CHARACTERIZATION OF EXPONENTIAL DIVERGENCE OF THE KALMAN FILTER FOR TIME-VARYING SYSTEMS

SIAM JOURNAL ON CONTROL AND OPTIMIZATION, v.48, n.5, p.2917-2944, 2009
http://producao.usp.br/handle/BDPI/15932

Downloaded from: Biblioteca Digital da Produção Intelectual - BDPI, Universidade de São Paulo
CHARACTERIZATION OF EXPONENTIAL DIVERGENCE OF THE
KALMAN FILTER FOR TIME-VARYING SYSTEMS*

EDUARDO F. COSTA† AND ALESSANDRO ASTOLFI‡

Abstract. This paper studies semistability of the recursive Kalman filter in the context of linear

time-varying (LTV), possibly nondetectable systems with incorrect noise information. Semistability

is a key property, as it ensures that the actual estimation error does not diverge exponentially. We

explore structural properties of the filter to obtain a necessary and sufficient condition for the filter to

be semistable. The condition does not involve limiting gains nor the solution of Riccati equations, as

they can be difficult to obtain numerically and may not exist. We also compare semistability with the

notions of stability and stability w.r.t. the initial error covariance, and we show that semistability in a

sense makes no distinction between persistent and nonpersistent incorrect noise models, as opposed

to stability. In the linear time invariant scenario we obtain algebraic, easy to test conditions for

semistability and stability, which complement results available in the context of detectable systems.

Illustrative examples are included.

Key words. Kalman filter stability, system structure, semistability

AMS subject classifications. 93E11, 93B99, 93D99

DOI. 10.1137/080717055

1. Introduction. It is a well-known fact that Kalman filters (KFs) may present

the phenomenon of divergence under incorrect model and noise information [24] in

such a manner that the state estimate may present a fast divergence from the actual

value, sometimes with exponential rate. The subject has deserved much attention, and

sufficient conditions for stability of the filter were obtained for the recursive KF (see,

e.g., [3, 19, 23, 24, 27]), as well as for alternative filters such as stationary/frozen filters

[4, 16, 27]. Furthermore, $H_\infty$ filters and a variety of robust filters that are inherently

stable have been developed in different contexts; see, e.g., [18, 22, 25, 26, 30]. However,

in spite of these achievements, there are still issues to be understood in the study of

divergence of the recursive KF, as explained later. The recursive KF is attractive

for many applications, in part due to the ease of implementation and in part due to

the fact that it is optimal for the nominal model that, in many cases, is the most

significant one, while other filters are often optimal for simplified or “worst-case”

models that may be of little relevance.

We study divergence of the standard KF as follows. Consider the discrete-time,

time-varying stochastic system defined in a fundamental probability space $(\Omega, \mathcal{F}, \mathbb{P})$ by

\[
\Phi : \begin{cases}
    x(k+1) = A_k x(k) + B_k w(k), & x(0) = x_0, \\
    y(k) = C_k x(k) + D_k v(k),
\end{cases}
\]

*Received by the editors February 25, 2008; accepted for publication (in revised form) August 3,
2009; published electronically November 5, 2009. This work was supported in part by FAPESP
grants 06/02004-0 and 06/04210-6, CNPq grants 304429/2007-4 and 482386/2007-0, and the EPSRC
Research Grant EP/E057438, Nonlinear Observation Theory with Applications to Markov Jump
Systems. This paper is based on two preliminary papers in Proceedings of the American Control
Conference, 2008 [9, 10].

†Departamento de Matemática Aplicada e Estatística, Universidade de São Paulo, C.P. 668, 13560-
970, São Carlos, SP, Brasil (efcosta@icmc.usp.br).

‡Electrical and Electronic Engineering Department, Imperial College London, SW7 2AZ, London,
UK, and Dipartimento di Informatica, Sistemi e Produzione, Università di Roma “Tor Vergata,”
00133 Roma, Italia (a.astolfi@imperial.ac.uk).

2917
where \( x(k) \in \mathbb{R}^n \) is the state, \( y(k) \in \mathbb{R}^r \) is the observed variable, \( w(k) \in \mathbb{R}^p \) and \( v(k) \in \mathbb{R}^q \) form independent noise processes, and \( x_0 \) is a (independent) zero-mean random variable satisfying \( \mathbb{E}\{x_0x_0^\prime\} = \Psi \). We assume \( D_kD_k^\prime > 0, k \geq 0 \) (nonsingular measurement noise) and work with a hypothesis of structural invariance of the stable/unstable spaces\(^1\) of \( A_k \), as detailed in section 2. The standard KF provides an estimate \( \hat{x}(k) \) for \( x(k) \) and is constructed via the Riccati difference equation (RDE), with no initialization correction or any procedure for ensuring stability. Assume the calculated error covariances \( P(k), k \geq 0 \), are bounded, however, possibly due to the incidence of nonmodeled disturbances at some time instant \( k \), the actual error covariance \( \tilde{X}(k) \) does not coincide with the calculated \( P(k) \). A fundamental question that arises is whether \( \tilde{X}(k) \) diverges exponentially in spite of \( P(k) \) being bounded. The problem is formulated as follows. Assume that the KF is calculated for some available data \( E_k \) and \( \Sigma \) instead of the actual values \( B_k \) and \( \Psi \), respectively. For each \( 0 \leq \xi < 1 \), \( \Psi \) and sequence \( B_k, k \geq 0 \), find the existence of \( \bar{X} \) such that \( \tilde{X}(k) \leq \xi^{-k}\bar{X} \). We say that a KF satisfying this condition is semistable. No additional hypothesis is taken into account apart from bounded \( P(k) \), and hence the system may be nondetectable, and no constraint is imposed on the noise uncertainties, as \( \Sigma - \Psi \) and \( E_k - B_k \) may be of arbitrary magnitude, which excludes many robust formulations.

Apart from being an involving technical question, semistability is relevant for applications in which polynomial divergence is acceptable. For instance, semistability ensures convergence to zero of the discounted estimation error with arbitrary discount rate \( 0 < \gamma < 1 \), since \( \mathbb{E} \{ (\gamma)^k ||x(k) - \hat{x}(k)||^2 \} = \gamma^k \text{tr}(X(k)) \leq (\gamma/\xi)^k \text{tr}(\bar{X}) \), converges to zero as \( k \to \infty \) for \( \xi > \gamma \); the total discounted estimation error converges, \( \sum_{k=0}^{\infty} \gamma^{-k}\text{tr}(\tilde{X}(k)) < \infty \). Moreover, in one of the most unfavorable situations for filtering, in which \( \mathbb{E} \{ ||x(k)||^2 \} \) diverges exponentially, semistability ensures that the relative error \( \mathbb{E} \{ ||x(k) - \hat{x}(k)||^2 \}/\mathbb{E} \{ ||\hat{x}(k)||^2 \} \) decreases exponentially as \( k \to \infty \). See Remark 4 and the illustrative Example 3.

Available results on the divergence of the KF consider simplified scenarios, imposing detectability-like conditions or the existence of a limiting stationary or periodic solution for \( P(k) \), as in [19, 23, 24, 27]. Limiting solutions can be conveniently employed for many evaluations by comparison and ordering of solutions. The problem is more complex when the recursive KF has no limiting stationary solutions, and it is still more difficult when it does not converge even to a periodic solution, because we cannot employ the aforementioned tools nor existing results on convergence of RDE. Note that the lack of a limiting stationary solution is not necessarily associated with time-varying or periodic systems; see Example 6. Detectability hypotheses are natural when dealing with stability of the KF but are conservative for semistability. See Example 6 for a nondetectable system with a semistable KF. Another gap is that existing stabilizability-like conditions for KF stability disregard \( \Sigma \) (see, e.g., [3]), and this is rather restrictive, as illustrated by Example 5. Moreover, available results address only stability, meaning that \( X(k) \) is bounded for any incorrect noise covariances \( E_k \) or \( \Sigma \), which is of course a sufficient condition for semistability.

The results in this paper do not rely on comparisons with RDE solutions. Indeed, our approach involves replacing the RDE variable \( P(k) \) with another quantity in the condition for semistability, as explained next. We start by deriving some links between \( P(k) \) and \( X(k) = \mathbb{E}\{x(k)x(k)^\prime\} \) (defined assuming \( B_k = E_k \) and \( \Psi = \Sigma \)), provided the former structurally describes the latter, in the sense of Lemma 4, which clarify the role of the condition \( \Sigma > 0 \). The fact that the null space of \( P(k) \) coincides with the null

---

\(^1\)Please see section 2 for definitions.
space of $X(k)$ together with an orthogonality property involving the Kalman gain (see Corollaries 2 and 3) allows us to "eliminate" the Kalman gain from the semistability condition: the KF is semistable if and only if, for each $0 \leq \xi < 1$, there exists $\hat{Z}$ such that $Z(k) \leq \xi^{-k}\hat{Z}$, where the symmetric and positive semidefinite matrix $Z(k)$, $k \geq 0$, is a component of the state of the system described by

$$
\Theta : \begin{cases}
Z(k+1) = (H_{k+1}A_k)Z(k)(H_{k+1}A_k)', \\
X(k+1) = A_kX(k)A_k' + E_kE_k', 
\end{cases} \quad k \geq 0,
$$

with initial condition $(Z(0), X(0)) = (H_0H_0', \Sigma)$, where $H_k$ is the projection onto the null space of $X(k)$. Note that this condition involves neither detectability assumptions, nor existence of limiting gains, nor calculations of the Kalman gains and of the solution of the associated RDE. We also show that semistability makes no distinction among the incorrect data being related to persistent or nonpersistent noise, in the sense that we can set $B_k = E_k$ and $\Psi = I$ in the analysis, as opposed to stability; see Lemma 1 and Example 1. We also provide, in Lemma 2, comparisons between semistability and stability notions.

In the linear time invariant (LTI) scenario, we combine the condition derived in the time-varying context with an available condition for semistability of a version of system $\Theta$ in (2) with $E = 0$ [11]. To make these conditions compatible, we employ $H$, the orthogonal projection onto the noncontrollable space of $(A, E)$, to obtain a modified system matrix $HA$ that incorporates $E$ in an adequate way. This allows us to show in Theorem 2 that the KF is semistable if and only if

$$\ker\{J_H\Sigma J_H'\} \cap J_H = \{0\},$$

where $J_H$ is such that $J_HAHJ_H^{-1}$ is in Jordan form and $J_H$ is the unstable subspace of $J_HAHJ_H^{-1}$. Condition (3) appears in [6, Theorem 1] as a necessary and sufficient condition for $\Sigma$ to belong to the basin of attraction of the strong solution of the algebraic Riccati equation (ARE) for detectable systems (and also in [15, Theorem 3] and [21, Corollary II.1] as a sufficient one, with a slightly different notation); from this standpoint, the present paper clarifies the meaning of the above condition in the scenario of nondetectable systems. We combine existing results [5, 6, 14, 15, 17, 21, 24] and Theorem 2 to obtain a necessary and sufficient condition for stability of the KF in Theorem 3.

The paper is organized as follows. Section 2 presents notation, some preliminary results, and the stability concepts. Section 3 addresses the role of the initial covariance matrices $\Sigma$ and $\Psi$ and some related results that are essential to obtain the main result in section 4. Section 5 is focused on the LTI scenario. Illustrative examples are presented in section 6. Finally, section 7 provides some conclusions, and the appendix contains some technical proofs.

2. Definitions and preliminary results. Let $\mathbb{D}$ (respectively, $\overline{\mathbb{D}}$) be the open (respectively, closed) unit disk in the complex plane. Let $\mathcal{R}^{r \times s}$ (respectively, $\mathcal{R}^r$) represent the normed linear space formed by all $r \times s$ (respectively, $r \times r$) real matrices, and $\mathcal{R}^{r \times s}(\mathcal{R}^{r \times 0})$ represents the cone $\{U \in \mathcal{R}^r : U = U'\}$ (the closed convex cone $\{U \in \mathcal{R}^r : U = U', U \geq 0\}$) where $U'$ denotes the transpose of $U$; $U \geq V$ signifies that $U - V \in \mathcal{R}^{r \times 0}$. For $U \in \mathcal{R}^n$, $U^*$ stands for the pseudo-inverse of $U$. $\lambda_i(U)$, $i = 1, \ldots, n$, stand for the eigenvalues of $U$. For a nontrivial $U \in \mathcal{R}^{n \times 0}$, $\lambda_{-}(U)$ is the smallest strictly positive eigenvalue of $U$. $\lambda_i(U)$ lying in $\mathbb{D}$ (respectively, $\overline{\mathbb{D}}$ and $\mathbb{D}/\mathbb{D}$) is referred to as a stable
(semistable\(^2\) and strictly semistable) eigenvalue of \(U\); the associated eigenvector \(v \in \mathbb{R}^n\) is stable (semistable and strictly semistable); otherwise it is unstable. The space spanned by the stable eigenvectors of \(U\) is referred to as the stable subspace of \(U\) and is referred to similarly for semistable, strictly semistable, and unstable spaces. Let \(H^{r,n}\) denote the linear space formed by the sequences of matrices \(H = \{H_i \in \mathbb{R}^{r,n}; i \in \mathbb{Z}\}\) such that \(\sup_{i \in \mathbb{Z}} ||H|| < \infty\). Let \(H^u = H^{n,n}\) and \(||H||_\infty = \sup_{i \in \mathbb{Z}} ||H_i||\). For \(U, V \in H^{r,n}\), let \(U + V = \{U_i + V_i, i \in \mathbb{Z}\}\), and similarly for other mathematical relations; e.g., \(U = 0\) means that \(U_i = 0, i \geq 0\).

Consider system \(\Phi\) in (1). Assume that \(A \in H^{n}, B \in H^{n,p}, C \in H^{r,n},\) and \(D \in H^{r,q}\), with \(DD' > \chi I\) for some \(\chi > 0\) (nonsingular measurement noise). \(w\) and \(v\) form stationary zero-mean independent white noise processes satisfying (without loss of generality) \(E\{w(k)w(k)\}' = I\) and \(E\{v(k)v(k)\}' = I\). The independent random variable \(x_0\) has Gaussian distribution with \(E\{x_0\} = \bar{x}\) (we assume \(\bar{x} = 0\) unless stated otherwise) and \(E\{x_0x_0\}' = \Psi\).

The standard recursive KF provides the estimates \(\hat{x}(0) = \bar{x}\) and

\[
\hat{x}(k+1) = A_k\hat{x}(k) + L_k[y(k) - C_k\hat{x}(k)],
\]

where the Kalman gain

\[
L_k = A_kP(k)C_k'[C_kP(k)C_k' + D_kD_k']^{-1}
\]

is calculated via the RDE

\[
P(k + 1) = A_k[P(k) - P(k)C_k'[C_kP(k)C_k' + D_kD_k']^{-1}C_kP(k)]A_k' + E_kE_k'
\]

with initial condition \(P(0) = \Sigma \in \mathbb{R}^{n_0}\), where \(E \in H^{n,p}\). \(\Sigma\) and \(E\) are the available information on the actual covariance matrices \(\Psi\) and \(B\), respectively. We consider the following assumptions.

**Assumption 1.** For each \(\Sigma \in \mathbb{R}^{n_0}\) there is \(\bar{X} \in \mathbb{R}^{n_0}\) such that \(P(k) \leq \bar{X}, k \geq 0\).

**Remark 1.** Assumption 1 holds trivially provided that \(P(k)\) in (6) converges as \(k \to \infty\) or, in particular, that \((A,C)\) is detectable; see, e.g., [1, 6]. The assumption is connected to the fact that we address divergence of the actual error in absence of divergence of the nominal error.

Consider the state transition matrix \(A(k + t, t) = A_{k+t} \cdots A_t\), \(k, t \geq 0\). For \(\tau > 0\), let \(M_{\tau}(k + t, t)\) stand for a projection onto the unstable space of \(\tau A(k + t, t)\) such that \(M_{\tau}(k + t, t)v = 0\) whenever \(v\) is in the semistable space of \(\tau A(k + t, t)(M_{\tau}\) is an orthogonal projection when \(A\) is in Jordan form).

**Assumption 2.** There exists \(T\) and \(0 < \tau < 1\) such that \(M_{\tau}((k + 1)T, kT) = M_{\tau}((k + 1)T, kT) = M_{\tau}(T, 0), k \geq 0\).

For simplicity, we write \(M = M_1(T, 0)\), where \(T\) is as in Assumption 2.

**Remark 2.** Assumption 2 holds trivially for time invariant or periodic systems, \(T\) being the period and with \(\zeta^{-1} < \tau < 1\), where \(\zeta\) is the modulus of the unstable eigenvalue of \(A(T + t, t)\) that is closer to \(\mathbb{D}\) (if any; otherwise \(0 < \tau < 1\)). The assumption can be weakened by considering variable step sizes, but we prefer to avoid the associated notational complexity. Assumption 2 requires that the unstable space of \(A\), as well as the semistable space of \(A\), be invariant along a sequence of time instants with step size \(T\). Moreover, the unstable spaces should present a "minimal

---

\(^2\)Following the terminology of [1].
One important drawback is the phenomenon of divergence, discussed hereafter. The minimal mean square estimator) and that it improves the signal-to-noise ratio [2].

A phenomenon of divergence, illustrated in Example 1.

The KF is stable (stable w.r.t. (ii)

Definition 1 (KF semistability).

(i) The KF is semistable if, for each $B \in \mathcal{H}^{n,p}$, $\Psi \in \mathcal{R}^{n0}$, and $0 \leq \xi < 1$, there exists $X$ such that $X(k, B, \Psi) \leq \xi^{-2k} X$, $k \geq 0$.

(ii) The KF is stable (stable w.r.t. $\Psi$) if, for each $B \in \mathcal{H}^{n,p}$ and $\Psi \in \mathcal{R}^{n0}$ (for $B = E$ and for each $\Psi \in \mathcal{R}^{n0}$), there exists $X$ such that $X(k, B, \Psi) \leq \bar{X}$, $k \geq 0$.

It is worth mentioning that we do not introduce a notion of semistability w.r.t. $\Psi$ because, as we shall see in section 3, it is equivalent to semistability.

Similarly to Proposition 1, if we define $X(0, 0, B, \Psi) = \Psi$ and

$$X(k + 1, t, B, \Psi) = A_k X(k, t) A_k^T + B_k B_k^T, \quad k \geq t \geq 0,$$

then $X(k, 0, B, \Psi) = E\{x(k)x(k)\}$. The homogeneous solutions of (8) and (9) are given as

$$X_h(k, t, \Psi) = (A_k - L_k C_k) \cdots (A_t - L_t C_t) \Psi (A_t - L_t C_t)^T \cdots (A_k - L_k C_k)^T,$$

$$X_h(k, t, \Psi) = A_k \cdots A_t \Psi A_t^T \cdots A_k, \quad k \geq t \geq 0.$$
For convenience we define \( \tilde{X}_h(k-1, k, \Psi) = \Psi \) and \( \tilde{X}_h(k-1, k+\ell, \Psi) = 0, \ell \geq 1 \). We omit the variables \( t, \Psi, B \) when \( t = 0, \Psi = \Sigma, \) and \( B = E \), respectively, except in contexts in which a more complete notation is preferred. For instance, we denote \( X(k, 0, E, \Sigma) \) simply by \( X(k) \).

**Remark 4.** It can be shown that a necessary condition for \( \tilde{X} \) to diverge exponentially as \( k \to \infty \) is that \( X \) diverges exponentially; in fact, exploiting Proposition 2 and setting \( 0 < \epsilon < 1 \), we write

\[
E\{\|\xi(k)\|^2\}/E\{\|\tilde{\xi}(k)\|^2\} = tr\{\tilde{X}(k, B, \Psi)\}/E\{(x(k) - \tilde{x}(k))' (x(k) - \tilde{x}(k))\}
\leq tr\{\tilde{X}(k, B, \Psi)\}/E\{(1 - \epsilon^2)x(k)'x(k) + (1 - \epsilon^{-2})\tilde{x}(k)'\tilde{x}(k)\} - tr\{\tilde{X}(k)\}
= tr\{\tilde{X}(k, B, \Psi)\}/((1 - \epsilon^2)tr\{X(k, B, \Psi)\} + (1 - \epsilon^{-2})tr\{\tilde{X}(k, B, \Psi)\})
\]

and \( X \) dominates \( \tilde{X} \) (the latter may diverge polynomially, at most). An interpretation in terms of signal-to-noise ratio (SNR) can be derived. For instance, assume that \( y(k) \) is scalar, and consider the SNR at the filter output \( E\{\|\hat{x}(k)\|^2\}/E\{\|\tilde{x}(k)\|^2\} \) decreases exponentially as \( k \to \infty \); in fact, exploiting Proposition 2 and setting \( 0 < \epsilon < 1 \), we write

\[
\frac{E\{\|\hat{x}(k)\|^2\}}{E\{\|\tilde{x}(k)\|^2\}} = tr\{\tilde{X}(k, B, \Psi)\}/E\{(x(k) - \hat{x}(k))' (x(k) - \hat{x}(k))\}
\leq tr\{\tilde{X}(k, B, \Psi)\}/E\{(1 - \epsilon^2)x(k)'x(k) + (1 - \epsilon^{-2})\hat{x}(k)'\hat{x}(k)\} - tr\{\tilde{X}(k)\}
= tr\{\tilde{X}(k, B, \Psi)\}/((1 - \epsilon^2)tr\{X(k, B, \Psi)\} + (1 - \epsilon^{-2})tr\{\tilde{X}(k, B, \Psi)\})
\]

The condition for semistability is described next.

**Condition 1.** Consider the system (which is the same as (2))

\[
\begin{align*}
\Theta : & \quad Z(k + 1) = (H_{k+1}A_k)Z(k)(H_{k+1}A_k)', \\
& \quad X(k + 1) = A_kX(k)A_k' + E_kE_k', \quad k \geq 0,
\end{align*}
\]

with initial condition \( (Z(0), X(0)) = (H_0H_0', \Sigma) \), where \( H_k, k \geq 0 \), stands for the orthogonal projection onto the null space of \( X(k) \). For each \( 0 \leq \xi < 1 \), \( \tilde{Z} \) exists such that \( Z(k) \leq \xi^{-k}Z, \quad k \geq 0 \).

One useful fact related to semistability is that, assuming the matrix \( A_k \) is semistable for each \( k \geq 0 \), the matrix \( \xi A_k \) is stable with a uniform margin of stability. For instance, the system \( x(k + 1) = (\xi A_k)x(k) + B_kw(k) \) accommodates persistent noise in the sense that \( E\{\|x(k)\|^2\} \) is bounded for any \( B \in \mathcal{H}^{n,p} \). To extend this idea to semistability of the KF in section 3, we define \( \tilde{X}_\xi(k, 0, B, \Psi) \in \mathcal{R}^{n_0}, k \geq 0 \), by \( \tilde{X}_\xi(0, 0, B, \Psi) = \Psi \) and

\[
\begin{align*}
\tilde{X}_\xi(k + 1, 0, B, \Psi) = & \quad (\xi(A_k - L_kC_k))\tilde{X}_\xi(k, 0, B_k, \Psi)(\xi(A_k - L_kC_k))' \\
& + L_kDD'L_k + B_kB_k', \quad k \geq 0.
\end{align*}
\]

We omit the variables \( t, \Psi, B \) when \( t = 0, \Psi = \Sigma, \) and \( B = E \), respectively.

We now gather some properties and inequalities, the proofs of which are given in Appendix A.

**Proposition 2.** The following statements hold.

(i) For each \( \Sigma \in \mathcal{R}^{n_0} \) there exists \( \rho \geq 0 \) such that \( \|L_k\| \leq \rho \) for \( k \geq 0 \).

(ii) Let \( V \in \mathcal{R}^{n_0} \). For each \( \Sigma \in \mathcal{R}^{n_0} \) there exists \( \delta \geq 0 \) such that \( \tilde{X}_h(k+1, k, V) \leq \delta\|V\| \).

(iii) Let \( 0 \leq \alpha \leq 1 \) and \( V_0, V_1 \in \mathcal{R}^{n_0} \) and assume \( V_1 \geq V_0 \). Then \( \tilde{X}_h(k, \alpha V_1) \geq \alpha \tilde{X}_h(k, V_0) \), \( k \geq 0 \). Similarly, \( \tilde{X}(k, \alpha V_1) \geq \alpha \tilde{X}(k, V_0) \) and \( \tilde{X}_\xi(k, \alpha V_1) \geq \alpha \tilde{X}_\xi(k, V_0), k \geq 0 \).
1) Assume that \( \tilde{X}(k) \) is \\( (1 - \varepsilon^2)U V V' + \varepsilon^2 (I - U)' V (I - U)' \) \( \leq \varepsilon \), and Proposition 2(iii) leads to \( \tilde{X}(k, 0, E, I) \). From (1) and (10) we obtain

\[
\tilde{X}(k, 0, E, I) = \xi^{2k} \tilde{X}(k, 0, I) + \sum_{\ell=1}^{k} \xi^{2(k-\ell)} \tilde{X}(k-1, \ell, \Upsilon_{\ell}).
\]

For the first term on the right-hand side of (14), it is simple to check that

\[
\xi^{2k} \tilde{X}(k, 0, I) \leq \xi^{2k} \tilde{X}(k, 0, B = 0, I) \leq \tilde{X}.
\]

For the second term on the right-hand side of (14), note from the RDE (6) that \( P(\ell + 1) \geq \Upsilon_{\ell} \), \( \ell \geq 0 \); furthermore, setting \( \rho \) as in Proposition 2(i), we obtain \( \tilde{X}(\ell, \ell, \Upsilon_{\ell}) \leq \xi \Upsilon_{\ell} \) where \( \xi = (\|A\| + \rho\|C\|)^2 \), and Proposition 2(iii) leads to

\[
\tilde{X}(k-1, \ell, \Upsilon_{\ell}) = \tilde{X}(k-1, \ell + 1, \tilde{X}(\ell, \ell, \Upsilon_{\ell})) \leq \tilde{X}(k-1, \ell + 1, \xi \Upsilon_{\ell})
\]

\[
= \xi \tilde{X}(k-1, \ell + 1, \Upsilon_{\ell}) \leq \xi \tilde{X}(k-1, \ell + 1, P(\ell + 1))
\]

\[
\leq \xi \tilde{X}(k-1, \ell + 1, P(\ell + 1)) = \xi P(k - 1) \leq \xi \tilde{X}.
\]

Substituting (15) and (16) in (14) provides \( \tilde{X}(k, 0, E, I) \leq (\xi^{2k} + \xi \sum_{\ell=1}^{k} \xi^{2(k-\ell)}) \tilde{X} \), and the result follows, since \( 0 \leq \xi < 1 \) (uniformly over \( k \)).

Sufficiency. Assume that \( B = E \), that is, there is no error in the persistent noise data. For each \( \Psi \in \mathcal{M}^{n_0} \) there exists \( \kappa \geq 1 \) for which \( \Psi \leq \kappa I \), and hence from Proposition 2(iii), (iv) we obtain

\[
\tilde{X}(k, \Psi) \leq \tilde{X}(k, \kappa I) \leq \kappa \xi^{-2k} \tilde{X}(k, I) \leq \xi^{-2k}(\kappa \tilde{X}).
\]

To incorporate the case \( B \neq E \), we again make use of the margin provided by the uniform \( \xi \); more specifically, for each \( 0 \leq \xi < 1 \) we set \( \xi < \xi < 1 \) and, by hypothesis, \( \tilde{X}(k, 0, E, I) \). Assume for now that (the proof of this inequality is in Appendix B)

\[
\tilde{X}(k + t, t, B B') \leq \delta^{2T} \xi^{-2kT} \tilde{X}(\|B\|)^2, \quad k, t \geq 0.
\]
The KF leads to bounded $\tilde{X}(k,0,B,\Psi)$ if $\Psi$ is stable with respect to $\Psi$, but not stable. Figure 1 shows that Lemma 2) and linearly divergent $\tilde{X}(k,0,B,\Psi)$, hence the definition of stability is analogous to the proof of Lemma 1, in the specific case in which $\Psi = 0$.

Moreover, note that replacing $0$ with $\Psi$ in Assumption 2 leads to the stability definition. These facts suggest that a result similar to the one in Lemma 1 does not hold when $\Psi = 1$.

The margin provided by $\tilde{X}$ in the proof of Lemma 1 does not hold when $\Psi = 1$. Moreover, note that replacing $0 < \Psi < 1$ with $\Psi = 1$ in the semistability definition leads to the stability definition. These facts suggest that a result similar to the one in Lemma 1 does not hold for stability, and hence there should be a distinction between accommodation of nonpersistent and persistent incorrect noise models, when stability is concerned. The next lemma makes this precise; see also Example 1.

**Lemma 2.** Stability w.r.t. $\Psi$ is implied by stability, implies semistability, and is equivalent to stability w.r.t. $I$.

**Proof.** The first statement is immediate. The second statement follows straightforwardly from Lemma 1, since $X_{\Psi}(k,E,I) \leq \tilde{X}(k,E,I)$. The proof of the third statement is analogous to the proof of Lemma 1, in the specific case in which $B = E$.

**Example 1** (strict stability w.r.t. $\Psi$). Consider system $\Phi$ with

$$A_i = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}, E_i = 0, \Sigma = C_i = D_i = I, i \geq 0.$$ 

The KF leads to bounded $\tilde{X}(k,\Psi = I)$ (it is enough to consider $\Psi = I$, according to Lemma 2) and linearly divergent $\tilde{X}(k,B = 0.2I,\Psi = I)$ (see Figure 1), yielding that the KF is stable w.r.t. $\Psi$ but, clearly, not stable. Figure 1 shows that $A_k - L_kC_k$ presents a vanishing margin of stability and tends to a semistable matrix as $k \to \infty$.

Figure 1 also provides an illustration of the fact that the implementation of the limiting gain $L_\infty$ [27] yields semistability (indeed, $L_\infty = 0$ is “overoptimistic”), and the implementation of a frozen gain $L_k = L_{k_F}$, $k \geq k_F$, for some $k_F$ [4, 16] yields stability (“overpessimistic”).

One important property of the KF is that, provided that at time $k$ the actual error covariance is structurally described by the nominal $P$, it remains described by $P$ at successive instants.

**Lemma 3.** Assume that $V \in \mathcal{R}^{n_0}$ and $t \geq 0$ are such that $\ker\{P(t)\} \subset \ker\{V\}$. Then there exists $X$ such that $\tilde{X}(k+t,t,V) \leq X$, $k \geq 0$. Moreover, $\ker\{P(k+t)\} \subset \ker\{\tilde{X}(k+t,t,V)\}$, $k \geq 0$.

---

3Terminology employed, for instance, in [19].
Proof. For $V$ such that $\ker \{ V \} \supset \ker \{ P(t) \}$ we can pick a scalar $0 < \kappa \leq 1$ such that $\kappa V \leq P(t)$. Therefore, Proposition 2(iii) (with $\alpha = 0$) and Assumption 1 yield

$$\tilde{X}(k + t, t, \kappa V) \leq \tilde{X}(k + t, t, P(t)) = P(k + t) \leq \tilde{X}, \quad k \geq 0. \tag{19}$$

We set $\alpha = \kappa$ and $V_0 = V_1 = V$ in Proposition 2(iii) to obtain $\kappa \tilde{X}(k + t, t, V) \leq \tilde{X}(k + t, t, \kappa V)$; this and (19) provide

$$\tilde{X}(k + t, t, V) \leq \kappa^{-1} \tilde{X}, \quad k \geq 0. \tag{20}$$

Regarding the second statement, it follows immediately from (19) that

$$\ker \{ P(k + t) \} = \ker \{ \tilde{X}(k + t, k, P(t)) \} \subset \ker \{ \tilde{X}(k + t, t, \kappa V) \}. \tag{21}$$

Now, for $\eta \in \ker \{ \tilde{X}(k + t, t, \kappa V) \}$ it is simple to check that $\eta \in \ker \{ \tilde{X}(k + t, t, \kappa V) \}$ (in fact, $0 = \eta' \tilde{X}(k + t, t, \kappa V) \eta = \eta' (\kappa \tilde{X}_h(k + t, t, V) + \tilde{X}(k + t, t, \Psi = 0)) \eta$ yields that $\tilde{X}_h(k + t, t, V) \eta = 0$ and $\tilde{X}(k + t, t, \Psi = 0) \eta = 0$). We have shown that $\ker \{ \tilde{X}(k + t, t, \kappa V) \} \subset \ker \{ \tilde{X}(k + t, t, V) \}$, and substituting this relation in (20) yields the result.

By fixing $t = 0$ and setting $P(0) = \Sigma > 0$, the condition in Lemma 3 is trivially satisfied for any $V$, yielding the next sufficient condition for stability w.r.t. $\Psi$ and making clear the role of $\Sigma > 0$. It retrieves the fact that $\Sigma > 0$ is a sufficient condition for semistability of the KF in periodic or time invariant scenarios; see, e.g., [17] and [24].

**Corollary 1.** If $\Sigma > 0$, then the KF is stable w.r.t. $\Psi$.

The upper bound on $\tilde{X}$ provided by Lemma 3 is dependent on $V$ and $t$. This lack of uniformity cannot be removed in general. However, for the homogeneous solution $\tilde{X}_h$ of (8) we can employ the second statement of Lemma 3 to derive the next result.

**Lemma 4.** Let $t \geq 0$ and $P(t)$ be given by (6). For each $0 \leq \xi < 1$ there exists $\tilde{X}$ such that $\tilde{X}_h(k + t, t, V) \leq \xi^{-2k} \| V \| \tilde{X}, \quad k \geq 0$, for all $V$ for which $\ker \{ V \} \supset \ker \{ P(t) \}$.

**Proof.** See Appendix C.

4. **Semistability and Condition 1.** This section explores the connections between the Kalman gain $L$ and the covariance matrices $X$ and $P$, seeking to remove the dependence on $L$ (and the associated RDE) from the condition of Lemma 1. We start...
by showing an orthogonality result involving $L_k C_k$ and the null space of $X(k + 1)$, with the interpretation that the KF behaves exactly as the plant in subspaces with no associated noise. Then we show that the null spaces of $P$ and $X$ coincide, thus extending the orthogonality result to $P$; in the time invariant case, this is to some extent analogous to [19, Theorem 1]. These results allow us to show that Condition 1 is necessary and sufficient for semistability of the KF.

**Lemma 5.** The following statements hold.
(i) $\ker\{P(k + 1)\} = \ker\{X(1, 0, P(k))\}$, $k \geq 0$.
(ii) For each $v \in \mathbb{R}^n$, $L_k C_k v$ is orthogonal to $\ker\{P(k + 1)\}$.

**Proof.** We start by showing that $\ker\{P(k + 1)\} \supset \ker\{X(1, 0, P(k))\}$. By optimality of the KF, we have that $P(k + 1) = \hat{X}(k + 1) \leq X(k + 1, k, \hat{X}(k)) = X(1, 0, P(k))$; hence for $\eta \in \ker\{X(1, 0, P(k))\}$,

$$\eta' P(k + 1) \eta \leq \eta' X(1, 0, P(k)) \eta = 0.\label{21}$$

Conversely, if we pick an arbitrary $\eta \in \ker\{P(k + 1)\}$, we can employ Proposition 1 to write

$$0 = \eta' P(k + 1) \eta = \eta' \hat{X}(k + 1) \eta = \eta' [(L_k^2 + L_k C_k) X(k)] (A_k + L_k C_k)' + L_k D_k D_k' L_k + E_k E_k' \eta \label{22}$$

which, in particular (and recalling that $DD' > 0$), means that

$$L_k \eta = 0, \quad \eta \in \ker\{P(k + 1)\},\label{23}$$

allowing us to re-evaluate (22) as

$$0 = \eta' [(A + L_k C) \hat{X}(k) (A_k + L_k C_k)' + L_k D_k D_k' L_k + E_k E_k'] \eta = \eta' [A_k \hat{X}(k) A_k' + E_k E_k'] \eta = \eta' X(1, 0, P(k)) \eta,$$

completing the proof of (i). For (ii), note that (23) immediately leads to $\eta' (L_k C_k v) = 0$; for each $\eta \in \ker\{P(k + 1)\}$; thus $L_k C_k v$ is orthogonal to $\ker\{P(k + 1)\}$. □

**Corollary 2.** $\ker\{P(k)\} = \ker\{X(k)\}$, $k \geq 0$.

**Proof.** We proceed inductively. For $k = 0$, $P(0) = X(0, 0, \Sigma) = \Sigma$ by definition. Now assume that $\ker\{P(k)\} = \ker\{X(k)\}$ holds for $k$, and note that

$$\ker\{A_k P(k) A_k' + E_k E_k'\} = \ker\{A_k X(k) A_k' + E_k E_k'\} = \ker\{X(k + 1)\}.\label{24}$$

Lemma 5 and (9) yield

$$\ker\{P(k + 1)\} = \ker\{X(k + 1, k, P(k))\} = \ker\{A_k P(k) A_k' + E_k E_k'\}.\label{25}$$

Hence (24) and (25) complete the induction. □

**Corollary 3.** The following statements hold.
(i) For each $v \in \mathbb{R}^n$, $L_k C_k v \perp \ker\{P(k + 1)\}$.
(ii) $H_{k + 1} L_k C_k = 0$, where $H_k \in \mathbb{R}^n$, $k \geq 0$, represents the orthogonal projection onto $\ker\{X(k)\}$.

**Proof.** Statement (i) follows immediately from Lemma 5(ii) and Corollary 2. Regarding (ii), since $H_{k + 1}$ is the orthogonal projection onto $\ker\{X(k + 1)\}$, as stated in Corollary 2, we have that $\ker\{H_{k + 1}\} \perp \ker\{P(k + 1)\}$. On the other hand, the statement (i) leads to $L_k C_k v \perp \ker\{P(k + 1)\}$. Then, $L_k C_k v \perp \ker\{H_{k + 1}\}$, and $H_{k + 1} L_k C_k v = 0$; hence we have the claim. □
Corollary 3 provides a link between the dynamics of system $\Theta$ defined in (12) and the homogeneous part of (8), namely,

\begin{equation}
Z(k+1) = (H_{k+1}A_k)Z(k)(H_{k+1}A_k)' = [H_{k+1}(A_k - L_kC_k)]Z(k)[H_{k+1}(A_k - L_kC_k)]',
\end{equation}

which leads to the evaluations involving $Z$ and $\hat{X}_\xi$ that are presented in the next lemma.

**Lemma 6.** The following inequalities hold for any $\epsilon \neq 0$.

(i) $\hat{X}_\xi(k,0,E,I) \leq \hat{X}_\xi(k,0,E,0) + \xi^{2k}(1 + \epsilon^{-2})k+1Z(k)$

\[ + \xi^{2k}(1 + \epsilon^2)\bar{X}_h(k-1,0,(I - H_0)I(I - H_0)'), \]

\[ + \xi^{2k}(1 + \epsilon^2)\sum_{\ell=1}^{k}((1 + \epsilon^{-2})^\ell \]

\[ \times \bar{X}_h(k-1,\ell,(I - H_\ell)\bar{X}_h(\ell - 1,\ell - 1,Z(\ell - 1))(I - H_\ell)'), \]

(ii) $\hat{X}_\xi(k,0,E,I) \geq \hat{X}_\xi(k,0,E,0) + \xi^{2k}(1/2)(1 - \epsilon^{-2})kZ(k)$

\[ - \xi^{2k}\bar{X}_h(k-1,0,(I - H_0)I(I - H_0)'), \]

\[ - \xi^{2k}(\epsilon^2 - 1)\sum_{\ell=1}^{k}((1 - \epsilon^{-2})^\ell \]

\[ \times \bar{X}_h(k-1,\ell,(I - H_\ell)\bar{X}_h(\ell - 1,\ell - 1,Z(\ell - 1))(I - H_\ell)'), \]

**Proof.** See Appendix D.

**Theorem 1.** Consider Assumptions 1 and 2. The KF is semistable if and only if Condition 1 holds.

**Proof (sufficiency).** For each $0 \leq \xi < 1$ set $\epsilon > 0$ and $0 \leq \xi_Z < 1$ in such a manner that

$$\xi^2\xi_Z^{-2}(1 + \epsilon^{-2}) < 1.$$ 

We start by taking into account separately each term on the right-hand side of statement (i) of Lemma 6. For the first term Proposition 2(iii) yields

\begin{equation}
\hat{X}_\xi(k,0) \leq \hat{X}_\xi(k,\Sigma) \leq P(k) \leq \hat{X}.
\end{equation}

For the second term it is simple to check from Condition 1 (with $\xi$ replaced by $\xi_Z$) that

\begin{equation}
(\xi^2(1 + \epsilon^{-2}))^k(1 + \epsilon^{-2})Z(k) \leq (\xi_Z^2\xi^2(1 + \epsilon^{-2}))^k(1 + \epsilon^{-2})\bar{Z} \leq (1 + \epsilon^{-2})\bar{Z}.
\end{equation}

For the third term, since $H_0$ is the projection onto $\ker\{P(0)\}$, we have that $\ker\{(I - H_0)I(I - H_0)\} \supset \ker\{P(0)\}$, yielding, from Lemma 4,

\begin{equation}
\xi^{2k}(1 + \epsilon^2)\bar{X}_h(k-1,0,(I - H_0)I(I - H_0)'), \]

\[ \leq \xi^{2k}(1 + \epsilon^2)\xi^{-2k}\|(I - H_0)I(I - H_0)\|\bar{X} \leq (1 + \epsilon^2)\bar{X}.
\end{equation}

For the last term, Condition 1 with $\xi$ replaced by $\xi_Z$ and Proposition 2(ii), (ii) lead to $\|\bar{X}_h(\ell - 1,\ell - 1,Z(\ell - 1))\| \leq \|\bar{X}_h(\ell - 1,\ell - 1,\xi_Z^{-2}\bar{Z})\| \leq \xi_Z^{-2\delta}\|\bar{Z}\|$, $\ell = 1, \ldots, k$, and

\[ \|\hat{X}_h(\ell - 1,\ell - 1,Z(\ell - 1))(I - H_\ell)' \| \leq \xi_Z^{-2\delta}\|\bar{Z}\|.
\]
Then, we employ Lemma 4 (note that \(\ker\{(I - H_\ell)\tilde{X}_h(\cdot)(I - H_\ell)'\}\supset\ker\{P(\ell)\}\)) with \(\xi\) replaced by \(\xi_Z\), and the above inequality, respectively, to evaluate

\[
(1 + \epsilon^2)\xi^2 \sum_{\ell=1}^{k} [(1+\epsilon^{-2})^\ell \times \tilde{X}_h(k - 1, \ell, (I - H_\ell)\tilde{X}_h(\ell - 1, \ell - 1, Z(\ell - 1))(I - H_\ell)')]
\leq (1 + \epsilon^2)\xi^2 \sum_{\ell=1}^{k} [(1+\epsilon^{-2})^\ell \xi_Z^{-2(k-1-\ell)} \times \|(I - H_\ell)\tilde{X}_h(\ell - 1, \ell - 1, Z(\ell - 1))(I - H_\ell)'\|X]
\leq (1 + \epsilon^2)\xi^2 \sum_{\ell=1}^{k} [(1+\epsilon^{-2})^\ell \xi_Z^{-2(k-1-\ell)} \xi_Z^{-2\ell} \|Z\|\tilde{X} \|X\|\xi/\xi^2 Z)^2 k \sum_{\ell=1}^{k} (1+\epsilon^{-2})^\ell
\leq (1 + \epsilon^2)\xi^2 \|Z\|\tilde{X} \|X\| k \sum_{\ell=1}^{k} (\xi^2\xi_Z^{-2}(1+\epsilon^{-2})^\ell \leq \zeta \tilde{X},
\]

where we have defined \(\zeta = (1 + \epsilon^2)\xi^2 \|Z\|q(1 - q)^{-1}\) with \(q = \xi^2\xi_Z^{-2}(1+\epsilon^{-2}) < 1\).

Substituting (27)–(30) in (i) of Lemma 6 we obtain, for \(k \geq 0\),

\[
\tilde{X}_\xi(k, 0, E, I) \leq \tilde{X} + (1 + \epsilon^{-2})\tilde{Z} + (1 + \epsilon^2)\tilde{X} + \zeta \tilde{X},
\]

and Lemma 1 leads to the result.

\textit{Necessity.} Assuming that the KF is semistable, we have in particular that for each \(0 \leq \xi < 1/2\) there exists \(X\) such that \(X_\xi(k, 0, E, I) \leq \tilde{X}, k \geq 0\), and from Lemma 6(ii), with \(\epsilon > 1\) such that \(\xi^2(1 - \epsilon^{-2}) < 1\), we obtain

\[
\tilde{X}_\xi(k, 0, E, I) \leq \tilde{X} + (1 + \epsilon^{-2})\tilde{Z} + (1 + \epsilon^2)\tilde{X} + \zeta \tilde{X},
\]

Similarly to the proof of the sufficiency, one can check that the third and fourth terms on the right-hand side of (32) are bounded from below by \(-\tilde{X}\) and \(-\eta \tilde{X}\), where \(\eta\) is set similarly to \(\zeta\) in the proof of the sufficiency. Note that \(\tilde{X}_\xi(k, 0, E, 0) \geq 0\). These elements can be combined to obtain \(\tilde{X} \geq (1/2)\xi^2k(1 - \epsilon^{-2})^k Z(k) - \zeta \tilde{X} - \eta \tilde{X}\), and setting \(\epsilon = 2\) and \(\xi_Z = 2\xi\), we have that for each \(0 \leq \xi_Z < 1\), \(Z(k) \leq 2\xi_Z^{-2k}(2 + \eta)\tilde{X}\), and hence Condition 1 holds.

\section{LTI systems.}

In this section we present simple conditions to test semistability, stability, and stability w.r.t. \(\Psi\) of the KF in the LTI scenario with \(A = \{A_i : A_i = A_0, i \in Z\}\), and similarly for \(B, C, D,\) and \(E\). For ease of notation, we denote \(A_0\) by \(A\), and we do the same for the other matrices. We need the following result
from [11]. For $U, W \in \mathcal{R}^n$ and $V \in \mathcal{R}^{n_0}$, consider the system

$$
\begin{align*}
Z(k + 1) &= H_k W Z(k) W' H_k', \\
X(k + 1) &= W X(k) W' + U U', \\
(Z(0), X(0)) &= (H_0 H_0', V),
\end{align*}
$$

(33) \quad \Theta(U, V, W) :

and denote its state with $(Z(k, V, W, U), X(k, V, W, U))$ to emphasize the dependence on the matrices $U$, $V$, and $W$.

**Proposition 3.** Consider the system $\Theta(0, \Sigma, A)$ and the corresponding variable $Z(k, 0, \Sigma, A)$. Let $J$ be such that $J A J^{-1}$ is in Jordan form, and let $\mathcal{J}$ stand for the unstable space of $J A J^{-1}$. For each $0 \leq \xi Z < 1$ there exists $\bar{Z} \in \mathcal{R}^{n_0}$ such that

$$
\xi^2 Z(k) \leq \bar{Z}, \quad k \geq 0, \quad \text{if and only if}
$$

$$
\ker\{J \Sigma J'\} \cap \mathcal{J} = \{0\}.
$$

(34)

Proposition 3 represents a simple way of testing the condition of Theorem 1, valid for the particular case when $E = 0$. In order to incorporate $E \neq 0$, we consider the orthogonal projection onto the noncontrollable subspace associated with the pair $(A, E)$,

$$
H = I - (C C')^* C C',
$$

(35)

where $C = [E \quad A E \quad \cdots \quad A^{n-1} E]$ is the controllability matrix associated with $(A, E)$, and we explore the auxiliary systems $\Theta(C, \Sigma, A)$ and $\Theta(0, \Sigma, H A)$. System $\Theta(0, \Sigma, H A)$ matches the setup of Proposition 3 with $A$ replaced by $H A$, and if we let $J_H$ be such that $J_H H A J_H^{-1}$ is in Jordan form and let $\mathcal{J}_H$ be the unstable subspace of $J_H H A J_H^{-1}$, then (34) reads as follows.

**Condition 2.** $\ker\{J_H \Sigma J'_H\} \cap \mathcal{J}_H = \{0\}$.

**Remark 5** (equivalent forms of Condition 2). Condition 2 is not equivalent to $\ker\{\Sigma\} \cap \mathcal{N} = \{0\}$, where $\mathcal{N}$ stands for the unstable subspace of $H A$. For example, consider

$$
A = \begin{bmatrix} 2 & 1 \\
0 & 1 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 0 & 0 \\
0 & 1 \end{bmatrix}, \quad E = 0, \quad J_H H A J_H^{-1} = \begin{bmatrix} 1 & 0 \\
0 & 2 \end{bmatrix}, \quad J_H \Sigma J'_H = \begin{bmatrix} 1 & 1 \\
1 & 1 \end{bmatrix}.
$$

We have $\ker\{\Sigma\} \cap \mathcal{N} \neq \{0\}$, but $\ker\{J_H \Sigma J'_H\} \cap \mathcal{J}_H = \{0\}$. This is due to the fact that, considering a rank-one decomposition of the form $\Sigma = \sum_{i=1}^n \sigma_i \sigma_i'$, we may have $A^k \sigma_i$ orthogonal to the unstable space of $A$ for some values of $k$ ($k = 0$ in the numerical example, with $\sigma_1 = [0 \quad 1]^T$) and nonorthogonal for other $k$ ($k \geq 1$ in the example); this difficulty is circumvented by considering the unstable space of $A'$: orthogonality of $A^k \sigma_i$ to this space holds for all $k \geq 0$ if it holds for $k = 0$. This can be employed to show that Condition 2 is equivalent to

$$
\ker\{\Sigma\} \cap \mathcal{L} = \{0\},
$$

where $\mathcal{L}$ stands for the unstable space of $A' H'$. Now it is easy to see that this condition appears in [6, Theorem 1] (in the context of the dual LQ-optimal control problem for continuous-time systems) and also in [15, Theorem 3] and [21, Corollary II.1], with a slightly different notation, as a condition for $P(k)$ to converge to a strong solution, assuming detectability of $(A, C)$. Condition 2 can be expressed by $\ker\{\Sigma\} \cap \ker\{C\} \cap \mathcal{L} = \{0\}$, where $\mathcal{L}$ stands for the unstable subspace of $A'$; note that $E$ is accounted
for through $C$ (and it is well known that $\ker\{E\} \supset \ker\{C\}$) whereas $\Sigma$ is directly accounted for. For instance, for $V \in \mathbb{R}^{n_0}$, it is more likely to satisfy the condition setting $\Sigma = 0$ and $E = V$ than $\Sigma = V$ and $E = 0$; see Example 5. One interpretation is that unstable modes of $A$ should be completely excited by nonpersistent noise and/or excited by persistent noise, where we say, respectively, “completely excited” and “excited” to emphasize how $\Sigma$ and $E$ are counted for.

The relations that provide the necessity of Condition 2 for the semistability of the KF are presented in the next lemma.

**Lemma 7.** Consider the systems $\Theta(E, \Sigma, A)$, $\Theta(C, \Sigma, A)$, and $\Theta(0, \Sigma, HA)$ and the associated state trajectories. Then $Z(k, E, \Sigma, A) \geq Z(k, C, \Sigma, A) = Z(k, 0, \Sigma, HA)$, $k \geq 0$.

**Proof.** See Appendix E.

The ordering $Z(k, E, \Sigma, A) \geq Z(k, C, \Sigma, A)$ in Lemma 7 is a consequence of the (easy to check) fact that $CC' \geq EE'$. A converse relation evidently does not hold, making the proof of sufficiency of Condition 2 more complex. We have to handle transients in $X$ that take place when dealing with general $E$, and we need to introduce the system $\Theta(C, A'\Sigma A'^n, A)$, which allows us to employ the following invariance result.

**Proposition 4.** Assume $U \in \mathbb{R}^n$ is in Jordan form, and let $J$ be the unstable space of $U$. If $\Sigma \in \mathbb{R}^{n_0}$ satisfies $\ker\{\Sigma\} \cap J = \{0\}$, then $\ker\{U^k\Sigma U^k\} \cap J = \{0\}$, $k \geq 0$.

**Lemma 8.** Consider the systems $\Theta(E, \Sigma, A)$ and $\Theta(0, \Sigma, HA)$ and the associated state trajectories. For each $0 \leq \xi < 1$ assume that there exists $\tilde{Z}$ such that $Z(k, 0, \Sigma, HA) \leq \xi^{2k} \tilde{Z}$. Then there exists $\tilde{Z}$ such that $Z(k, E, \Sigma, A) \leq \xi^{2k} \tilde{Z}$.

**Proof.** See Appendix F.

We are ready to present the main result. Recall from Remark 2 that Assumption 2 holds trivially in the present context.

**Theorem 2.** Consider Assumption 1. The KF is semistable if and only if Condition 2 holds.

**Proof (sufficiency).** The condition of Proposition 3 holds with $A$ replaced by $HA$; thus for each $0 \leq \xi < 1$ there exists $\tilde{Z} \in \mathbb{R}^{n_0}$ such that $\xi^{2k} Z(k, 0, \Sigma, HA) \leq \tilde{Z}$, $k \geq 0$. Then, Lemma 8 provides the fact that there exists $\tilde{Z}$ such that $Z(k, E, \Sigma, A) \leq \xi^{2k} \tilde{Z}$, and Theorem 1 leads to the result.

**Necessity.** From Theorem 1 we have that for each $0 \leq \xi < 1$ there exists $\tilde{Z}$ such that $Z(k, E, \Sigma, A) \leq \xi^{2k} \tilde{Z}$. Lemma 7 immediately leads to $Z(k, 0, \Sigma, HA) \leq \xi^{2k} \tilde{Z}$, and Proposition 3 with $A$ replaced by $HA$ yields the claim.

Consider the standard conditions that $(A, C)$ is detectable and that the modes of $A$ unreachable by $E$ do not lie on the unit circle. These conditions arise when studying existence, convergence, and stabilizability of solutions for the limiting Riccati equation; see, e.g., [5, 12, 13, 17, 21]. Different terminology and notation can be found in the literature, and we mention that the semistabilizing solution of the limiting Riccati is usually referred to as the strong solution, and that the above condition involving the unreachable modes of $A$ is equivalent to requiring that $A$ has no left eigenvector $z$ corresponding to an eigenvalue on the unit circle such that $zE = 0$, as appearing in [24], or that $S_H = \{0\}$ where $S_H$ stands for the strictly semistable space of $J_H HA J^-1_H$; we employ the latter notation because it is concise (and compatible with previous notation). If we assume detectability and $S_H = \{0\}$ in addition to Condition 2, then stability of the KF is obtained. The sufficiency of this stronger condition follows by combining the next available results.
(i) Condition 2 is a necessary and sufficient condition for $P(k)$ to converge to a strong solution (see [6, Theorem 1]) in the context of the dual LQ-optimal control problem for continuous-time systems;

(ii) the strong solution is stabilizing if and only if $S_H = \{0\}$ (see, e.g., [5, 14, 17, 21]);

(iii) if $P(k)$ converges to a stabilizing solution, then $\tilde{X}(k)$ is bounded; i.e., the KF is stable [24, Theorem 1]. This stronger condition for stability is also a necessary one. In fact, necessity of detectability is obvious. The necessity of $S_H = \{0\}$ is shown in [24, Theorem 4.2] assuming $P_k$ converges (and convergence of $P_k$ is ensured by Condition 2, as stated, for instance, in [6, Theorem 1]). It remains to show the necessity of Condition 2; in principle, if it does not hold, then $P(k)$ does not converge to the strong solution of the ARE, but it is still needed to show that $\tilde{X}$ diverges. The results on divergence of $\tilde{X}$ in [24, Theorems 4.1 and 4.2] do not give the complete picture, as illustrated in Example 5. Fortunately, necessity of Condition 2 follows trivially from Theorem 2; otherwise the filter is not even semistable. This yields the following characterization of stability of the KF. Recall from Remarks 1 and 2 that Assumptions 1 and 2 hold for detectable LTI systems.

**Theorem 3.** The KF is stable if and only if $(A, C)$ is detectable, Condition 2 holds, and $S_H = \{0\}$. $S_H = \{0\}$ further requires the strictly semistable spaces of $A$ to be excited by $E$. Indeed, to obtain stability it is not enough to excite semistable modes via $\Sigma$, since the noise may vanish as time evolves and the KF may tend to a semistable one, as illustrated in Example 1. The situation in which some of the semistable modes of $A$ are excited only by $\Sigma$ yields stability w.r.t. $\Psi$. The proof of this fact involves several adaptations of the results in [9, 11] and therefore is not presented.

**Proposition 5.** If $\ker\{J_H \Sigma J_H^T \} \cap (J_H \cup S_H) = \{0\}$, then the KF is stable w.r.t. $\Psi$.

**Remark 6.** The conditions of Theorems 1–3 and Proposition 5 can be employed in the problem of stabilization of the KF. For instance, assuming $E = 0$, we have that $\Sigma = \sigma_1 \sigma_1^T + \cdots + \sigma_r \sigma_r^T$, with $r = \dim(J_H)$, is such that the KF is semistable if and only if $\sigma_j$, $0 \leq j \leq r$, are linearly independent vectors with nontrivial projections onto $J_H$, and a similar condition holds for stability and stability w.r.t. $\Psi$. As an illustration, for the system $\Theta$ of Example 1, for (strict) semistability we set $\Sigma = 0$, for stability w.r.t. $\Psi$ we set $\Sigma = I$, and for stability we set $E = [0 \ 0 \ 1]^T$.

The links with classical conditions for stability of the KF are as follows.

**Lemma 9.** For LTV systems, the following statement holds.

(i) Consider Assumptions 1 and 2. $(A, E)$ uniformly stabilizable implies Condition 1.

For LTI systems, the following statements hold.

(ii) $(A, E)$ stabilizable (in the usual sense for LTI systems) implies $J_H = S_H = \{0\}$.

(iii) $(A, E)$ semistabilizable (in the sense that the modes of $A$ unreachable from $E$ are semistable) implies $J_H = \{0\}$.

(iv) $\Sigma > 0$ implies $\ker\{J_H \Sigma J_H^T \} = \{0\}$.

**Proof.** (i) We have from [3, Theorems 4.3 and 5.3], under Assumption 1, that $(A, E)$ uniformly stabilizable implies that the KF is stable, and Theorem 1 completes the proof.

(ii) Provided $(A, E)$ is stabilizable, it is a straightforward matter to check that $HA$ is a stable matrix, recalling that the projection $H$ “cancels” controllable dynamics of...
A, and stabilizability of \((A, E)\) yields that the remaining dynamics are stable. Thus, \(J_H \cup \mathcal{S}_H = \{0\}\). In particular, for controllable \((A, E)\) one can easily check that \(H = 0\).

(iii) Similarly as above, semistabilizable \((A, E)\) implies semistable \(HA\); hence \(J_H = \{0\}\).

(iv) \(\Sigma > 0\) trivially implies \(\ker\{J_H \Sigma J_H'\} = \{0\}\). □

6. Illustrative examples.

Example 2 (Example 1 continued). Consider the system \(\Phi\) in Example 1. Direct evaluation of \(\tilde{X}\) in Example 1 has shown that the KF is not stable. Since \((A, C)\) is detectable and the system is LTI, Assumptions 1 and 2 hold (see Remarks 1 and 2) and stability w.r.t. \(\Psi\) follows from Corollary 1, since \(\Sigma = I\), or alternatively from Proposition 5 and Lemma 9. Moreover, \(\Sigma > 0\) yields \(H_0 = 0\), and from (12) it follows that \(Z(k) = 0, k \geq 0\); thus Condition 1 holds and semistability is confirmed by Theorem 1; \(\Sigma > 0\) also implies Condition 2, and Theorem 2 leads to the same result.

Example 3 (absolute and relative estimation error). Consider the system \(\Phi\) with

\[
A_i = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, \quad C_i = D_i = I, \quad \Psi = 10^{-2}I, \quad B_i = 10^{-3}I.
\]

We consider \(\tilde{x} = [1 \quad 0]'\) and \(\hat{x}(0) = \tilde{x}\). Assumptions 1 and 2 hold. Consider the following setups.

(i) \(\Sigma = \sigma\sigma'\) with \(\sigma = [0 \quad 1]'\) and \(E_i = 0, i \geq 0\). Condition 2 holds and Theorem 2 states that the KF is semistable. The actual error \(\tilde{X}(k, \Psi, B)\) presents a slow, linear divergence, e.g.,

\[
\tilde{X}(100) \approx \begin{bmatrix} 3.040 & 0.020 \\ 0.020 & 0.0101 \end{bmatrix} \quad \text{and} \quad \tilde{X}(500) \approx \begin{bmatrix} 3.042 & 0.020 \\ 0.020 & 0.0105 \end{bmatrix}.
\]

The relative error \(E = \{\|\tilde{x}(k)\|^2\}/E = \{\|\hat{x}(k)\|^2\}\) converges exponentially to zero. In a typical realization we have obtained \(x(100) \approx [5.699 \times 10^{29} \quad 0.954]'\) and \(\hat{x}(100) \approx [5.699 \times 10^{29} \quad 0.899]'\), a fairly good estimate.

(ii) \(\Sigma = E_i = 0\). The conditions of Theorem 2 do not hold and the KF is not semistable: \(\tilde{X}(k, \Psi, E)\) diverges exponentially (e.g., now \(\|\tilde{X}(100)\| \approx 8.035 \times 10^{57}\)). As for the relative error we obtained \(\lim_{k \to \infty} E = \{\|\tilde{x}(k)\|^2\}/E = \{\|\hat{x}(k)\|^2\} \approx 0.020\). A typical realization provides \(x(100) \approx [7.299 \times 10^{29} \quad 1.075]'\) and \(\hat{x}(100) \approx [6.338 \times 10^{29} \quad 1.000]'\), which presents absolute error of order \(10^9\) but is still reasonable in terms of relative error. Figure 2 illustrates trajectories associated with the above realizations, making clear how they differ; we present “scaled” trajectories because of the large quantities involved and because they are associated with the relative error.

(iii) \(\Sigma = 0\) and \(E_i = \sigma\sigma', i \geq 0\), with \(\sigma\) as above. The conditions of Theorem 3 hold and the filter is stable. However, we have

\[
\tilde{X}(100) \approx \begin{bmatrix} 3.848 & 0.562 \\ 0.562 & 0.416 \end{bmatrix},
\]

representing a poor estimate when compared to (i). We observed that the estimates get better than the ones in (i) at much larger instants (e.g., \(k \approx 10^6\)). In fact, the filter performance may be more sensitive to a high magnitude of \(E - B\) (related to the forcing terms of (6) and (8)), as in (iii), than to a high magnitude of \(\Sigma - \Psi\) (related to the initial condition of (6) and (8)), as in (i).

Example 4 (polynomial divergence). Consider the system \(\Phi\) of Example 3 with \(A_i(1,1) = 1, B_i = \Psi = I, i \geq 0\). Figure 3 illustrates the norm of the actual
error covariance of the KF for different values of $\Sigma (0, 10^{-3}I, I, 10^3I)$. The graphs highlight that choosing $\Sigma$ is not an efficient tool for reducing the degree of polynomial divergence; one needs to choose an appropriate $E$.

**Example 5** (noise excitation: $\Sigma$ versus $E$). Consider the system $\Phi$ with

$$A_i = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \quad C_i = D_i = I, \quad i \geq 0.$$  

Consider the following setups. (i) $\Sigma = I$ and $E_i = 0$, $i \geq 0$. Assumptions 1 and 2 hold (since $(A_i, C_i)$ is detectable and the system is LTI; see Remarks 1 and 2). It is simple to check that the conditions of Theorem 3 hold, and hence the KF is stable. (ii) $\Sigma = \sigma^t \sigma$ with $\sigma = \begin{bmatrix} 0 & 1 \end{bmatrix}$ and $E_i = 0$, $i \geq 0$. We have that $C = 0$, $H = I$, $HA$ is in Jordan form, and $J_H = I$, leading to $\ker(\Sigma) \cap \mathbb{R}^2 = [\mu]$ where $\mu = [1 \ 0]^t$, and thus Condition 2 does not hold and the KF is not semistable (despite $(A_i, \Sigma)$ being stabilizable). The conditions for error divergence in [24, Theorems 4.1 and 4.2] do not hold (there is no eigenvector of $A'$ in the null space of $\Sigma'$, nor on the unit circle), illustrating the conservativeness of those results. (iii) $\Sigma = 0$ and $E_i = \sigma^t \sigma$, $i \geq 0$, with $\sigma = \begin{bmatrix} 0 & 1 \end{bmatrix}$. Stabilizable $(A, E)$ is enough to provide stability: $C = I$, $H = 0$,
and $HA = 0$, yielding $J_H = \{0\}$ (confirming Lemma 9) and Condition 2. Note that $\Sigma$ in (ii) equals $E$ in (iii); this illustrates in what sense $\Sigma$ has to completely excite the unstable spaces of $A$, whereas $E$ has only to excite them; see Remark 5.

Example 6 (LTI system with periodic RDE solution). Consider the system $\Phi$ with

$$A_i = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, B_i = C_i = 0, D_i = I, i \geq 0.$$ 

Implementing the KF, one can check that $P(k)$ is periodic (and available results on convergence of $P$ cannot be employed). Thus $P(k)$ is bounded and satisfies Assumption 1, although $(A,C)$ is not detectable. Assumption 2 holds; see Remark 2. It is simple to check that $Z(k)$ is bounded (in fact, $Z(k) = P(k + 1)$), and hence Condition 1 holds and the KF is semistable, as stated in Theorem 1. Condition 2 also holds, and Theorem 2 confirms semistability. This illustrates that detectability of $(A,C)$ and convergence of $P$ are not required for semistability of the KF.

Example 7. Consider the system $\Phi$ with

$$A_i = 2 - 0.99 \cos\left(\frac{i}{1 + (i/1000)}\right), E_i = 0, \Sigma = B_i = C_i = D_i = 1, i \geq 0.$$ 

Assumptions 1 and 2 hold with $\bar{X} = 8, M = 1, T = 1, \text{and } 1.01^{-1} < \tau < 1$. Since $\Sigma > 0$, we have that $Z(k) = 0$ and Condition 1 holds, ensuring that the KF is semistable. See Figure 4 for the behavior of $\bar{X}(k,B,\Sigma)$. Now, modify the system in such a manner that $A_i$ does not present a uniform minimal “margin of instability” (thus violating Assumption 2; see Remark 2) by replacing $A_i$ with $A_{M,i} = 2 - (1 - e^{-i/100})\cos(i(1 + (i/1000))^{-1})$. See the associated $\bar{X}_M(k,B,\Sigma)$ in Figure 4. Despite $A_{M,i}$ getting closer to one than $A_i$ for certain $i \geq 0$ (the “margin of instability” is reduced), $\bar{X}_M$ diverges faster than linearly, which is surprising in the sense that, for fixed unidimensional $A_i = A$, $\bar{X}$ can diverge at most linearly.

7. Concluding remarks. In this paper we have explored the structure of the KF from the perspective of divergence of the actual error covariance $\bar{X}(k)$ under incorrect noise measurements, assuming that the calculated error covariance $P(k)$ is bounded. We have shown that stability w.r.t. $\Psi$ (accommodation of incorrect $\Sigma$, exclusively) is weaker than stability, whereas semistability makes no distinction between imprecise $\Sigma$, $E$, or both; see Lemma 1. This result and some structural properties of
the KF (as the orthogonality property of the Kalman gain expressed by the condition $H_{k+1}L_kC_k = 0$) allow us to derive Condition 1 and to show that it is necessary and sufficient for semistability of the KF. Semistability is equivalent to avoid exponential divergence of the actual error covariance \( \tilde{X}(k) \), and it ensures that, when \( X(k) \) diverges exponentially, the actual relative error \( E\{\|\tilde{z}(k)\|^2\}/E\{\|\tilde{z}(k)\|^2\} \) converges exponentially as the time evolves; see Remark 4. Condition 1 involves the relevant data \( A, E, \) and \( \Sigma \) and involves neither conditions on \( C \) (hence Condition 1 is valid for nondetectable systems) nor calculations of RDEs. These results provide the elements for obtaining the algebraic necessary and sufficient Condition 2 for semistability of the KF in the LTI context, with the interpretation that the nonpersistent noise (characterized by \( \Sigma \)) has to completely excite unstable modes of \( A \) that are not already excited by \( E \); see Remark 5 and the illustrative Example 5. The result is valid independently of the convergence of \( P(k) \) or any conditions on \( C \), hence clarifying the meaning of Condition 2 for nondetectable systems (or nonstabilizable systems in the dual control scenario), complementing available results for detectable systems [6]. Still in the LTI context, we combine Condition 1 with some available results to obtain a necessary and sufficient condition for stability of the KF, as well as a condition for stability w.r.t. \( \Psi \), and we compare these conditions with classical ones in Lemma 9. The above results give a rather complete picture of conditions for semistability and stability for the KF and can be employed in the problem of stabilization of the KF via a suitable choice of \( \Sigma \) and \( E \), for both stability and semistability; see Remark 6.

**Appendix A. Proof of Proposition 2.**

(i) From (5), Assumption 1, and recalling that we assume \( DD' \geq \chi I \) for some \( \chi > 0 \), we have \( \|L_k\| = \|AP(k)C'CP(k)C' + DD'\|^{-1} \leq \|A\| \|\tilde{X}\| \|C\| \chi^{-1} \).

(ii) Employing Proposition 1 and (i) above yields \( \|\tilde{X}(k+1,1,\Sigma)\| \leq \|(A-L_kC)\Sigma (A-L_kC)' + \|L_kDD'L_k' + BB'\| \leq \|(A+\rho\|C\|)^2 \|\Sigma\| + \rho^2 \|DD'\| + \|BB'\| \).

(iii) We proceed inductively. For \( k = 0 \) the result is immediate. Now assume that the assertion of the proposition holds true for \( k > 0 \), in such a manner that \( \tilde{X}(k,\alpha\Sigma_1) - \alpha\tilde{X}(k,\Sigma_0) \geq 0 \). This and Proposition 1 yield

\[
\tilde{X}(k+1,\alpha\Sigma_1) - \alpha\tilde{X}(k+1,\Sigma_0) = (A-L_kC)\tilde{X}(k,\alpha\Sigma_1)(A-L_kC)' + L_kDD'L_k' + BB' - \alpha(A-L_kC)\tilde{X}(k,\Sigma_0)(A-L_kC)' - \alpha L_kDD'L_k' - \alpha BB' = (A-L_kC)(\tilde{X}(k,\alpha\Sigma_1) - \alpha\tilde{X}(k,\Sigma_0))(A-L_kC)' + (1-\alpha)L_kDD'L_k' + (1-\alpha)BB' \geq 0.
\]

(iv) We proceed inductively. The statement is trivial for \( k = 0 \). Assuming \( \tilde{X}(k,U,V) \leq \xi^{-2k}X_\xi(k,U,V) \), from (8) and (13) we obtain \( \tilde{X}(k+1,U,V) = (A_k - L_kC_k)(\tilde{X}(k,U,V) - \xi^{-2k}X_\xi(k,U,V))(A_k - L_kC_k)' \leq 0 \). The result is now immediate, since \( X_\xi(\cdot) \leq \xi^{-2k}X_\xi(\cdot) \).

(v) The result follows from \( \Gamma T' \geq 0 \) with \( \Gamma = \epsilon HV^{1/2} + \epsilon^{-1}(I-H)V^{1/2} \).

(vi) For the first statement, let \( v \) be a semistable eigenvector of \( A((k+t)T,tT) \) (if no such eigenvector exists, then \( M = I \) and the statement holds trivially). We need to show that \( X_h((k+t)T,tT,vv') \leq \xi^{-2kT}vv'\|\tilde{X}\| \) for some \( \tilde{X} \). The linear dependence of the bound on \( vv' \) follows from the linear dependence of \( \tilde{X}_h \) on \( vv' \). Then we can assume without loss of generality that \( vv' = 1 \). If \( \tilde{X}_h(\ell,tT,vv') = (I-M)\tilde{X}_h(\ell,tT,vv')(I-M)' \), \( 0 \leq \ell \leq (k+t)T \), one can show that \( \tilde{X}_h(\ell,tT,vv') \leq \xi^{-2\ell}\tilde{X} \), where the first inequality involves the filter optimality and the second inequality is immediate from the fact that \( v \) is a semistable eigenvector of \( A((k+t)T,tT) \); otherwise, more complex evaluations are required, involving the
form \( \bar{X}_h(\ell, t T, \nu') \leq (1 + \epsilon)M \bar{X}_h(\ell, t T, \nu')M' + (1 + \epsilon^{-1})(I - M)\bar{X}_h(\ell, t T, \nu')(I - M') \) and showing that \( P_t \) is such that its null space is a subset of the null space of \( M \bar{X}_h(\ell, t T, \nu')M' \); see the details in [8]. Now consider the second statement. Consider the case \( T = 1 \). We start by showing inductively that there exists \( \mu > 0 \) such that

\[
(37) \quad \bar{X}(k, 0, 0, I) \geq \mu M M', \quad k \geq 0.
\]

Equation (37) is trivial for \( k = 0 \). Assuming this inequality holds for \( k > 0 \), we obtain

\[
\bar{X}(k + 1, 0, 0, I) - L_k D_k D_k' L_k' = (A_k + L_k C_k) \bar{X}(k, 0, 0, I)(A_k + L_k C_k)' \geq \mu (1 - \epsilon^2)A_k M M' A_k' - \mu(\epsilon^2 - 1)L_k C_k M M' C_k' L_k' \geq \mu (1 - \epsilon^2)\tau^{-1}M M' - \mu(\epsilon^2 - 1)L_k C_k M M' C_k' L_k'.
\]

Since \( L_k C_k M M' C_k' L_k' \leq \beta L_k D_k D_k' L_k' \) for some \( \beta > 0 \), we obtain for sufficiently small \( \epsilon \) and \( \mu \) that \( \bar{X}(k + 1, 0, 0, I) \geq \mu M M' \), completing the induction. The proof for the second statement of (vi) is inductive. For ease of notation, we set \( A_k = A_k + L_k C_k, k \geq 0 \). The result is immediate for \( k = 0 \). Assuming it holds for \( k > 0 \) and employing (37), we obtain \( \bar{X}(k + 1, 0, 0, I) - L_k D_k D_k' L_k' = \delta A_k \bar{X}(k, 0, 0, I)A_k' \geq \delta^{k+1}A_k \bar{X}(k, 0, 0, I)A_k' \geq (1 - \epsilon^2)(\delta - 1)\mu A_k M M' A_k' - (\epsilon^2 - 1)(\delta - 1)\mu L_k C_k M M' C_k' L_k'. \) Using the fact that \( A_k M M' A_k' \geq \tau^{-1}M M' \) and defining \( \beta \) as above, for sufficiently small \( \epsilon \) and \( \mu \) we evaluate \( \bar{X}(k + 1, 0, 0, I) - L_k D_k D_k' L_k' \geq \delta^{k+1}A_k \bar{X}(k, 0, 0, I)A_k' \geq (1 - \epsilon^2)(\delta - 1)\mu A_k M M' A_k' \), hence we have the claim. See the details and the case \( T > 1 \) in [8]. Finally, consider the third statement. Assume \( E = 0 \). Let the eigenvectors of \( M \Sigma M' \) be denoted by \( m_i \) and be ordered in such a manner that \( m_0, \ldots, m_{\ell} \), corresponding to eigenvalues different from zero. For each \( m_i \) we pick \( v_{i,t} \) in the unstable space of \( A'(t T, 0) \) such that \( M' m_i = A'(t T, 0) v_{i,t} \) (existence and uniqueness of \( v \) follow from the fact that \( M' m_i \) is in the unstable space of \( A((k + 1)T, k T) \), \( k \geq 0 \), by Assumption 2; one can also check that \( \|v_{i,t}\| < \tau \|M m_i\| \)). First we show that \( v_{i,t} M P_{TT} M' v_{i,t} \geq \nu \|v_{i,t}\|^2 \), \( 0 \leq t \leq \ell \), for some \( \nu > 0 \) and all \( t \geq 0 \). To prove this fact we establish a contradiction. Assume that for each \( \nu > 0 \) there exists \( t \geq 0 \) such that \( v_{i,t} M P_{TT} M' v_{i,t} < \nu \|v_{i,t}\|^2 \). From a control problem perspective (adjoint to the filtering problem considered in this paper) we have that, for the cost functional \( J^T(v) = \inf_{u} [\sum_{t=0}^{T-1} \|u_k\|^2 + \sum_{t=k}^{T-1} z_k E_k' E_k z_k + u_k D_k' D_k u_k] \) subject to \( z_{k+1} = A_{T-k} z_k + C_{T-k} u_k \), \( z_0 = v \), \( P_k \) represents the optimal cost in the sense that \( v^* P_{TT} v = J^T(v) \). One can employ the fact that \( DD' \geq \chi \) to get that \( \sum_{t=0}^{T-1} \|u_k\|^2 \leq \chi^{-1} J^T(v_{i,t}) = \chi^{-1} v_{i,t} M P_{TT} M' v_{i,t} < \chi^{-1} \nu \|v_{i,t}\|^2 \), in such a manner that \( A'(t T, 0) \) approaches the “noncontrolled” \( A'(t T, 0) v_{i,t} \) when \( \nu \) is small enough; this can be employed to show that

\[
\sum_{t=0}^{T-1} z_k E_k' E_k z_k > \tau^{2t} v_{i,t}^* A(t T, 0) \sum_{t=0}^{T-1} A'(t T, 0) v_{i,t}
\]

for each \( t \geq 1 \), for a sufficiently small \( \nu \). These evaluations yield

\[
(38) \quad \|v_{i,t}\|^2 > \tau^{2t} \nu \|z_{k,t}\|^2 > \tau^{2t} v_{i,t}^* A(t T, 0) \sum_{t=0}^{T-1} A'(t T, 0) v_{i,t} \geq \tau^{2t} \mu \|M m_i\|^2 \geq \tau^{2t} \lambda^2 (\Sigma) \|M m_i\|^2,
\]

hence \( \nu > \tau^{2t} \lambda^2 (\Sigma) \|M m_i\|^2 \|v_{i,t}\|^2 > \tau^{2t} \lambda^2 (\Sigma) \tau^{-2t} = \lambda^2 (\Sigma) \), which is absurd for sufficiently small \( \nu \). This can be easily extended to any linear combination \( v_t = \sum_{i=1}^n \alpha_i v_{i,t} \)
\[ \sum_{t} \alpha_i v_{i,t}; \] in particular, we mention that \( m_i \) in (38) is replaced by \( \sum_{t} \alpha_i m_i \). So far we have shown that \( v_t^\prime M P_t^\prime T M' v_t \geq \nu \| v_t \|^2 \) for some \( \nu > 0 \) and all \( t \geq 1 \), where \( v_t \) is in the space spanned by \( v_{i,t}, 0 \leq i \leq \ell \). This is easily extended to \( t \geq 0 \) by setting \( \nu = \min(\nu, \lambda_-(M \Sigma M')) \). For \( \ell + 1 \leq i \leq n \), it is simple to check that \( v_t^\prime M P_t^\prime T M' v_{i,t} = 0, t \geq 0 \), which completes the proof of the third statement for \( E = 0 \). For \( E \neq 0 \) we proceed as follows. For each \( i = \ell + 1, \ldots, n \) such that \( v_t^\prime M P_t^\prime T M' v_{i,t} > 0 \) for some \( t_i < \infty \) we can replace \( \Sigma \) in the above by \( P_{i,t} \) and obtain evaluations similar to (38) for \( t \geq t_i \); see the details in [8].

**Appendix B. Proof of Lemma 1 (continued).** We prove inequality (18). From Proposition 2(v) we have that \( BB' \leq (1 + \varepsilon^2) M B B' M' + (1 + \varepsilon^{-2}) (I - M) B B'(I - M)' \), and employing Proposition 2(iii) yields

\[
(39) \quad \tilde{X}_h((k + t)T, tT, BB') \leq \tilde{X}_h((k + t)T, tT, (1 + \varepsilon^2) M B B' M' + (1 + \varepsilon^{-2}) (I - M) B B'(I - M')).
\]

For the first term on the right-hand side of (39), setting \( \mu > 0 \) as in Proposition 2(vi) with \( \xi < \phi < 1 \) and employing Proposition 2(iii), we obtain

\[
\tilde{X}_h((k + t)T, tT, \mu(1 - \phi) M B B' M') \leq \| B \|^2 \tilde{X}_h((k + t)T, tT, \mu(1 - \phi) M M')
\leq \| B \|^2 \tilde{X}_h((k + t)T, 0, \mu(1 - \phi) M M', 0) \leq \phi^{-(k+t)T} \| B \|^2 \tilde{X}_h((k + t)T, 0, 0, I)
\leq \phi^{-(k+t)T} \| B \|^2 \tilde{X}_h((k + t)T, 0, B = E, I).
\]

As a result, from (17) with \( \kappa = 1, \Psi = I, \) and \( \xi \) replaced by \( \tilde{\xi} \) such that \( \xi < \tilde{\xi} \phi^{1/2} < 1 \), we obtain

\[
(40) \quad \tilde{X}_h((k + t)T, tT, M B B' M') \leq \phi^{-(k+t)T} \mu^{-1}(1 - \phi)^{-1} \| B \|^2 \tilde{X}_h((k + t)T, 0, B = E, I)
\leq (\tilde{\xi} \phi^{1/2})^{-2(k+t)T} \mu^{-1}(1 - \phi)^{-1} \| B \|^2 \tilde{X}.
\]

For the second term on the right-hand side of (39) we have, from Proposition 2(vi), that

\[
(41) \quad X_h((k + t)T, tT, (I - M) B B'(I - M')) \leq \tilde{\xi}^{-2kT} \| B \|^2 \tilde{X}_h.
\]

Substituting (40) and (41) in (39), we obtain

\[
(42) \quad \tilde{X}_h((k + t)T, tT, BB') \leq \tilde{\xi}^{-2(k+t)T} \| B \|^2 \tilde{X},
\]

where \( \tilde{\xi} = \phi^{1/2} \tilde{\xi} \) and \( \tilde{X} = (1 + \varepsilon^2) \mu^{-1}(1 - \phi)^{-1} \tilde{X} + (1 + \varepsilon^{-2}) \tilde{X}_h \). Note that (42) represents an evaluation for the maximal expansion of \( \tilde{X}_h \) along time intervals of the form \([tT, (k + t)T]\). The inequality (18) follows by extending the above evaluation to general intervals \([t, k + t]\), defining \( \tilde{k} \) as the largest integer for which \( \tilde{k}T \leq k + t \) and replacing \( k + t \) in (42) by \( \tilde{k} \), combined with the evaluation for the expansion of \( \tilde{X} \) provided in Proposition 2(ii).
Appendix C. Proof of Lemma 4. Consider the set \( \mathcal{U} = \{ U \in \mathcal{R}^n, \| U \| = 1, \ker \{ U \} \supset \ker \{ P(tT) \} \} \). Similarly to (39), for \( U \in \mathcal{U} \),

\[
X_h((k + t)T, tT, U) \leq (1 + \varepsilon^2)X_h((k + t)T, tT, MUM') + (1 + \varepsilon^{-2})\tilde{X}_h((k + t)T, tT, (I - M)U(I - M)'),
\]

and similarly to (41),

\[
X_h((k + t)T, tT, (I - M)U(I - M')) \leq \xi^{-2kT}\| U \| \tilde{X}_h \leq I \leq \xi^{-2kT}X_h.
\]

We need an evaluation for the first term on the right-hand side of (43). Two cases arise: for \( MP(tT)M' \neq 0 \), Proposition 2(vi) yields \( \lambda_\nu(MP(tT)M') \geq \nu \), and if we let \( N' \) be such that \( N' \) is the orthogonal projection onto the space spanned by eigenvectors associated with eigenvalues of \( MP(tT)M' \) different from zero, then \( MP(tT)M' \geq \nu NN' \).

Since \( \ker \{ U \} \supset \ker \{ P(tT) \} \), we have that \( \ker \{ MUM' \} \supset \ker \{ MP(tT)M' \} \). One can check that \( NMM'N' = MUM' \). These evaluations allow us to write

\[
\nu N\nu M' = \nu NN'M' \leq \nu NN' \leq MP(tT)M'.
\]

Then, employing Proposition 2(iii) and Assumption 1, we obtain

\[
X_h((k + t)T, tT, MUM') = \nu^{-1}X_h((k + t)T, tT, \nu MUM')
\]

\[
\leq \nu^{-1}\tilde{X}((k + t)T, tT, \nu MUM') \leq \nu^{-1}\tilde{X}((k + t)T, tT, MP(tT)M') \leq \nu^{-1}\tilde{X}.
\]

By substituting (44) and (45) in (43), and setting \( \tilde{X} = (1 + \varepsilon^2)\nu^{-1}\tilde{X} + (1 + \varepsilon^{-2})\tilde{X}_h \),

we get that \( X_h((k + t)T, tT, U) \leq \xi^{-2kT}\tilde{X} \), which allows us to evaluate, for \( V \in \mathcal{R}^n \),

such that \( V \neq 0 \) and \( \ker \{ V \} \supset \ker \{ P(tT) \} \) (possibly with \( \| V \| \neq 1 \)),

\[
X_h((k + t)T, tT, V) = \| V \| X_h((k + t)T, tT, \| V \|^{-1}V) \leq \xi^{-2kT}\| V \| \tilde{X}.
\]

We can extend (46) to general intervals \([t, k + t] \), analogously to the proof of Lemma 1. Indeed, we write for general \( k, t \geq 0 \),

\[
\tilde{X}_h(k + t, t, V) = \tilde{X}_h(k + t, kT + 1, \tilde{X}_h(kT, (k - m)T, \tilde{X}_h((k - m)T - 1, t, V))),
\]

where \( k \) and \( m \) are the largest integers such that \( k \leq k + t \) and \( (k - m)T \geq t \). However, in order to substitute (46) in the above equation, we have to check that \( \tilde{X}_h((k - m)T - 1, t, V)) \supset \ker \{ P((k - m)T - 1) \} \); by hypothesis \( \ker \{ P(t) \} \subset \ker \{ V \} \),

and thus Lemma 3 yields \( \ker \{ P((k - m)T - 1) \} \subset \ker \{ \tilde{X}((k - m)T - 1, t, V) \} \) and the result follows from the fact that \( \ker \{ \tilde{X}((k - m)T - 1, t, V) \} \subset \ker \{ \tilde{X}_h((k - m)T - 1, t, V) \} \) (since \( \tilde{X} \supset \tilde{X}_h \)). Substituting (46) in the above equation and employing the evaluation for the expansion of \( \tilde{X} \) provided in Proposition 2(ii), we obtain

\[
\tilde{X}_h(k + t, kT + 1, \tilde{X}_h(kT, (k - m)T, \tilde{X}_h((k - m)T - 1, t, V)))
\]

\[
\leq \tilde{X}_h(k + t, kT + 1, \tilde{X}_h(kT, (k - m)T, \delta \| V \|))
\]

\[
\leq \tilde{X}_h(k + t, kT + 1, \xi^{-2mT} \delta \| V \| \tilde{X}) \leq \xi^{-2mT} \delta^2 \| V \| \tilde{X} \leq \xi^{-2k} \tilde{X},
\]
where \( \bar{X} = \delta^{2T}\|\!\| V \|\!\| \bar{X} \). \( \square \)

**Appendix D. Proof of Lemma 6.**

*Proof of (i).* We proceed inductively. For \( k = 0 \) we employ Proposition 2(v) with \( V = I \) to write

\[
\tilde{X}_\xi(0, 0, I) = I \leq (1 + \epsilon^{-2})H_0 H_0' + (1 + \epsilon^2)(I - H_0)(I - H_0)'.
\]

Now, assume that statement (i) of the lemma holds for some \( k > 0 \). If we adapt Proposition 2(iii) to the situation in which \( A \) and \( C \) are replaced by \( \xi A \) and \( \xi C \), respectively, we get for \( V_0, V_1 \in \mathcal{R}^{n_0}, V_1 \geq V_0 \), that \( (\xi(A + L_k C))V_1(\xi(A + L_k C))' \geq (\xi(A + L_k C))V_1(\xi(A + L_k C))' \), and we can write

\[
\tag{47}
\tilde{X}_\xi(k + 1, 0, I) = (\xi(A + L_k C))X_\xi(k, 0, I)(\xi(A + L_k C))' + L_k DD'L_k + BB'
\leq (\xi(A + L_k C))\left( X_\xi(k, 0, 0) + \xi^{2k}(1 + \epsilon^{-2})^{k+1} Z(k) + \xi^{2k}(1 + \epsilon^2)\sum_{\ell=1}^{k}((1 + \epsilon^{-2})^{\ell+1} \times \tilde{X}_h(k - 1, \ell, I - H_\ell)\tilde{X}_h(\ell - 1, \ell - 1, Z(\ell - 1))(I - H_\ell)'(\xi(A + L_k C)') + L_k DD'L_k + BB'.
\]

Now we evaluate the terms on the right-hand side of (47); for the first term,

\[
\tag{48}
(\xi(A + L_k C))\tilde{X}_\xi(k, 0, 0)(\xi(A + L_k C))' + L_k DD'L_k + BB' = \tilde{X}_\xi(k + 1, 0, 0).
\]

For the second term, we employ Proposition 2(v) and (26) to write

\[
\tag{49}
(\xi(A + L_k C))(\xi^{2k}(1 + \epsilon^{-2})^{k+1} Z(k))(\xi(A + L_k C))' \leq (1 + \epsilon^{-2})H_{k+1}(\xi(A + L_k C))(\xi^{2k}(1 + \epsilon^{-2})^{k+1} Z(k))(\xi(A + L_k C))' H_{k+1} + (1 + \epsilon^2)(I - H_{k+1})(\xi(A + L_k C))(\xi^{2k}(1 + \epsilon^{-2})^{k+1} Z(k)) \times (\xi(A + L_k C))' (I - H_{k+1})' = \xi^{2(k+1)}(1 + \epsilon^{-2})^{k+1} Z(k + 1)
\]

\[
+ (1 + \epsilon^2)(I - H_{k+1})\xi^2\tilde{X}_h(k, k, (\xi^{2k}(1 + \epsilon^{-2})^{k+1} Z(k)))(I - H_{k+1})' = \xi^{2(k+1)}(1 + \epsilon^{-2})^{k+1} Z(k + 1) + (1 + \epsilon^2)(I - H_{k+1})\tilde{X}_h(k, k, Z(k))(I - H_{k+1})'.
\]

For the third term on the right-hand side of (47), we have

\[
\tag{50}
(\xi(A + L_k C))(\xi^{2k}(1 + \epsilon^2)\tilde{X}_h(k - 1, 0, (I - H_0)I(I - H_0)')(\xi(A + L_k C))' = \xi^{2(k+1)}(1 + \epsilon^2)\tilde{X}_h(k, 0, (I - H_0)I(I - H_0)').
\]
and for the fourth term,

\[(51) \quad (\xi(A + L_k C))(\xi^2(1 + \epsilon^2) \sum_{\ell=1}^{k} [(1 + \epsilon^{-2})^{\ell} \times \bar{X}_h(k - 1, \ell, (I - H_\ell)\bar{X}_h(\ell - 1, \ell - 1, Z(\ell - 1))(I - H_\ell)'(\xi(A + L_k C))']\]

\[= \xi^{2(k+1)}(1 + \epsilon^2) \sum_{\ell=1}^{k} [(1 + \epsilon^{-2})^{\ell} \times \bar{X}_h(k, \ell, (I - H_\ell)\bar{X}_h(\ell - 1, \ell - 1, Z(\ell - 1))(I - H_\ell)'(\xi(A + L_k C))']\]

The induction is completed by substituting (48)–(51) in (47).

Proof of (ii). The proof is similar to that of (i) and will not be presented. We mention only that for \(k = 0\) we employ Proposition 2 (v) with \(V = I\) and \(\epsilon = \sqrt{2}\) to write

\[\bar{X}_\xi(0, 0, I) = I \geq (1/2)H_0 H_0' - (I - H_0)(I - H_0)',\]

and then an induction similar to that in (i) starts with \(k = 1\), now employing Proposition 2(v) with \(V = I\). 

Appendix E. Proof of Lemma 7. For ease of notation, let \(Z(k, E, \Sigma, A), Z(k, C, \Sigma, A),\) and \(Z(k, 0, \Sigma, H A)\) be denoted, respectively, as \(Z(k), \bar{Z}(k), \) and \(\bar{Z}(k),\) and use similar notation for the variables \(X\) and projections \(H;\) e.g., \(H_k\) refers to the projection onto \(\text{ker}\{X(k, E, \Sigma, A)\}.\) We shall need the following auxiliary results. From (33) it follows that

\[(52) \quad X(k, U, V, W) = W^k V W^{k'} + \sum_{\ell=0}^{k-1} W^\ell U U' W^{\ell'}, \quad k \geq 0.\]

Consider system \(\Theta(C, \Sigma, A).\) Let \(\bar{X}_{nf}(k)\) and \(\bar{X}_f(k)\) stand for the free and forced solutions, given by

\[\bar{X}_{nf}(k) = A^k \Sigma A^k, \quad k \geq 0,\]

\[\bar{X}_f(k) = \sum_{\ell=0}^{k-1} A^\ell C C' A^\ell, \quad k \geq 1, \quad X_{f,0} = 0.\]

From (52) we have \(\bar{X}(k) = X(k, C, \Sigma, A) = \bar{X}_{nf}(k) + \bar{X}_f(k).\) Introduce the orthogonal projections \(H_{f, k} \in \mathbb{R}^n\) as

\[H_{f, k} = I - (X_f(k))^*(X_f(k)).\]

The next result follows from basic matrix properties [20].

**Lemma 10.** \(H_k(C, \Sigma, A) = (I - (H_{f, k}\bar{X}_{nf}(k) H_{f, k}')^*(H_{f, k}\bar{X}_{nf}(k) H_{f, k}')) H_{f, k}.\)

**Lemma 11.** The following statements hold.

(i) \(\text{ker}\{X_f(k)\} = \text{ker}\{CC'\}, \quad k \geq 0.\)

(ii) \(H_{f, k} = H, \quad k \geq 1, \) and \(H_{f,0} = I.\)
Proof. (i) From (53) it follows that \( \bar{X}_t(1) = \sum_{t=0}^\infty A^t CC'A'^t = CC' \), so it is enough to show that \( \ker\{\bar{X}_t(k)\} = \ker\{\bar{X}_t(1)\} \). For each \( v \in \mathbb{R}^n \) such that \( v'\bar{X}_t(1)v = 0 \), we have that

\[
0 = v'\bar{X}_t(1)v = v'CC'v = v'(EE' + AEE'A' + \cdots + A^{n-1}EE'A^{n-1})v,
\]

leading to \( v'A^tEE'A'^tv = 0, \ 0 \leq t \leq n-1 \). Employing the Cayley–Hamilton theorem, one can check that \( v'A^tEE'A'^tv = 0, \ t \geq 0 \), yielding

\[
v'\bar{X}_t(k)v = v' \left[ \sum_{t=0}^{k-1} A^t CC'A'^t \right] v = v' \left[ \sum_{t=0}^{k-1} A^t(EE' + AEE'A' + \cdots + A^{n-1}EE'A^{n-1})A'^t \right] v = 0,
\]

and we conclude that \( \ker\{\bar{X}_t(1)\} \subset \ker\{\bar{X}_t(k)\} \). For the converse relation, note that

\[
\bar{X}_t(k) = \sum_{t=0}^{k-1} A^t CC'A'^t = A \left( \sum_{t=0}^{k-2} A^t CC'A'^t \right) A' + CC',
\]

and hence for each \( v \in \mathbb{R}^n \), such that \( v'\bar{X}_t(k)v = 0 \), we can write \( v'\bar{X}_t(1)v = v'CC'v = 0 \), and thus \( \ker\{y\bar{X}_t(1)\} \supset \ker\{\bar{X}_t(k)\} \).

(ii) The result is trivial for \( k = 0 \), as \( H_{0,0} \) is the projection onto the null space of \( \bar{X}_t(0) = 0 \); for \( k \geq 1 \) it follows immediately from assertion (i).

Lemma 12. \( (HA)^{k}\sigma = HA^{k}\sigma, \ k \geq 1 \).

Proof. For \( k = 1 \) the result is trivial. From (35), \( H \) is the orthogonal projection onto \( \ker\{CC'\} = \ker\{C'\} \). Then, for each \( v \in \mathbb{R}^n \), \( (I - H)v \perp \ker\{C'\} \). Since the space orthogonal to \( \ker\{C\} \) is \( A \)-invariant \([7]\), we have that \( A(I - H)v \perp \ker\{C'\} \) and \( HA(I - H)v = 0 \). Therefore, setting \( v = A\sigma \) yields

\[
(HA)^{k}\sigma = (HA)^{k-1}[A\sigma] + (HA)^{k-1}[(H - I)A\sigma] = (HA)^{k-1}A\sigma, \ k \geq 2,
\]

and the result follows in a straightforward manner.

Lemma 13. The following statements hold:

(i) \( H_{0}(C, \Sigma, A) = I - \Sigma^{*}\Sigma; \)
(ii) \( H_{k}(C, \Sigma, A) = (I - (HA)^{k}\Sigma(HA)^{k})^{*}(HA)^{k}\Sigma(HA)^{k^{*}})H, \ k \geq 1. \)

Proof. (i) is immediate from the definition of \( H_{0}. \)

(ii) We employ Lemma 12 and assertion (ii) of Lemma 11, respectively, to write

\[
(HA)^{k}\Sigma(HA)^{k^{*}} = HA^{k}\Sigma A^{k^{*}}H = H_{k,k}A^{k}\Sigma A^{k^{*}}H_{k,k} = H_{k,k}X_{nl}(k)H_{k,k},
\]

and Lemma 10 completes the proof.

Proof of Lemma 7 (continued). Since \( CC' = EE' + \cdots + A^{n-1}EE'A^{n-1} \), one has that \( EE' \leq CC' \), which yields \( A^kEE'A^{k^{*}} \leq A^k CC'A^{k^{*}} \), \( k \geq 0 \), allowing us to evaluate, employing (52),

\[
X(k) = A^k\Sigma A^{k^{*}} + \sum_{t=0}^{k-1} A^t EE'A'^t \leq A^k\Sigma A^{k^{*}} + \sum_{t=0}^{k-1} A^t CC'A'^t = \bar{X}(k).
\]

Then, for each \( v \) such that \( \bar{X}(k)v = 0 \) we have that \( v'X(k)v \leq v'\bar{X}(k,v) = 0 \), allowing us to conclude that \( X(k)v = 0 \). This means that

\[
\ker\{X(k)\} \supset \ker\{\bar{X}(k)\}
\]
and it is simple to check that for each \( v \in \mathbb{R}^n \), \( H_k v = \bar{H}_k v + r \) for some \( r \perp \bar{H}_k v \). This allows us to evaluate, for each \( V \in \mathbb{R}^{n_0} \), \( H_k V H_k \geq \bar{H}_k V \bar{H}_k \) and to obtain \( Z(k) \geq \bar{Z}(k) \), completing the proof for the first statement. For the second statement, it is simple to check by inspection that \( X(k) = (H A)^k \Sigma (H A)^{kt} \) and

\[
\bar{H}_k = I - \bar{X}_k \bar{A}_k = I - ((H A)^k \Sigma (H A)^{kt})^* (H A)^k \Sigma (H A)^{kt}
\]

in such a manner that from Lemma 13 we have that \( \bar{H}_k = \bar{H}_k H \). This and (33) allow us to evaluate

\[
\bar{Z}(k + 1) - \bar{Z}(k + 1) = \bar{H}_k AZ(k) A' \bar{H}_k' - \bar{H}_k (H A) \bar{Z}(k) (H A)' \bar{H}_k', \tag{56}
\]

\[
= \bar{H}_k (H A) (Z(k) - \bar{Z}(k)) (H A)' \bar{H}_k', \quad k \geq 0.
\]

Since \( \bar{Z}(0) - \bar{Z}(0) = \bar{H}_0 \bar{H}_0' - \bar{Z}(0) = (I - \Sigma^* \Sigma)(I - \Sigma^* \Sigma)' - \bar{Z}(0) = \bar{H}_0 \bar{H}_0' - \bar{Z}(0) = 0 \), employing (56) in a recursive fashion for \( k = 0, 1, \ldots \) yields \( \bar{Z}(k) - \bar{Z}(k) = 0 \), \( k \geq 0 \). \( \square \)

**Appendix F. Proof of Lemma 8.** Since \( Z(k, 0, \Sigma, H A) \leq \xi^{2k} \bar{X} \), Proposition 3 with \( A \) replaced by \( H A \) implies that Condition 2 holds; hence, Proposition 4 with \( A \) and \( V \), respectively, replaced by \( A = J_H HA J_H^{-1} \) and \( V = J_H \Sigma J_H^{-1} \) yields

\[
\ker \{ \bar{A}^n (J_H \Sigma J_H^{-1}) A^n \} \cap J_H = \{ 0 \}. \tag{57}
\]

Note that \( \bar{A}^n (J_H \Sigma J_H^{-1}) A^n = J_H (H A)^n \Sigma (H A)^n J_H' \) and (57) leads to

\[
\ker \{ J_H (H A)^n \Sigma (H A)^n J_H' \} \cap J_H = \{ 0 \}.
\]

Employing the fact that \( (H A)^n = H A^n \) (see Lemma 12 in Appendix E) we obtain

\[
\ker \{ J_H \bar{\Sigma} J_H' \} \cap J_H = \{ 0 \},
\]

where \( \bar{\Sigma} = H A^n \Sigma A^n H' \). Proposition 3 implies that, for each \( 0 \leq \xi < 1 \) there exists \( \bar{X} \) such that \( Z(k, 0, \bar{\Sigma}, H A) \leq \xi^{-2k} \bar{X} \), and from Lemma 7 we get that

\[
Z(k, c, \bar{\Sigma}, A) \leq \xi^{-2k} \bar{X}. \tag{58}
\]

Since \( H \) is an orthogonal projection, we have that \( \bar{\Sigma} = H A^n \Sigma A^n H' \leq A^n \Sigma A^n \leq A^n \Sigma A^n + CC' \), and Proposition 2(iii) yields \( X(k, c, \bar{\Sigma}, A) \leq X(k, c, A^n \Sigma A^n + CC', A) \), in such a manner that \( \ker \{ X(k, c, \bar{\Sigma}, A) \} \supset \ker \{ X(k, c, A^n \Sigma A^n + CC', A) \} \). This and (58), respectively, provide

\[
Z(k, c, A^n \Sigma A^n + CC', A) \leq Z(k, c, \bar{\Sigma}, A) \leq \xi^{-2k} \bar{X}. \tag{59}
\]

The link with the system \( \Theta(A, E, \Sigma) \) is that

\[
X(n, E, \Sigma, A) = A^n \Sigma A^n + \sum_{t=0}^{n-1} A^t E E' A'^t
\]

\[
= A^n \Sigma A^n + CC' = X(0, c, A^n \Sigma A^n + CC', A), \tag{60}
\]

and, in order to obtain a similar relation for the variable \( Z \), we consider a version \( \hat{\Theta} \) of system \( \hat{\Theta}(c, A^n \Sigma A^n + CC', A) \) with the initial condition \( Z(0) = H_0 H_0' \) replaced by \( Z(0) = H_0 \Delta \Delta' H_0' \), where

\[
\Delta = A (H_{n-1}(E, \Sigma, A) A) \cdots (H_1(E, \Sigma, A) A) H_0(E, \Sigma, A).
\]
Let \((\hat{Z}(k), \hat{X}(k))\) and \(\hat{H}_k\), respectively, stand for the state of \(\Theta\) and the projection onto the null space of \(\hat{X}_k\). Note that

\[
\hat{X}(k) = X(k, \mathcal{C}, A^n\Sigma A^n + CC', A),
\]

with the interpretation that the variable \(X\) is not modified by the introduction of \(\Delta\) in the variable \(Z\). From (60) we obtain \(\hat{X}(0) = X(n, E, \Sigma, A)\) and \(\hat{H}_0 = H_0(n, E, \Sigma, A)\) and, by substituting \(\Delta\) as above in \(\hat{Z}(0) = \hat{H}_0\Delta\hat{H}_0'\), we obtain

\[
\hat{Z}(0) = Z(n, E, \Sigma, A).
\]

Omitting details, we have that \(\ker\{\hat{X}(k)\} = \ker\{X(k + n, E, \Sigma, A)\}, \ k \geq 0\), and (62) allows us to obtain

\[
\hat{Z}(k) = Z(k + n, E, \Sigma, A), \quad k \geq 0.
\]

Moreover, for \(\Delta\) defined as above we have that \(\Delta\hat{Z} = |A|^{-2n}I\), and hence \(\hat{Z}(0) = \hat{H}_0\Delta\hat{H}_0' \leq |A|^{-2n}\hat{H}_0\hat{H}_0' = |A|^{-2n}Z(0, \mathcal{C}, A^n\Sigma A^n + CC', A)\). This, together with (61), provides that

\[
\hat{Z}(k) \leq |A|^{-2n}Z(k, \mathcal{C}, A^n\Sigma A^n + CC', A).
\]

From (59), (63), and (64) we get that \(Z(k + n, E, \Sigma, A) \leq |A|^{-2k}\hat{Z}\), which yields \(Z(k, E, \Sigma, A) \leq |A|^{-2k}\|\hat{Z}\|\), \(\|\hat{Z}(0, E, \Sigma, A)\|, \ldots, \|\hat{Z}(n - 1, E, \Sigma, A)\|\) \(k \geq 0\).

\[\blacksquare\]

REFERENCES