathbf{q}^T(\mathbf{s})\mathbf{q}(\mathbf{s})$ using Young's inequality, we obtain the last inequality in (21).

[6]Note that $\beta_c\mathbf{L}_{k+1}(\mathbf{I}+\beta_c\mathbf{L}_{k+1}^T\mathbf{L}_{k+1})^{-1}\mathbf{L}_{k+1}^T(\mathbf{M}_{k+1}+\mathbf{L}_{k+1}(\mathbf{I}+\beta_c\mathbf{L}_{k+1}^T\mathbf{L}_{k+1})^{-1}\mathbf{L}_{k+1}^T)^{-1}\mathbf{L}_{k+1}(\mathbf{I}+\beta_c\mathbf{L}_{k+1}^T\mathbf{L}_{k+1})^{-1}\mathbf{L}_{k+1}^T \preceq \beta_c\mathbf{L}_{k+1}(\mathbf{I}+\beta_c\mathbf{L}_{k+1}^T\mathbf{L}_{k+1})^{-1}\mathbf{L}_{k+1}^T \preceq \mathbf{I}$ leads to the inequality in (22).

Proof: We prove Lemma 7 in a similar way to Lemma 6. Let $\mathbf{Y}_{k+1}$ be an arbitrary matrix. From (6) and (18), we have that

$$\begin{aligned}&\left(\mathbf{P}_{k+1}^{-1}+\mathbf{Y}_{k+1}^T\mathbf{Y}_{k+1}\right)^{-1}\\&=\left(\left(\mathbf{A}_k\mathbf{P}_k\mathbf{A}_k^T+\mathbf{G}_k\mathbf{Q}_k\mathbf{G}_k^T\right)^{-1}+\mathbf{Z}_{k+1}^T\mathbf{Z}_{k+1}\right)^{-1},\end{aligned} \tag{29}$$

where the matrix $\mathbf{Z}_{k+1}$ is defined such that

$$\mathbf{Z}_{k+1}^T\mathbf{Z}_{k+1} = \mathbf{C}_{k+1}^T\mathbf{R}_{k+1}^{-1}\mathbf{C}_{k+1}+\mathbf{Y}_{k+1}^T\mathbf{Y}_{k+1}. \tag{30}$$

The first inequality in (31) follows from (29) and the bounds in (13) and Assumption 1, with $\beta_o = c_{\mathbf{G}}c_{\mathbf{Q}2}$. The three subsequent equalities in (31) are obtained using the matrix inversion lemma, after which the last inequality in (31) follows from Young's inequality. We obtain from (31) that

$$\begin{aligned}&\left(\mathbf{P}_{k+1}^{-1}+\mathbf{Y}_{k+1}^T\mathbf{Y}_{k+1}\right)^{-1}\\&\preceq 2\mathbf{A}_k\left(\mathbf{P}_k^{-1}+\alpha_{o,k+1}\mathbf{A}_k^T\mathbf{Z}_{k+1}^T\mathbf{Z}_{k+1}\mathbf{A}_k\right)^{-1}\mathbf{A}_k^T+2\beta_o\mathbf{I},\end{aligned} \tag{32}$$

with $\alpha_{o,k+1} = \|\mathbf{I}+\beta_o\mathbf{Z}_{k+1}\mathbf{Z}_{k+1}^T\|^{-1}$. Now, considering any $k \geq 0$, define $\mathbf{Y}_j$ such that

$$\mathbf{Y}_j^T\mathbf{Y}_j = \alpha_{o,j+1}\mathbf{A}_j^T\mathbf{Z}_{j+1}^T\mathbf{Z}_{j+1}\mathbf{A}_j \tag{33}$$

for $j \in \{k, k+1, \dots, k+N_o-1\}$ and $\mathbf{Y}_{k+N_o}^T\mathbf{Y}_{k+N_o} = \mathbf{0}$. We recursively obtain from (32) and (33) that

$$\begin{aligned}\mathbf{P}_{k+N_o}&\preceq 2^{N_o}\mathbf{\Phi}(k+N_o,k)\left(\mathbf{P}_k^{-1}+\mathbf{Y}_k^T\mathbf{Y}_k\right)^{-1}\\&\times\mathbf{\Phi}^T(k+N_o,k)+2\beta_o\sum_{i=0}^{N_o-1}2^i\\&\times\mathbf{\Phi}(k+N_o,k+N_o-i)\mathbf{\Phi}^T(k+N_o,k+N_o-i).\end{aligned} \tag{34}$$

From (13), (10), (30) and (33), it follows that

$$\mathbf{Y}_k^T\mathbf{Y}_k \succeq \gamma_o\mathbf{W}_o(k+N_o,k), \tag{35}$$

with $\gamma_o = \frac{1}{c_{\mathbf{R}2}}\prod_{j=k}^{k+N_o-1}\alpha_{o,j+1}$. Combining (34), (35) and the bounds in Assumptions 1 and 3 leads to

$$\mathbf{P}_{k+N_o} \preceq \left(\frac{2^{N_o}}{\gamma_o c_o}+2\beta_o\sum_{i=0}^{N_o-1}(2c_{\mathbf{A}2})^i\right)\mathbf{I}. \tag{36}$$

The lemma follows directly from (36). ∎

The bounds on $\mathbf{P}_k$ in Lemmas 6 and 7 are independent of the initial condition $\mathbf{P}_0$. We note that the $\mathbf{P}_0$ can be chosen such that such that the bounds in Lemmas 6 and 7 hold for all $k \geq 0$. It is important to note that the lower bound in Lemma 6 and the upper bound in Lemma 7 are derived independently of each other; the lower bound on $\mathbf{P}_k$ does not depend on the observability of the system, and the upper bound on $\mathbf{P}_k$ is independent of the controllability of the system. By contrast, the bounds on $\mathbf{P}_k$ in [8], [12], [15], [16] are not derived independently.

B. Main result

We are now ready to present our main result. The result in Theorem 8 is obtained by a Lyapunov analysis using the candidate function in (17) and the bounds on $\mathbf{P}_k$ in Section IV-A.

Theorem 8: Under Assumptions 1-3, there exist constants $c_{\mathbf{e}}, \rho_{\mathbf{e}}, c_{\mathbf{w}}, c_{\mathbf{v}} \in \mathbb{R}_{>0}$, with $\rho_{\mathbf{e}} < 1$, such that

$$\begin{aligned}\|\mathbf{e}_k\| \leq \max\Big\{&c_{\mathbf{e}}\rho_{\mathbf{e}}^{k-k_0}\|\mathbf{e}_{k_0}\|,\ c_{\mathbf{w}}\max_{k_0\leq i\leq k-1}\|\mathbf{w}_i\|,\\&c_{\mathbf{v}}\max_{k_0\leq j\leq k-1}\|\mathbf{v}_{j+1}\|\Big\}\end{aligned} \tag{37}$$

$$
\begin{aligned}
&\left(\mathbf{P}_{k+1}+\mathbf{L}_{k+1}\mathbf{L}_{k+1}^T\right)^{-1} \preceq \left(\left(\mathbf{M}_{k+1}^{-1}+\beta_c\mathbf{I}\right)^{-1}+\mathbf{L}_{k+1}\mathbf{L}_{k+1}^T\right)^{-1} = \left(\frac{1}{\beta_c}\mathbf{I}-\frac{1}{\beta_c^2}\left(\mathbf{M}_{k+1}+\frac{1}{\beta_c}\mathbf{I}\right)^{-1}+\mathbf{L}_{k+1}\mathbf{L}_{k+1}^T\right)^{-1}\\
&= \frac{1}{\beta_c^2}\left(\frac{1}{\beta_c}\mathbf{I}+\mathbf{L}_{k+1}\mathbf{L}_{k+1}^T\right)^{-1}\left(\mathbf{M}_{k+1}+\frac{1}{\beta_c}\mathbf{I}-\frac{1}{\beta_c^2}\left(\frac{1}{\beta_c}\mathbf{I}+\mathbf{L}_{k+1}\mathbf{L}_{k+1}^T\right)^{-1}\right)^{-1}\left(\frac{1}{\beta_c}\mathbf{I}+\mathbf{L}_{k+1}\mathbf{L}_{k+1}^T\right)^{-1}\\
&\quad+\left(\frac{1}{\beta_c}\mathbf{I}+\mathbf{L}_{k+1}\mathbf{L}_{k+1}^T\right)^{-1}\\
&= \frac{1}{\beta_c^2}\left(\beta_c\mathbf{I}-\beta_c^2\mathbf{L}_{k+1}\left(\mathbf{I}+\beta_c\mathbf{L}_{k+1}^T\mathbf{L}_{k+1}\right)^{-1}\mathbf{L}_{k+1}^T\right)\left(\mathbf{M}_{k+1}+\mathbf{L}_{k+1}\left(\mathbf{I}+\beta_c\mathbf{L}_{k+1}^T\mathbf{L}_{k+1}\right)^{-1}\mathbf{L}_{k+1}^T\right)^{-1}\\
&\quad\times\left(\beta_c\mathbf{I}-\beta_c^2\mathbf{L}_{k+1}\left(\mathbf{I}+\beta_c\mathbf{L}_{k+1}^T\mathbf{L}_{k+1}\right)^{-1}\mathbf{L}_{k+1}^T\right)+\beta_c\mathbf{I}-\beta_c^2\mathbf{L}_{k+1}\left(\mathbf{I}+\beta_c\mathbf{L}_{k+1}^T\mathbf{L}_{k+1}\right)^{-1}\mathbf{L}_{k+1}^T\\
&\preceq 2\beta_c^2\mathbf{L}_{k+1}\left(\mathbf{I}+\beta_c\mathbf{L}_{k+1}^T\mathbf{L}_{k+1}\right)^{-1}\mathbf{L}_{k+1}^T\left(\mathbf{M}_{k+1}+\mathbf{L}_{k+1}\left(\mathbf{I}+\beta_c\mathbf{L}_{k+1}^T\mathbf{L}_{k+1}\right)^{-1}\mathbf{L}_{k+1}^T\right)^{-1}\mathbf{L}_{k+1}\left(\mathbf{I}+\beta_c\mathbf{L}_{k+1}^T\mathbf{L}_{k+1}\right)^{-1}\mathbf{L}_{k+1}^T\\
&\quad+2\left(\mathbf{M}_{k+1}+\mathbf{L}_{k+1}\left(\mathbf{I}+\beta_c\mathbf{L}_{k+1}^T\mathbf{L}_{k+1}\right)^{-1}\mathbf{L}_{k+1}^T\right)^{-1}+\beta_c\mathbf{I}-\beta_c^2\mathbf{L}_{k+1}\left(\mathbf{I}+\beta_c\mathbf{L}_{k+1}^T\mathbf{L}_{k+1}\right)^{-1}\mathbf{L}_{k+1}^T
\end{aligned}
\tag{21}
$$

$$
\begin{aligned}
&\left(\mathbf{P}_{k+1}^{-1}+\mathbf{Y}_{k+1}^T\mathbf{Y}_{k+1}\right)^{-1} \preceq \left(\left(\mathbf{A}_k\mathbf{P}_k\mathbf{A}_k^T+\beta_o\mathbf{I}\right)^{-1}+\mathbf{Z}_{k+1}^T\mathbf{Z}_{k+1}\right)^{-1}\\
&= \left(\frac{1}{\beta_o}\mathbf{I}-\frac{1}{\beta_o^2}\mathbf{A}_k\left(\mathbf{P}_k^{-1}+\frac{1}{\beta_o}\mathbf{A}_k^T\mathbf{A}_k\right)^{-1}\mathbf{A}_k^T+\mathbf{Z}_{k+1}^T\mathbf{Z}_{k+1}\right)^{-1}\\
&= \frac{1}{\beta_o^2}\left(\frac{1}{\beta_o}\mathbf{I}+\mathbf{Z}_{k+1}^T\mathbf{Z}_{k+1}\right)^{-1}\mathbf{A}_k\left(\mathbf{P}_k^{-1}+\frac{1}{\beta_o}\mathbf{A}_k^T\mathbf{A}_k-\frac{1}{\beta_o^2}\mathbf{A}_k^T\left(\frac{1}{\beta_o}\mathbf{I}+\mathbf{Z}_{k+1}^T\mathbf{Z}_{k+1}\right)^{-1}\mathbf{A}_k\right)^{-1}\mathbf{A}_k^T\\
&\quad\times\left(\frac{1}{\beta_o}\mathbf{I}+\mathbf{Z}_{k+1}^T\mathbf{Z}_{k+1}\right)^{-1}+\left(\frac{1}{\beta_o}\mathbf{I}+\mathbf{Z}_{k+1}^T\mathbf{Z}_{k+1}\right)^{-1}\\
&= \frac{1}{\beta_o^2}\left(\beta_o\mathbf{I}-\beta_o^2\mathbf{Z}_{k+1}^T\left(\mathbf{I}+\beta_o\mathbf{Z}_{k+1}\mathbf{Z}_{k+1}^T\right)^{-1}\mathbf{Z}_{k+1}\right)\mathbf{A}_k\left(\mathbf{P}_k^{-1}+\mathbf{A}_k^T\mathbf{Z}_{k+1}^T\left(\mathbf{I}+\beta_o\mathbf{Z}_{k+1}\mathbf{Z}_{k+1}^T\right)^{-1}\mathbf{Z}_{k+1}\mathbf{A}_k\right)^{-1}\mathbf{A}_k^T\\
&\quad\times\left(\beta_o\mathbf{I}-\beta_o^2\mathbf{Z}_{k+1}^T\left(\mathbf{I}+\beta_o\mathbf{Z}_{k+1}\mathbf{Z}_{k+1}^T\right)^{-1}\mathbf{Z}_{k+1}\right)+\beta_o\mathbf{I}-\beta_o^2\mathbf{Z}_{k+1}^T\left(\mathbf{I}+\beta_o\mathbf{Z}_{k+1}\mathbf{Z}_{k+1}^T\right)^{-1}\mathbf{Z}_{k+1}\\
&\preceq 2\beta_o^2\mathbf{Z}_{k+1}^T\left(\mathbf{I}+\beta_o\mathbf{Z}_{k+1}\mathbf{Z}_{k+1}^T\right)^{-1}\mathbf{Z}_{k+1}\mathbf{A}_k\left(\mathbf{P}_k^{-1}+\mathbf{A}_k^T\mathbf{Z}_{k+1}^T\left(\mathbf{I}+\beta_o\mathbf{Z}_{k+1}\mathbf{Z}_{k+1}^T\right)^{-1}\mathbf{Z}_{k+1}\mathbf{A}_k\right)^{-1}\mathbf{A}_k^T\mathbf{Z}_{k+1}^T\\
&\quad\times\left(\mathbf{I}+\beta_o\mathbf{Z}_{k+1}\mathbf{Z}_{k+1}^T\right)^{-1}\mathbf{Z}_{k+1}+2\mathbf{A}_k\left(\mathbf{P}_k^{-1}+\mathbf{A}_k^T\mathbf{Z}_{k+1}^T\left(\mathbf{I}+\beta_o\mathbf{Z}_{k+1}\mathbf{Z}_{k+1}^T\right)^{-1}\mathbf{Z}_{k+1}\mathbf{A}_k\right)^{-1}\mathbf{A}_k^T+\beta_o\mathbf{I}\\
&\quad-\beta_o^2\mathbf{Z}_{k+1}^T\left(\mathbf{I}+\beta_o\mathbf{Z}_{k+1}\mathbf{Z}_{k+1}^T\right)^{-1}\mathbf{Z}_{k+1}
\end{aligned}
\tag{31}
$$

for all $k \geq k_0 \geq \max\{N_c, N_o\}$.

Proof: Consider any $k \geq \max\{N_c, N_o\}$. Let $\mathbf{S}_{k+1}$ be an arbitrary matrix. We obtain the equality in (38) from (6), (15), (18) and the matrix inversion lemma. Subsequently, the first inequality in (38) follows from Young's inequality. The second inequality is obtained by applying the matrix inversion lemma, with $\alpha_{t,k+1} = \|\mathbf{I}+\frac{1}{c_{\mathbf{P}1}}\mathbf{S}_{k+1}^T\mathbf{S}_{k+1}\|^{-1}$ (and the help of (18) and Lemma 6). The third inequality in (38) follows from (6) and Young's inequality, where $\varepsilon_t \in \mathbb{R}_{>0}$ is an arbitrary constant. From (13) and (38), it follows that

$$
\mathbf{e}_{k+1}^T\mathbf{P}_{k+1}^{-1}\mathbf{e}_{k+1} \leq (1+\varepsilon_t)\,\mathbf{e}_k^T\mathbf{P}_k^{-1}\mathbf{e}_k+\frac{1+\varepsilon_t}{\varepsilon_t c_{\mathbf{Q}1}}\|\mathbf{w}_k\|^2 + \frac{2}{c_{\mathbf{R}1}}\|\mathbf{v}_{k+1}\|^2. \tag{39}
$$

Also, by defining

$$
\mathbf{S}_k\mathbf{S}_k^T = \mathbf{A}_k^{-1}\left(\mathbf{G}_k\mathbf{Q}_k\mathbf{G}_k^T+\mathbf{S}_{k+1}\mathbf{S}_{k+1}^T\right)\mathbf{A}^{-T}, \tag{40}
$$

we get that

$$
\begin{aligned}
\mathbf{e}_{k+1}^T\left(\mathbf{P}_{k+1}+\mathbf{S}_{k+1}\mathbf{S}_{k+1}^T\right)^{-1}\mathbf{e}_{k+1} &\leq \frac{1+\varepsilon_t}{\varepsilon_t c_{\mathbf{Q}1}}\|\mathbf{w}_k\|^2\\
&+\left(1-\frac{\alpha_{t,k+1}}{2}\right)(1+\varepsilon_t)\,\mathbf{e}_k^T\mathbf{P}_k^{-1}\mathbf{e}_k+\frac{2}{c_{\mathbf{R}1}}\|\mathbf{v}_{k+1}\|^2\\
&+\frac{\alpha_{t,k+1}}{2}(1+\varepsilon_t)\,\mathbf{e}_k^T\left(\mathbf{P}_k+\mathbf{S}_k\mathbf{S}_k^T\right)^{-1}\mathbf{e}_k.
\end{aligned}
\tag{41}
$$

By letting $\mathbf{S}_{k+N_c}\mathbf{S}_{k+N_c}^T = \mathbf{0}$, we obtain recursively from (41) that

$$
\begin{aligned}
\mathbf{e}_{k+N_c}^T\mathbf{P}_{k+N_c}^{-1}\mathbf{e}_{k+N_c} &\leq (1-\gamma_t)(1+\varepsilon_t)^{N_c}\,\mathbf{e}_k^T\mathbf{P}_k^{-1}\mathbf{e}_k\\
&+\gamma_t(1+\varepsilon_t)^{N_c}\,\mathbf{e}_k^T\left(\mathbf{P}_k+\mathbf{S}_k\mathbf{S}_k^T\right)^{-1}\mathbf{e}_k+\frac{(1+\varepsilon_t)^{N_c}}{\varepsilon_t c_{\mathbf{Q}1}}\\
&\times\sum_{i=k}^{k+N_c-1}\|\mathbf{w}_i\|^2+\frac{2(1+\varepsilon_t)^{N_c-1}}{c_{\mathbf{R}1}}\sum_{j=k}^{k+N_c-1}\|\mathbf{v}_{j+1}\|^2
\end{aligned}
\tag{42}
$$

$$
\begin{aligned}
&\mathbf{e}_{k+1}^T\left(\mathbf{P}_{k+1}+\mathbf{S}_{k+1}\mathbf{S}_{k+1}^T\right)^{-1}\mathbf{e}_{k+1}=\left(\mathbf{A}_k\mathbf{e}_k-\mathbf{G}_k\mathbf{w}_k\right)^T\mathbf{M}_{k+1}^{-1}\left(\mathbf{A}_k\mathbf{e}_k-\mathbf{G}_k\mathbf{w}_k\right)+\mathbf{v}_{k+1}^T\mathbf{R}_{k+1}^{-1}\mathbf{v}_{k+1}\\
&\quad-\left(\mathbf{C}_{k+1}\left(\mathbf{A}_k\mathbf{e}_k-\mathbf{G}_k\mathbf{w}_k\right)-\mathbf{v}_{k+1}\right)^T\left(\mathbf{C}_{k+1}\mathbf{M}_{k+1}\mathbf{C}_{k+1}^T+\mathbf{R}_{k+1}\right)^{-1}\left(\mathbf{C}_{k+1}\left(\mathbf{A}_k\mathbf{e}_k-\mathbf{G}_k\mathbf{w}_k\right)-\mathbf{v}_{k+1}\right)\\
&\quad-\left(\mathbf{M}_{k+1}^{-1}\left(\mathbf{A}_k\mathbf{e}_k-\mathbf{G}_k\mathbf{w}_k\right)+\mathbf{C}_{k+1}^T\mathbf{R}_{k+1}^{-1}\mathbf{v}_{k+1}\right)^T\mathbf{S}_{k+1}\left(\mathbf{I}+\mathbf{S}_{k+1}^T\left(\mathbf{M}_{k+1}^{-1}+\mathbf{C}_{k+1}^T\mathbf{R}_{k+1}^{-1}\mathbf{C}_{k+1}\right)\mathbf{S}_{k+1}\right)^{-1}\mathbf{S}_{k+1}^T\\
&\quad\times\left(\mathbf{M}_{k+1}^{-1}\left(\mathbf{A}_k\mathbf{e}_k-\mathbf{G}_k\mathbf{w}_k\right)+\mathbf{C}_{k+1}^T\mathbf{R}_{k+1}^{-1}\mathbf{v}_{k+1}\right)\\
&\leq\left(\mathbf{A}_k\mathbf{e}_k-\mathbf{G}_k\mathbf{w}_k\right)^T\mathbf{M}_{k+1}^{-1}\left(\mathbf{A}_k\mathbf{e}_k-\mathbf{G}_k\mathbf{w}_k\right)+2\mathbf{v}_{k+1}^T\mathbf{R}_{k+1}^{-1}\mathbf{v}_{k+1}\\
&\quad-\frac{1}{2}\left(\mathbf{A}_k\mathbf{e}_k-\mathbf{G}_k\mathbf{w}_k\right)^T\mathbf{M}_{k+1}^{-1}\mathbf{S}_{k+1}\left(\mathbf{I}+\mathbf{S}_{k+1}^T\left(\mathbf{M}_{k+1}^{-1}+\mathbf{C}_{k+1}^T\mathbf{R}_{k+1}^{-1}\mathbf{C}_{k+1}\right)\mathbf{S}_{k+1}\right)^{-1}\mathbf{S}_{k+1}^T\mathbf{M}_{k+1}^{-1}\left(\mathbf{A}_k\mathbf{e}_k-\mathbf{G}_k\mathbf{w}_k\right)\\
&\leq\left(1-\frac{\alpha_{t,k+1}}{2}\right)\left(\mathbf{A}_k\mathbf{e}_k-\mathbf{G}_k\mathbf{w}_k\right)^T\mathbf{M}_{k+1}^{-1}\left(\mathbf{A}_k\mathbf{e}_k-\mathbf{G}_k\mathbf{w}_k\right)\\
&\quad+\frac{\alpha_{t,k+1}}{2}\left(\mathbf{A}_k\mathbf{e}_k-\mathbf{G}_k\mathbf{w}_k\right)^T\left(\mathbf{M}_{k+1}+\mathbf{S}_{k+1}\mathbf{S}_{k+1}^T\right)^{-1}\left(\mathbf{A}_k\mathbf{e}_k-\mathbf{G}_k\mathbf{w}_k\right)+2\mathbf{v}_{k+1}^T\mathbf{R}_{k+1}^{-1}\mathbf{v}_{k+1}\\
&\leq\left(1-\frac{\alpha_{t,k+1}}{2}\right)\left(1+\varepsilon_t\right)\mathbf{e}_k^T\mathbf{P}_k^{-1}\mathbf{e}_k+\frac{\alpha_{t,k+1}}{2}\left(1+\varepsilon_t\right)\mathbf{e}_k^T\left(\mathbf{P}_k+\mathbf{A}_k^{-1}\left(\mathbf{G}_k\mathbf{Q}_k\mathbf{G}_k^T+\mathbf{S}_{k+1}\mathbf{S}_{k+1}^T\right)\mathbf{A}_k^{-T}\right)^{-1}\mathbf{e}_k\\
&\quad+\frac{1+\varepsilon_t}{\varepsilon_t}\mathbf{w}_k^T\mathbf{Q}_k^{-1}\mathbf{w}_k+2\mathbf{v}_{k+1}^T\mathbf{R}_{k+1}^{-1}\mathbf{v}_{k+1}
\end{aligned}\tag{38}
$$

with $\gamma_t=\frac{1}{2N_c}\prod_{i=k}^{k+N_c-1}\alpha_{t,i+1}$. The corresponding matrix $\mathbf{S}_k\mathbf{S}_k^T$ can be computed in a recursive manner using (40). Using (13) and (9), $\mathbf{S}_k\mathbf{S}_k^T$ can be bounded by

$$\mathbf{S}_k\mathbf{S}_k^T\succeq c_{\mathbf{Q}1}\mathbf{\Phi}^{-1}(k+N_c,k)\mathbf{W}_c(k+N_c,k)\mathbf{\Phi}^{-T}(k+N_c,k).\tag{43}$$

By applying the bounds in Assumptions 1 and 2, and in Lemma 7, we subsequently obtain from (43) that

$$\mathbf{S}_k\mathbf{S}_k^T\succeq\frac{c_{\mathbf{Q}1}c_c}{c_{\mathbf{A}2}^{N_c}c_{\mathbf{P}2}}\mathbf{P}_k.\tag{44}$$

It follows from (17) and (39) that

$$
\begin{aligned}
V_k\leq\left(1+\varepsilon_t\right)^{k-k_0}V_{k_0}+\frac{\left(1+\varepsilon_t\right)^{k-k_0}}{\varepsilon_t c_{\mathbf{Q}1}}\sum_{i=k_0}^{k-1}\|\mathbf{w}_i\|^2\\
+\frac{2(1+\varepsilon_t)^{k-k_0-1}}{c_{\mathbf{R}1}}\sum_{j=k_0}^{k-1}\|\mathbf{v}_{j+1}\|^2
\end{aligned}\tag{45}
$$

for all $k\geq k_0\geq\max\{N_c,N_o\}$. Moreover, we get from (13), (17), (42) and (44) that

$$
\begin{aligned}
V_{k+N_c}\leq\eta_t V_k+\frac{\left(1+\varepsilon_t\right)^{N_c}}{\varepsilon_t c_{\mathbf{Q}1}}\sum_{i=k}^{k+N_c-1}\|\mathbf{w}_i\|^2\\
+\frac{2(1+\varepsilon_t)^{N_c-1}}{c_{\mathbf{R}1}}\sum_{j=k}^{k+N_c-1}\|\mathbf{v}_{j+1}\|^2
\end{aligned}\tag{46}
$$

for all $k\geq\max\{N_c,N_o\}$, where the constant η_t is given by $\eta_t=\left(1-\frac{\gamma_t c_{\mathbf{Q}1}c_c}{c_{\mathbf{Q}1}c_c+c_{\mathbf{A}2}^{N_c}c_{\mathbf{P}2}}\right)(1+\varepsilon_t)^{N_c}$. Without loss of generality, we assume that ε_t is sufficiently small such that $\eta_t<1$. Combining (45) and (46) yields

$$
\begin{aligned}
V_k\leq\left(\frac{(1+\varepsilon_t)^{N_c}}{\eta_t}\right)^{\frac{N_c-1}{N_c}}\eta_t^{\frac{k-k_0}{N_c}}V_{k_0}\\
+\frac{N_c(1+\varepsilon_t)^{N_c}}{\varepsilon_t c_{\mathbf{Q}1}(1-\eta_t)}\max_{k_0\leq i\leq k-1}\|\mathbf{w}_i\|^2\\
+\frac{2N_c(1+\varepsilon_t)^{N_c-1}}{c_{\mathbf{R}1}(1-\eta_t)}\max_{k_0\leq j\leq k-1}\|\mathbf{v}_{j+1}\|^2
\end{aligned}\tag{47}
$$

for all $k\geq k_0\geq\max\{N_c,N_o\}$. We obtain from (17) and Lemmas 6 and 7 that

$$\frac{1}{c_{\mathbf{P}2}}\|\mathbf{e}_k\|^2\leq V_k\leq\frac{1}{c_{\mathbf{P}1}}\|\mathbf{e}_k\|^2\tag{48}$$

for all $k\geq\max\{N_c,N_o\}$. The proof of the theorem follows from (47) and (48), with $c_{\mathbf{e}}=\sqrt{\frac{3c_{\mathbf{P}2}}{c_{\mathbf{P}1}}\left(\frac{(1+\varepsilon_t)^{N_c}}{\eta_t}\right)^{\frac{N_c-1}{2N_c}}}$, $\rho_{\mathbf{e}}=\eta_t^{\frac{1}{2N_c}}$, $c_{\mathbf{w}}=\sqrt{\frac{3c_{\mathbf{P}2}N_c(1+\varepsilon_t)^{N_c}}{\varepsilon_t c_{\mathbf{Q}1}(1-\eta_t)}}$ and $c_{\mathbf{v}}=\sqrt{\frac{6c_{\mathbf{P}2}N_c(1+\varepsilon_t)^{N_c-1}}{c_{\mathbf{R}1}(1-\eta_t)}}$. ■

Similar to the results in [8], [12], [16], Theorem 8 only provides a bound on the estimation $\mathbf{e}_k$ for $k\geq\max\{N_c,N_o\}$. Using Assumption 1, it can be shown that the estimation error $\mathbf{e}_k$ also remains bounded for values of k smaller than $\max\{N_c,N_o\}$, assuming bounded disturbances. Yet, any such bound depends on the initial condition $\mathbf{P}_0$. We note that, if $\mathbf{P}_0$ is chosen such that the bounds on the error covariance matrix in Lemmas 6 and 7 hold for all $k\geq 0$, then (37) of Theorem 8 is satisfied for all $k\geq k_0\geq 0$ and the estimation error is exponentially input-to-state stable as defined in Definition 5.

V. Conclusion

In this note, we have proved exponential input-to-state stability of the estimation error of the Kalman filter for deterministic discrete-time linear systems under suitable conditions. Because input-to-state stability is an important prerequisite for many contemporary analysis tools for cascaded and interconnected systems, this note provides a fundamental step in the analysis of systems of which the Kalman filter is a subsystem. Future research will entail the overall analysis of such systems.

References

[1] B. D. O. Anderson and J. B. Moore. Detectability and stabilizability of time-varying discrete-time linear systems. *SIAM Journal on Control and Optimization*, 19(1):20–32, 1981.

[2] M. Athans and E. Tse. A direct derivation of the optimal linear filter using the maximum principle. *IEEE Transactions on Automatic Control*, 12(6):690–698, 1967.

[3] M. Athans, R. Wishner, and A. Bertolini. Suboptimal state estimation for continuous-time nonlinear systems from discrete noisy measurements. *IEEE Transactions on Automatic Control*, 13(5):504–514, 1968.

[4] G. Besançon, H. Hammouri, and S. Benamor. State equivalence of discrete-time nonlinear control systems to state affine form up to input/output injection. *Systems & Control Letters*, 33(1):1–10, 1998.

[5] R. G. Brown and P. Y. C. Hwang. *Introduction to random singals and applied Kalman filtering*. John Wiley & Sons, 4th edition, 2012.

[6] H. Cox. On the estimation of state variables and parameters for noisy dynamic systems. *IEEE Transactions on Automatic Control*, 9(1):5–12, 1964.

[7] S. Dashkovskiy, B. S. Rüffer, and F. R. Wirth. An ISS small gain theorem for general networks. *Mathematics of Control, Signals, and Systems*, 19(2):93–122, 2007.

[8] J. J. Deyst and C. F. Price. Conditions for asymptotic stability of the discrete minimum-variance linear estimator. *IEEE Transactions on Automatic Control*, 13(6):702–705, 1968.

[9] W. W. Hager and L. L. Horowitz. Convergence and stability properties of the discrete Riccati operator equation and the associated optimal control and filtering problems. *SIAM Journal on Control and Optimization*, 14(2):295–312, 1976.

[10] J. P. Hespanha. *Linear systems theory*. Princeton University Press, Princeton, New Jersey, 2009.

[11] K. L. Hitz, T. E. Fortmann, and B. D. O. Anderson. A note on Bounds on solutions of the Riccati equation. *IEEE Transactions on Automatic Control*, 17(1):178–178, 1972.

[12] A. H. Jazwinski. *Stochastic processes and filtering theory*. Academic Press, New York, New York, 1970.

[13] Z.-P. Jiang and Y. Wang. Input-to-state stability for discrete-time nonlinear systems. *Automatica*, 37:857–869, 2001.

[14] R. E. Kalman. Contributions to the theory of optimal control. *Boletin de la Sociedad Matematica Mexicana*, 5:102–119, 1960.

[15] W. Li, G. Wei, D. Ding, Y. Liu, and F. E. Alsaadi. A new look at boundedness of error covariance of Kalman filtering. *IEEE Transactions on Systems, Man and Cybernetics*, 48(2):309–314, 2018.

[16] T. P. McGarty. *Stochastic systems and state estimation*. Wiley, New York, NY, 1974.

[17] J. B. Moore and B. D. O. Anderson. Coping with singular transition matrices in estimation and control stability theory. *International Journal of Control*, 31(2):571–586, 1980.

[18] D. Nešić and A. Loría. On uniform asymptotic stability of time-varying parameterized discrete-time cascades. *IEEE Transactions on Automatic Control*, 49(6):875–887, 2004.

[19] C. V. Rao, J. B. Rawlings, and J. H. Lee. Contrained linear state esimation – a moving horizon approach. *Automatica*, 37:1619–1628, 2001.

[20] C. V. Rao, J. B. Rawlings, and D. Q. Mayne. Constrained state estimation for nonlinear discrete-time systems: stability and moving horizon approximations. *IEEE Transactions on Automatic Control*, 48(2):246–258, 2003.

[21] P. Swerling. Modern state estimation methods from the viewpoint of the method of least squares. *IEEE Transactions on Automatic Control*, 16(6):707–719, 1971.

[22] J. C. Willems. Deterministic least squares filtering. *Journal of Econometrics*, 118(1-2):341–373, 2004.