MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

Generalized KYP Lemma: Unified Characterization of Frequency Domain Inequalities with Applications to System Design

Tetsuya IWASAKI and Shinji HARA

(Communicated by Kazuo Murota)

METR 2003–27 August 2003

DEPARTMENT OF MATHEMATICAL INFORMATICS
GRADUATE SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF TOKYO
BUNKYO-KU, TOKYO 113-8656, JAPAN

WWW page: http://www.i.u-tokyo.ac.jp/mi/mi-e.htm

Generalized KYP Lemma: Unified Characterization of Frequency Domain Inequalities with Applications to System Design

T. Iwasaki[*] and S. Hara[†]

August 8, 2003

Abstract

The cerebrated Kalman-Yakubovič-Popov (KYP) lemma establishes the equivalence between a frequency domain inequality (FDI) and a linear matrix inequality (LMI), and has played one of the most fundamental roles in systems and control theory. This paper generalizes the KYP lemma in two aspects — the frequency range and the class of systems — and unifies various existing versions by a single theorem. In particular, our result covers FDIs in finite frequency intervals for both continuous/discrete-time settings as opposed to the standard infinite frequency range. The class of systems for which FDIs are considered is no longer constrained to be proper, and nonproper transfer functions including polynomials can also be treated. We study implications of this generalization, and develop a proper interface between the basic result and various engineering applications. Specifically, it is shown that our result allows us to solve a certain class of system design problems with multiple specifications on the gain/phase properties in several frequency ranges. The method is illustrated by numerical design examples of digital filters and PID controllers.

1 Introduction

One of the most fundamental results in the field of dynamical systems analysis, feedback control, and signal processing, is the Kalman-Yakubovič-Popov (KYP) lemma [1, 2]. Various properties of dynamical systems can be characterized in terms of a set of inequality constraints in the frequency domain. The KYP lemma establishes equivalence between such frequency domain inequality (FDI) for a transfer function and a linear matrix inequality (LMI) for its state space realization. The basic roles of the KYP lemma are two fold: it provides (i) insights into analytical approaches to systems theory, and (ii) a framework for numerical approaches to systems analysis and synthesis.

The KYP lemma [2] states that, given matrices A, B, and a Hermitian matrix Θ, the FDI

$$\begin{bmatrix} (j\omega I - A)^{-1}B \\ I \end{bmatrix}^* \Theta \begin{bmatrix} (j\omega I - A)^{-1}B \\ I \end{bmatrix} < 0, \quad \forall\, \omega \in \mathbb{R} \cup \{\infty\} \tag{1}$$

holds if and only if the LMI

$$\begin{bmatrix} A & B \\ I & 0 \end{bmatrix}^* \begin{bmatrix} 0 & P \\ P & 0 \end{bmatrix} \begin{bmatrix} A & B \\ I & 0 \end{bmatrix} + \Theta < 0 \tag{2}$$

[*]Department of Mechanical and Aerospace Engineering, University of Virginia, 122 Engineer's Way, Charlottesville, VA 22904-4746, Email: iwasaki@virginia.edu

[†]Department of Information Physics and Computing, Graduate School of Information Science and Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-8656, Japan, E-mail: hara@crux.t.u-tokyo.ac.jp

admits a Hermitian solution P. Thus, the infinitely many inequalities (1) parametrized by ω can be checked by solving the finite dimensional convex feasibility problem (2). Appropriate choices of Θ in (1) allows us to represent various system properties including positive-realness and bounded-realness.

While the KYP lemma has been a major machinery for developing systems theory, it is not completely compatible with practical requirements. In particular, each design specification is often given not for the entire frequency range but rather for a certain frequency range of relevance. For instance, a closed-loop shaping control design typically requires small sensitivity in a low frequency range and small complementary sensitivity in a high frequency range. Thus a set of specifications would generally consists of different requirements in various frequency ranges. On the other hand, the standard KYP lemma treats FDIs for the entire frequency range only.

The current state of the art for fixing the incompatibility is to introduce the so-called weighting functions. A low/band/high-pass filter would be added to the system in series as a weight that emphasizes a particular frequency range and then the design parameters are chosen such that the weighted system norm is small. The weighting method has proven useful in practice, but suffers from nontrivial deficiencies. First, the additional weights tend to increase the system complexity (e.g., controller order), and the amount of increased complexity is positively correlated with the complexity of the weights. Second, the process of selecting appropriate weights is tedious and can be time-consuming, especially when the designer has to shoot for a good trade-off between the complexity of the weights and the accuracy in capturing desired specifications. Finally, the weighting method is effective only for gain specifications and it is not clear how to treat constraints on phase properties.

An alternative approach is to grid the frequency axis. In this case, infinitely many FDIs are approximated by a finite number of FDIs at selected frequency points. This approach has a practical significance especially when the system is well damped and the frequency response (after the design) is expected to be "smooth" (i.e., no sharp peaks). The resulting computational burden often seems moderate. However, it is difficult to determine *a priori* how fine the grid should be to achieve a certain performance. The method also suffers from the lack of a rigorous performance guarantee in the design process as the violation of the specifications may occur at a frequency between grid points.

Another approach that avoids both weighting functions and frequency gridding is to generalize the fundamental machinery, the KYP lemma, in such a way that FDIs in finite frequency ranges can be treated directly. There are several results in the literature along this line. The finite frequency KYP lemma obtained in [3] states that, given a scalar $\varpi > 0$, matrices A, B, and a Hermitian matrix Θ, the FDI in the low frequency range

$$\begin{bmatrix} (j\omega I - A)^{-1}B \\ I \end{bmatrix}^* \Theta \begin{bmatrix} (j\omega I - A)^{-1}B \\ I \end{bmatrix} < 0, \quad \forall\, |\omega| \le \varpi \tag{3}$$

holds if and only if the LMI

$$\begin{bmatrix} A & B \\ I & 0 \end{bmatrix}^* \begin{bmatrix} -Q & P \\ P & \varpi^2 Q \end{bmatrix} \begin{bmatrix} A & B \\ I & 0 \end{bmatrix} + \Theta < 0 \tag{4}$$

admits Hermitian solutions P and $Q > 0$. The nonstrict inequality version of this result has been obtained in [4] in the context of integrated design of structure/control systems. Other generalizations of the KYP lemma include those in [5, 6] where FDIs for polynomials are considered in the discrete-time setting. Specifically, algebraic conditions are given to characterize polynomials that are positive on an arc of the unit circle in the complex plane. These results are developed for the purpose of digital filter design.

The main objective of this paper is to further generalize the KYP lemma and to unify all the previous extensions mentioned above by a single theorem. There are two basic aspects in our generalization — the frequency range and the system class. In particular, we consider the frequency intervals characterized by two quadratic forms of the frequency variable λ. This characterization captures curve(s) on the complex plane and encompasses low/middle/high frequency conditions for both continuous-time and discrete-time systems. Thus, the result includes the previous results on the low/middle frequency conditions for continuous-time systems [4] as special cases, and adds new contributions to the cases involving high frequency conditions and discrete-time systems. The system class is generalized by, roughly speaking, replacing the term $(j\omega I - A)^{-1}B$ in (1) and (3) with a new term of the form $(\lambda E - A)^{-1}(B + \lambda F)$. It turns out that this allows us to capture FDIs for descriptor systems as well as polynomials.

Another contribution of the paper is to develop a useful interface between the generalized KYP lemma and various engineering applications. We will show that the design specifications encompassed by our result can be summarized, in the case of single-input single-output (SISO) transfer function designs, as follows: The Nyquist plot of the transfer function within a prescribed frequency interval must lie in the intersection of prescribed conic sections on the complex plane. It will also be shown that several classes of important engineering problems can be formulated in terms of such specifications, and can be solved exactly through LMI optimizations. Finally, design examples of digital filters and proportional-integral-derivative (PID) controllers will illustrate the procedures and advantages of our approach in comparison with existing ones.

This paper is organized as follows. We will first discuss in Section 2 system design problems that are naturally described by FDI specifications in various frequency ranges. These problems motivate our generalization of the KYP lemma to treat finite frequency ranges. Section 3 gives a characterization of general frequency ranges (curves on the complex plane) in terms of quadratic forms of the frequency variable. Section 4 presents our main results that generalize the standard KYP lemma. We will then show in Section 5 how various FDI specifications for systems design can be described within our framework. Finally, Section 6 shows what classes of design problems can be solved via convex optimizations, where analytical discussions will be followed by design examples.

We use the following notation. For a matrix M, its transpose, complex conjugate transpose, and the Moore-Penrose inverse are denoted by M^{T}, M^*, and $M^\dagger$, respectively. For a Hermitian matrix, $M > (\geq) 0$ and $M < (\leq) 0$ denote positive (semi)definiteness and negative (semi)definiteness. The real and the imaginary parts of M are denoted by $\Re(M)$ and $\Im(M)$. The symbol $\mathbf{H}_n$ stands for the set of $n \times n$ Hermitian matrices. The subscript n will be omitted if $n = 2$. For matrices Φ and P, $\Phi \otimes P$ means their Kronecker product. For $G \in \mathbb{C}^{n\times m}$ and $\Pi \in \mathbf{H}_{n+m}$, a function $\sigma : \mathbb{C}^{n\times m} \times \mathbf{H}_{n+m} \to \mathbf{H}_m$ is defined by

$$\sigma(G, \Pi) := \begin{bmatrix} G \\ I_m \end{bmatrix}^* \Pi \begin{bmatrix} G \\ I_m \end{bmatrix}.$$

Given a positive integer N, let $\mathbf{Z}_N$ be the set of nonnegative integers up to N, i.e., $\mathbf{Z}_N := \{ 0, 1, \ldots, N \}$.

2 Motivation: systems design

We will first motivate our research through several examples of design problems for which our generalized KYP lemma is naturally suited and can be useful. Later in the paper, some of these problems will be discussed in more detail with solution procedures and numerical design examples.

Structure/control design integration: Control performance of mechanical structures can be significantly enhanced if the designs of the structure and the controller are integrated [4, 7–9]. It has been shown that the finite frequency positive-real (FFPR) property in the low frequency range

$$G(j\omega) + G(j\omega) \geq 0, \quad \forall\, |\omega| \leq \varpi \tag{5}$$

is an important requirement for mechanical structure design to guarantee the existence of controllers that achieve high servo-bandwidth [4, 10], where $G(s)$ is the transfer function of the mechanical system to be designed. These references demonstrated that structures of practical significance can be designed to achieve the FFPR property by solving LMI feasibility problem with the aid of an existing version of the finite frequency KYP lemma [3, 4]. In particular, the method has been applied to a shape design of a swing arm for hard disk drives [4] and a smart arm design using piezo-electric film [10].

Digital filter design: The problem is to find a stable SISO transfer function $F(z)$ satisfying a set of frequency domain specifications. Typical design requirements for band-pass filters are given by the following Chebyshev approximation setting (see e.g. [11–14]):

$$\begin{array}{lll} |F(e^{j\theta})| \le \epsilon_\ell & : & |\theta| \le \vartheta_\ell \ , \\ |F(e^{j\theta})| \le \epsilon_h & : & \vartheta_h \le |\theta| \le \pi \ , \\ |F(e^{j\theta}) - M(e^{j\theta})| \le \epsilon_p & : & \vartheta_1 \le |\theta| \le \vartheta_2 \ , \\ |F(e^{j\theta})| \le 1 + \epsilon_o & : & \text{all } \theta \end{array} \tag{6}$$

where $\vartheta_\ell < \vartheta_1 < \vartheta_2 < \vartheta_h$ and $M(z)$ is a given function that has desired gain/phase properties in the pass-band $[\vartheta_1, \vartheta_2]$. An ideal response $M(z)$ would have the unity magnitude and the linear phase in the pass-band, i.e., $M(z) = z^{-d}$. The first three conditions specify the stop-bands and the pass-band, while the last condition is added to suppress overshoot in the transition bands $[\vartheta_\ell, \vartheta_1]$ and $[\vartheta_2, \vartheta_h]$.

Sensitivity-shaping: This is a typical control design with specifications on the closed-loop transfer functions [15,16]. For a plant $P(s)$ and a controller $K(s)$, the sensitivity and the complementary sensitivity functions $S(s)$ and $T(s)$ are defined by

$$S(s) := \frac{1}{1 - P(s)K(s)}, \quad T(s) := \frac{P(s)K(s)}{1 - P(s)K(s)}.$$

The sensitivity-shaping problem typically consists of the following type of requirements:

$$\begin{array}{lll} |S(j\omega)| \le \epsilon_\ell & : & |\omega| \le \varpi_\ell \ , \\ |T(j\omega)| \le \epsilon_h & : & |\omega| \ge \varpi_h \ , \\ |S(j\omega)| \le 1 + \epsilon_s & : & \text{all } \omega \ , \\ |T(j\omega)| \le 1 + \epsilon_t & : & \text{all } \omega \end{array}$$

where $\varpi_\ell < \varpi_h$. The first constraint may represent good tracking of a reference signal with spectral contents in the low frequency range, while the second may be interpreted as a robustness requirement against unmodeled high frequency dynamics of the plant. The last two constraints tend to avoid oscillatory time responses.

Open-loop shaping: This is a classical control design problem of determining the controller parameters to meet specifications on the open-loop transfer function. Given a marginally stable SISO plant $P(s)$, a set of specifications on the controller $K(s)$ is given in terms of the Nyquist plot of the open-loop transfer function $L(s) := K(s)P(s)$. Fig. 1 shows typical design requirements, where the shaded regions indicate where the Nyquist plot should lie in various frequency ranges. The half plane constraint on $L(j\omega)$ at the lower right part of the figure corresponds to the high-gain requirement in the low frequency range for the sensitivity reduction. The servo bandwidth requirement in the middle frequency range might be represented by an elliptic constraint as shown in the figure, where we may maximize the bandwidth. The two straight lines lying to the right of the point $-1 + j0$ provide gain/phase stability margins, and the small circle centered at the origin relates to the roll-off requirement in the high frequency range for robust stability.

3 Frequency range characterization

3.1 General framework

One of the key developments in this paper is a unified characterization of finite frequency ranges for both continuous-time and discrete-time systems. In general, a frequency range is visualized as a curve (or curves) on the complex plane. For instance, a low frequency range in the continuous-time setting is a line segment on

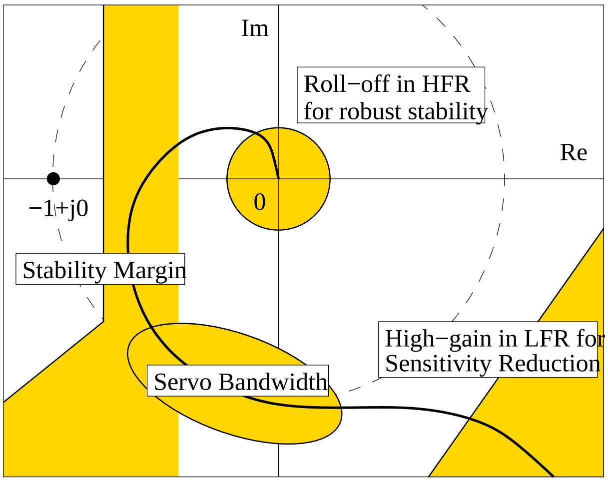


Figure 1: Open-loop shaping specifications

the imaginary axis containing the origin. A frequency range in the discrete-time setting would be an arc of the unit circle. The objective of this section is to propose a general framework for representing various curves (frequency ranges), which will later be used for characterizing FDIs. Proofs of the lemmas in this section are given in Appendix C.

Let us first clarify what we mean by curves on the complex plane.

Definition 1 *A curve on the complex plane is a collection of infinitely many points* $\lambda(t) \in \mathbb{C}$ *continuously parametrized by* t *for* $t_o \le t \le t_f$ *where* $t_o, t_f \in \mathbb{R} \cup \{\pm\infty\}$ *and* $t_o < t_f$*. A set of complex numbers* $\mathbf{\Lambda} \subseteq \mathbb{C}$ *is said to represent curves if it is a union of finitely many curves.*

We consider the following set of complex numbers that represent a certain class of curves:

$$\mathbf{\Lambda} := \{\, \lambda \in \mathbb{C} \mid \sigma(\lambda, \Phi) = 0,\ \sigma(\lambda, \Psi) \ge 0 \,\} \tag{7}$$

where $\Phi, \Psi \in \mathbf{H}$ are given matrices. We will explain what type of curves can be represented by $\mathbf{\Lambda}$ through what choices of Φ and Ψ. In order to exclude the cases which are out of our interests, we shall define the notion of trivial set as follows.

Definition 2 *A set of complex numbers* $\mathbf{\Lambda} \subseteq \mathbb{C}$ *is said to be trivial if one of the following holds: (a) it is empty; (b) it consists of a finite number of elements; (c) its closure is equal to* $\mathbb{C}$*. It is said to be nontrivial if it is not trivial.*

Note that the set $\mathbf{\Lambda}$ is the intersection of the following two:

$$\begin{aligned} \mathbf{\Lambda}_\Phi &:= \{\, \lambda \in \mathbb{C} \mid \sigma(\lambda, \Phi) = 0 \,\}, \\ \mathbf{\Lambda}_\Psi &:= \{\, \lambda \in \mathbb{C} \mid \sigma(\lambda, \Psi) \ge 0 \,\}. \end{aligned} \tag{8}$$

The result below shows when each of these sets is nontrivial, and how they look on the complex plane when nontrivial.

Lemma 1 *Let* $\Phi, \Psi \in \mathbf{H}$ *be given. The following equivalence relations hold:*

$$\begin{array}{lcl}
\mathbf{\Lambda}_\Phi \textit{ is nontrivial} & \Leftrightarrow & \det(\Phi) < 0 \\
\mathbf{\Lambda}_\Phi = \emptyset \textit{ or } \{\ \lambda_o\ \} & \Leftrightarrow & (\Phi \geq 0 \textit{ or } \Phi \leq 0) \textit{ and } \Phi \neq 0 \\
\mathbf{\Lambda}_\Phi = \mathbb{C} & \Leftrightarrow & \Phi = 0 \\
\mathbf{\Lambda}_\Psi \textit{ is nontrivial} & \Leftrightarrow & \det(\Psi) < 0 \\
\mathbf{\Lambda}_\Psi = \emptyset \textit{ or } \{\ \lambda_o\ \} & \Leftrightarrow & \Psi \leq 0 \textit{ and } \Psi \neq 0 \\
\mathbf{\Lambda}_\Psi = \mathbb{C} & \Leftrightarrow & \Psi \geq 0
\end{array}$$

When $\mathbf{\Lambda}_\Phi$ and $\mathbf{\Lambda}_\Psi$ are nontrivial, $\mathbf{\Lambda}_\Psi$ is given by one of the following regions, and $\mathbf{\Lambda}_\Phi$ is given by their boundaries:

Case $p > 0$: On or outside of $\mathbf{C}(-q/p, \gamma)$
Case $p < 0$: On or inside of $\mathbf{C}(-q/p, \gamma)$
Case $p = 0$: The half plane specified by $2\Re(q^*\lambda) + r \leq 0$

where $\mathbf{C}(-q/p, \gamma)$ is the circle with the center $-q/p$ and the radius γ, and, with $\Delta := \Phi$ or Ψ,

$$\begin{bmatrix} p & q \\ q^* & r \end{bmatrix} := \Delta, \quad \gamma := \sqrt{-\det(\Delta)}/|p|.$$

The above lemma shows that under the assumption of nontriviality of $\mathbf{\Lambda}_\Phi$ and $\mathbf{\Lambda}_\Psi$, i.e., $\det(\Psi) < 0$ and $\det(\Psi) < 0$, $\mathbf{\Lambda}_\Phi$ represents a circle or a line in the complex plane, while $\mathbf{\Lambda}_\Psi$ represents a half plane or a region inside or outside of a circle in the complex plane.

The following result gives an intuitive condition for $\mathbf{\Lambda}$, the intersection of $\mathbf{\Lambda}_\Phi$ and $\mathbf{\Lambda}_\Psi$, to be trivial.

Lemma 2 *Let $\Phi, \Psi \in \mathbf{H}$ be given and consider the set $\mathbf{\Lambda}$ defined in (7). Suppose $\det(\Phi) < 0$. Then the following statements are equivalent.*

(i) *$\mathbf{\Lambda}$ is trivial.*

(ii) *$\Psi \neq k\Phi$ for all $k \in \mathbb{R}$ and there exists $\tau \in \mathbb{R}$ such that $\Psi \leq \tau\Phi$.*

Clearly, if $\Psi = k\Phi$ for some k, then $\mathbf{\Lambda} = \mathbf{\Lambda}_\Phi$ because $\mathbf{\Lambda}_\Phi \subseteq \mathbf{\Lambda}_\Psi$, and hence $\mathbf{\Lambda}$ is nontrivial. When $\Psi \neq k\Phi$ for any $k \in \mathbb{R}$, it is easy to see that $\Psi \leq \tau\Phi$ implies triviality of $\mathbf{\Lambda}$ as follows. For $\lambda \in \mathbf{\Lambda}$, we have

$$0 \leq \sigma(\lambda, \Psi) \leq \tau\sigma(\lambda, \Phi) = 0$$

and thus $\mathbf{\Lambda}_\Psi$ is a circle, a line, a single point, or empty. Since $\mathbf{\Lambda}_\Phi$ is a circle or a line which is different from $\mathbf{\Lambda}_\Psi$ due to $\Psi \neq k\Phi$, its intersection with $\mathbf{\Lambda}_\Psi$ is either empty or a few points, showing triviality of $\mathbf{\Lambda}$. The converse statement can be proved through a version of the S-procedure (Lemma 9 in Appendix A). The detail is given in Appendix C.

The following lemma gives a closed-form characterization of the triviality condition in Lemma 2. To state the result, let us define a matrix Ψ_ϕ for a given pair of $\Phi, \Psi \in \mathbf{H}$ as follows:

$$\Psi_\phi := \Re[(JK)^{-*}\Psi(JK)^{-1}] \tag{9}$$

where K is an arbitrary matrix such that

$$\Phi = K^* \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} K, \quad J := \begin{bmatrix} 1 & 0 \\ 0 & j \end{bmatrix}. \tag{10}$$

The factorization of Φ as in (10) is possible if and only if $\det(\Phi) < 0$. The matrix Ψ_ϕ is not unique due to the freedom in the choice of K, but the result will be valid for any fixed choice.

Lemma 3 *Let $\Phi, \Psi \in \mathbf{H}$ be given. Suppose $\det(\Phi) < 0$ and define Ψ_ϕ by (9). Then the following statements hold true.*

(a) $\Psi_\phi \leq 0 \quad \Leftrightarrow \quad$ *There exists $\tau \in \mathbb{R}$ such that $\Psi \leq \tau\Phi$.*

(b) $\Psi_\phi \geq 0 \quad \Leftrightarrow \quad$ *There exists $\tau \in \mathbb{R}$ such that $\Psi \geq \tau\Phi$.*

A necessary and sufficient condition for $\mathbf{\Lambda}$ to represent curves can now be given as follows.

Proposition 1 *Let $\Phi, \Psi \in \mathbf{H}$ be given and define the matrix Ψ_ϕ by (9) and the sets $\mathbf{\Lambda}$, $\mathbf{\Lambda}_\Phi$, and $\mathbf{\Lambda}_\Psi$ by (7) and (8). Then the set $\mathbf{\Lambda}$ represents curves on the complex plane if and only if the following two conditions hold:*

- $\det(\Phi) < 0$,
- *Either $\Psi_\phi \geq 0$ or $\det(\Psi_\phi) < 0$ holds.*

In which case, there are two possibilities:

I. $\Psi_\phi \geq 0$: $\mathbf{\Lambda}$ is a circle or a line specified by $\mathbf{\Lambda}_\Phi$ since $\mathbf{\Lambda}_\Phi \subseteq \mathbf{\Lambda}_\Psi$.

II. $\det(\Psi_\phi) < 0$: $\mathbf{\Lambda}$ is a partial segment (or segments) of a circle or a line specified by $\mathbf{\Lambda}_\Phi \cap \mathbf{\Lambda}_\Psi$.

Proof. It is clear from Lemmas 2 and 3 that $\mathbf{\Lambda}$ is nontrivial if and only if $\Psi = \tau\Phi$ for some $\tau \in \mathbb{R}$ or $\lambda_{\max}(\Psi_\phi) > 0$. Note that the former condition is equivalent to $\Psi_\phi = 0$. This can be seen as follows. If $\Psi = \tau\Phi$, then Lemma 3 implies $0 \leq \Psi_\phi \leq 0$ from which we conclude $\Psi_\phi = 0$. Conversely, if $\Psi_\phi = 0$, then Lemma 3 implies existence of τ_1 and τ_2 such that $\tau_1\Phi \leq \Psi \leq \tau_2\Phi$. We then have $(\tau_2-\tau_1)\Phi \geq 0$ but, since $\det(\Phi) < 0$, this happens only if $\tau_1 = \tau_2$. Therefore $\tau_1\Phi = \Psi = \tau_2\Phi$ must be true. Thus we have shown that $\mathbf{\Lambda}$ is nontrivial if and only if $\Psi_\phi = 0$ or $\lambda_{\max}(\Psi_\phi) > 0$, which can readily be shown to be equivalent to $\Psi_\phi \geq 0$ or $\det(\Psi_\phi) < 0$. The former condition implies $\mathbf{\Lambda}_\Phi \subseteq \mathbf{\Lambda}_\Psi$ through Lemma 3 (b), and leads to Case I. The latter condition gives Case II as follows. Note that $\det(\Psi_\phi) < 0$ implies $\det(\Psi) < 0$ and hence $\mathbf{\Lambda}_\Psi$ is nontrivial. We need only to show that $\mathbf{\Lambda}_\Phi \not\subseteq \mathbf{\Lambda}_\Psi$. Since $\det(\Psi_\phi) < 0$, there exists a vector $\eta \in \mathbb{R}^2$ such that $\eta^\mathsf{T}\Psi_\phi\eta < 0$. Let $[\,\lambda_1 \;\; \lambda_2\,]^\mathsf{T} := (JK)^{-1}\eta$. Consider the case where $\lambda_2 = 0$. Let $\delta \in \mathbb{R}^2$ be such that $[01](JK)^{-1}\delta \neq 0$, where we note that existence of such δ is guaranteed by nonsingularity of JK. Defining $\eta_e := \eta + e\delta$ for a sufficiently small nonzero real scalar e, we have $\eta_e^\mathsf{T}\Psi_\phi\eta_e < 0$ and the second entry of the vector $(JK)^{-1}\eta_e$ is nonzero. Thus, we may assume $\lambda_2 \neq 0$ without loss of generality. Then it is straightforward to verify that $\lambda := \lambda_1/\lambda_2$ is such that $\lambda \in \mathbf{\Lambda}_\Phi$ but $\lambda \notin \mathbf{\Lambda}_\Psi$. ■

The following proposition provides another necessary and sufficient condition for $\mathbf{\Lambda}$ to represent curves through a common congruence transformation of Φ and Ψ. The proposition will play a key role for the derivation of our main result in the next section.

Proposition 2 *Let $\Phi, \Psi \in \mathbf{H}$ be given. Suppose $\det(\Phi) < 0$. Then there exists a common congruence transformation such that*

$$\Phi = M^*\Phi_o M, \quad \Psi = M^*\Psi_o M, \quad \Phi_o := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \Psi_o = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix} \tag{11}$$

where $\alpha, \beta, \gamma \in \mathbb{R}$ and $M \in \mathbb{C}^{2\times 2}$. In particular, α and γ are the eigenvalues of Ψ_ϕ in (9), which can be ordered to satisfy $\alpha \leq \gamma$. In this case, the set $\mathbf{\Lambda}$ in (7) represents curves on the complex plane if and only if either $\alpha < 0 < \gamma$ or $0 \leq \alpha \leq \gamma$ holds.

Proof. We need a series of lemmas on simultaneous matrix factorization appeared in Appendix B. Since $\det(\Phi) < 0$, there exists a nonsingular matrix K such that (10) holds. Let $\Upsilon := K^{-*}\Psi K^{-1}$, then Lemma 15 in Appendix guarantees it can be decomposed as in (59) with $N \in \mathbf{N}_o$. Note that α and γ are the eigenvalues of $\Psi_\phi = \Re[J\Upsilon J^*]$. Then the above factorizations of Φ and Ψ are obtained from (10) and (59) by noting the relation (58) and defining $M := NK$. Finally, the last statement follows from Proposition 1, where the former and the latter conditions correspond to $\det(\Psi_\phi) < 0$ and $\Psi_\phi \geq 0$, respectively. ■

3.2 Low/middle/high frequency ranges

By an appropriate choice of Φ and Ψ in (7), the set $\mathbf{\Lambda}$ can be specialized to define a certain range of the frequency variable λ. Let us start with the continuous-time case. The frequency variable $s = j\omega$ can be characterized by

$$\sigma(s, \Phi) = 0, \quad \Phi := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Consider the low frequency range $\lambda \in \mathbf{\Lambda}_{c\ell}$ defined by

$$\mathbf{\Lambda}_{c\ell} := \{ \, j\omega \mid \omega \in \mathbb{R}, \ |\omega| \leq \varpi \, \}$$

with $\varpi \in \mathbb{R}$ being a given positive scalar. The set $\mathbf{\Lambda}_{c\ell}$ is equal to $\mathbf{\Lambda}$ in (7) if we choose

$$\Psi = \begin{bmatrix} -1 & 0 \\ 0 & \varpi^2 \end{bmatrix}.$$

Similarly, the high frequency range can be characterized by

$$\mathbf{\Lambda}_{ch} := \{ \, j\omega \mid \omega \in \mathbb{R}, \ |\omega| \geq \varpi \, \}, \quad \Psi = \begin{bmatrix} 1 & 0 \\ 0 & -\varpi^2 \end{bmatrix}.$$

For the middle frequency range, we have

$$\mathbf{\Lambda}_{cm} := \{ \, j\omega \mid \omega \in \mathbb{R}, \ \varpi_1 \leq \omega \leq \varpi_2 \, \}, \quad \Psi := \begin{bmatrix} -1 & j\varpi_c \\ -j\varpi_c & -\varpi_1\varpi_2 \end{bmatrix}, \quad \varpi_c := \frac{\varpi_1 + \varpi_2}{2}.$$

This can be verified once we notice that

$$\sigma(j\omega, \Psi) = (\varpi_1 - \omega)(\omega - \varpi_2).$$

We now consider the discrete-time case. The frequency variable $z = e^{j\theta}$ can be characterized by

$$\sigma(z, \Phi) = 0, \quad \Phi := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Consider the low frequency range defined by

$$\mathbf{\Lambda}_{d\ell} := \{ \, e^{j\theta} \mid \theta \in \mathbb{R}, \ |\theta| \leq \vartheta \, \}.$$

Noting that $|\theta| \leq \vartheta$ if and only if $z = e^{j\theta}$ satisfies

$$z + z^* \geq 2\cos\vartheta =: \gamma,$$

we choose

$$\Psi = \begin{bmatrix} 0 & 1 \\ 1 & -\gamma \end{bmatrix}.$$

The high frequency range can be treated similarly:

$$\mathbf{\Lambda}_{dh} := \{\, e^{j\theta} \mid \theta \in \mathbb{R},\ \vartheta \le |\theta| \le \pi \,\}, \quad \Psi = \begin{bmatrix} 0 & -1 \\ -1 & \gamma \end{bmatrix}$$

where γ is defined as before. Finally, we consider the middle frequency range

$$\mathbf{\Lambda}_{dm} := \{\, e^{j\theta} \mid \theta \in \mathbb{R},\ \vartheta_1 \le \theta \le \vartheta_2 \,\}.$$

Note that the frequency interval condition can be written as

$$|\theta - \vartheta_c| \le \vartheta \quad \text{or} \quad \cos(\theta - \vartheta_c) \ge \cos\vartheta$$

where $0 \le \vartheta \le \pi$ and

$$\vartheta_c := \frac{\vartheta_2 + \vartheta_1}{2}, \quad \vartheta := \frac{\vartheta_2 - \vartheta_1}{2}.$$

This condition can be written as $\sigma(e^{j\theta}, \Psi) \ge 0$ with

$$\Psi := \begin{bmatrix} 0 & e^{j\vartheta_c} \\ e^{-j\vartheta_c} & -\gamma \end{bmatrix}, \quad \gamma := 2\cos\vartheta.$$

In summary, for the continuous-time setting, we have

$$\Phi = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{\Lambda} = \{\, j\omega \mid \omega \in \mathbf{\Omega} \,\}$$

where $\mathbf{\Omega}$ is a subset of real numbers specified by an additional choice of Ψ, for instance, as follows:

	LF	MF	HF
$\mathbf{\Omega}$	$\|\omega\| \le \varpi_\ell$	$\varpi_1 \le \omega \le \varpi_2$	$\|\omega\| \ge \varpi_h$
Ψ	$\begin{bmatrix} -1 & 0 \\ 0 & \varpi_\ell^2 \end{bmatrix}$	$\begin{bmatrix} -1 & j\varpi_c \\ -j\varpi_c & -\varpi_1\varpi_2 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & -\varpi_h^2 \end{bmatrix}$

where $\varpi_c := (\varpi_1 + \varpi_2)/2$, and LF, HF, and MF stand for low, high, and middle frequency ranges, respectively. Similarly, for the discrete-time setting, we have

$$\Phi = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{\Lambda} = \{\, e^{j\theta} \mid \theta \in \mathbf{\Theta} \,\}$$

where $\mathbf{\Theta}$ is a subset of real numbers specified by Ψ:

	LF	MF	HF
$\mathbf{\Theta}$	$\|\theta\| \le \vartheta_\ell$	$\vartheta_1 \le \theta \le \vartheta_2$	$\|\theta\| \ge \vartheta_h$
Ψ	$\begin{bmatrix} 0 & 1 \\ 1 & -2\cos\vartheta_\ell \end{bmatrix}$	$\begin{bmatrix} 0 & e^{j\vartheta_c} \\ e^{-j\vartheta_c} & -2\cos\vartheta \end{bmatrix}$	$\begin{bmatrix} 0 & -1 \\ -1 & 2\cos\vartheta_h \end{bmatrix}$

where $\vartheta_c := (\vartheta_1 + \vartheta_2)/2$, and $\vartheta := (\vartheta_2 - \vartheta_1)/2$.

4 Main results

4.1 S-procedure

We shall first present a version of the S-procedure [17]. The result is instrumental to the proof of the generalized KYP lemma to be stated later, and will be of independent interest as a tool for developments of systems theory. To this end, let us consider a subset $\mathbf{S}$ of Hermitian matrices and define certain properties that ensure exactness of the related S-procedure.

Definition 3 *For $\mathbf{S} \subset \mathbf{H}_\ell$, we define the following properties:*

(a) $\mathbf{S}$ *is a closed convex cone with vertex at the origin, i.e., it is a closed convex set such that*

$$S \in \mathbf{S} \ \Rightarrow \ \tau S \in \mathbf{S} \ ^\forall \tau \geq 0.$$

(b) For each matrix $H \in \mathbf{H}_\ell$ with rank $r > 0$ such that

$$H \geq 0, \quad \text{tr}(SH) \geq 0 \ ^\forall S \in \mathbf{S},$$

there exist vectors $\xi_i \in \mathbb{C}^\ell$ ($i = 1, \ldots, r$) such that

$$H = \sum_{i=1}^{r} \xi_i \xi_i^*, \quad \xi_i^* S \xi_i \geq 0 \ ^\forall S \in \mathbf{S}. \tag{12}$$

(c) $S \in \mathbf{S}$ and $S \leq 0$ imply $S = 0$.

Properties (a) and (b) have been used in [3] to define the notion of the lossless set, which leads to a nonconservative S-procedure with strict inequalities. The additional property (c) is necessary to keep exactness when nonstrict inequalities are used to describe the S-procedure. It turns out that Property (c) corresponds to the notion of controllability when we consider a special class of $\mathbf{S}$ tuned for systems analysis. The following is a generalized version of the S-procedure. The result with strict inequality appeared in [3], while that with nonstrict inequality is new.

Theorem 1 (S-procedure) *Let $\Theta \in \mathbf{H}_\ell$ and $\mathbf{S} \subset \mathbf{H}_\ell$ be given and define*

$$\mathbf{G} := \{\xi \in \mathbb{C}^q : \ \xi \neq 0, \ \xi^* S \xi \geq 0 \ ^\forall S \in \mathbf{S}\}. \tag{13}$$

Suppose $\mathbf{S}$ possesses Properties (a) and (b) in Definition 3. Then the following statements are equivalent.

(i) $\xi^* \Theta \xi < 0 \ ^\forall \xi \in \mathbf{G}$.

(ii) There exists $S \in \mathbf{S}$ such that $\Theta + S < 0$.

If in addition $\mathbf{S}$ possesses Property (c), then the following statements are equivalent.

(i) $\xi^* \Theta \xi \leq 0 \ ^\forall \xi \in \mathbf{G}$.

(ii) There exists $S \in \mathbf{S}$ such that $\Theta + S \leq 0$.

Proof. We shall prove the nonstrict inequality case only. See [3] for a proof of the strict inequality case. First note that (ii) $\Rightarrow$ (i) is obvious. To show the converse, let us suppose that (ii) does not hold, i.e., there is no $S \in \mathbf{S}$ such that $\Theta + S \leq 0$ holds. Then, from Lemma 10 in Appendix A, Properties (a) and (c) guarantee that there exists an $\epsilon > 0$ such that $\Theta + S \leq \epsilon I$ has no solution $S \in \mathbf{S}$. From the separating hyperplane argument (Lemma 11 in Appendix A), there exists a nonzero $H \in \mathbf{H}_\ell$ such that

$$H \geq 0, \quad \text{tr}[(\Theta + S - \epsilon I)H] \geq 0, \quad \forall\, S \in \mathbf{S}.$$

Since $\mathbf{S}$ contains the origin, we have

$$\text{tr}(\Theta H) \geq \epsilon \text{tr}(H) > 0, \quad \text{tr}(SH) \geq 0, \quad \forall\, S \in \mathbf{S}.$$

Due to Property (b), the latter condition implies the existence of $\xi_i \in \mathbb{C}^\ell$ such that (12) holds. Note that $\xi_i \in \mathbf{G}$ for all i. Moreover, $\text{tr}(\Theta H) = \sum_{i=1}^r \xi_i^* \Theta \xi_i > 0$ implies that $\xi_i^* \Theta \xi_i > 0$ must be true for some index i. This means that (i) with nonstrict inequality does not hold, and we conclude that (i) $\Rightarrow$ (ii). ∎

4.2 Generalized KYP lemma

This section derives a generalized KYP lemma using the S-procedure (Theorem 1); an FDI will be specified by statement (i) of Theorem 1, and statement (ii) then provides an LMI condition for satisfaction of the FDI. In particular, $\mathbf{G}$ in statement (i) captures the input/output graph of a system in a certain frequency range. To elaborate on this point, let us first discuss the case of the standard KYP lemma. For the FDI in (1), the set $\mathbf{G}$ would be given by

$$\mathbf{G} = \left\{ \begin{bmatrix} (j\omega I - A)^{-1} B \\ I \end{bmatrix} \eta \mid \eta \in \mathbb{C}^m,\ \eta \neq 0,\ \omega \in \mathbb{R} \cup \{\infty\} \right\}.$$

This set can also be characterized as

$$\mathbf{G} = \{\, \xi \in \mathbf{G}_\lambda \mid \lambda \in \bar{\mathbf{\Lambda}} \}, \quad \mathbf{G}_\lambda := \{\, \xi \in \mathbb{C}^{n+m} \mid \xi \neq 0,\ \Gamma_\lambda F \xi = 0 \,\}, \tag{14}$$

where $\bar{\mathbf{\Lambda}} := j\mathbb{R} \cup \{\infty\}$ and

$$\Gamma_\lambda := \begin{cases} \begin{bmatrix} I_n & -\lambda I_n \end{bmatrix} & \text{if} \quad \lambda \in \mathbb{C} \\ \begin{bmatrix} 0 & -I_n \end{bmatrix} & \text{if} \quad \lambda = \infty \end{cases}, \quad F := \begin{bmatrix} A & B \\ I_n & 0 \end{bmatrix}. \tag{15}$$

We now generalize this characterization of the FDI with respect to the frequency range. Recall that the set $\mathbf{\Lambda}$ in (7) captures finite frequency ranges of our interest. Hence we consider the general frequency range $\mathbf{\Lambda}$ in (7) for the definition of $\mathbf{G}$ in (14) where $\bar{\mathbf{\Lambda}}$ is defined as

$$\bar{\mathbf{\Lambda}} := \mathbf{\Lambda} \quad (\text{if } \mathbf{\Lambda} \text{ is bounded}), \quad \bar{\mathbf{\Lambda}} := \mathbf{\Lambda} \cup \{\infty\} \quad (\text{otherwise}). \tag{16}$$

Finally, the system description is also generalized by considering a completely arbitrary matrix F rather than the one with the structure as in (15).

It can now be seen that the main technical steps to arrive at the generalized KYP lemma are to express the set $\mathbf{G}$ in (14) as in (13) by an appropriate choice of $\mathbf{S}$ and to show that the chosen set $\mathbf{S}$ possesses the properties in Definition 3 so that the S-procedure is exact. The following two lemmas provide these steps. Proofs are given in Appendices D and E.

Lemma 4 *Let $F \in \mathbb{C}^{2n \times (n+m)}$ and $\Phi, \Psi \in \mathbf{H}$ be given such that $\mathbf{\Lambda}$ in (7) represents curves. Define $\bar{\mathbf{\Lambda}}$ and Γ_λ by (16) and (15), respectively. Then the set $\mathbf{G}$ defined in (14) can be characterized by (13) with*

$$\mathbf{S} := \{\, F^*(\Phi \otimes P + \Psi \otimes Q)F \mid P, Q \in \mathbf{H}_n,\ Q \geq 0 \,\}. \tag{17}$$

Lemma 5 *Let matrices* $F \in \mathbb{C}^{2n\times(n+m)}$ *and* $\Phi, \Psi \in \mathbf{H}$ *be given such that* $\mathbf{\Lambda}$ *in (7) represents curves. Define the set* $\mathbf{S}$ *by (17) and the matrix-valued mapping* Γ_λ *by (15). The following statements hold true.*

(i) The set $\mathbf{S}$ *has Properties (a) and (b) in Definition 3.*

(ii) The set $\mathbf{S}$ *has Property (c) and the matrix* F *is a minimal realization of the set* $\mathbf{S}$ *in the sense that*

$$F^*(\Phi \otimes P + \Psi \otimes Q)F = 0, \quad P, Q \in \mathbf{H}_n, \quad Q \geq 0 \quad \Rightarrow \quad P = Q = 0,$$

if and only if

$$\text{rank}(\Gamma_\lambda F) = n, \quad \forall\, \lambda \in \mathbb{C} \cup \{\infty\}. \tag{18}$$

We are now ready to state and prove a general theorem.

Theorem 2 (Generalized KYP) *Let matrices* $F \in \mathbb{C}^{2n\times(n+m)}$, $\Theta \in \mathbf{H}_{n+m}$, *and* $\Phi, \Psi \in \mathbf{H}$ *be given and define* $\mathbf{\Lambda}$ *and* $\bar{\mathbf{\Lambda}}$ *by (7) and (16), respectively. Suppose* $\mathbf{\Lambda}$ *represents curves on the complex plane. Denote by* N_λ *the null space of* $\Gamma_\lambda F$ *where* Γ_λ *is defined in (15). The following statements are equivalent.*

(i) $N_\lambda^* \Theta N_\lambda < 0 \quad \forall\, \lambda \in \bar{\mathbf{\Lambda}}$.

(ii) There exist $P, Q \in \mathbf{H}_n$ *such that* $Q > 0$ *and*

$$F^*(\Phi \otimes P + \Psi \otimes Q)F + \Theta < 0. \tag{19}$$

Moreover, if the rank condition in (18) is satisfied, then the following statements are equivalent.

(i) $N_\lambda^* \Theta N_\lambda \leq 0 \quad \forall\, \lambda \in \bar{\mathbf{\Lambda}}$.

(ii) There exist $P, Q \in \mathbf{H}_n$ *such that* $Q \geq 0$ *and*

$$F^*(\Phi \otimes P + \Psi \otimes Q)F + \Theta \leq 0. \tag{20}$$

Proof. We will prove the strict inequality case. The nonstrict inequality case can be shown similarly. Note that (i) holds if and only if the following equivalent statements hold; For each $\lambda \in \bar{\mathbf{\Lambda}}$,

$$\begin{aligned} & \xi^*\Theta\xi < 0, \ \ \forall\, \xi \text{ such that } \xi = N_\lambda \eta \text{ for some } \eta \neq 0 \\ \Leftrightarrow \quad & \xi^*\Theta\xi < 0, \ \ \forall\, \xi \neq 0 \text{ such that } \Gamma_\lambda F \xi = 0 \\ \Leftrightarrow \quad & \xi^*\Theta\xi < 0, \ \ \forall\, \xi \in \mathbf{G}_\lambda \end{aligned}$$

where $\mathbf{G}_\lambda$ is defined in (14). From Lemma 4, the last statement holds for all $\lambda \in \bar{\mathbf{\Lambda}}$ if and only if

$$\xi^*\Theta\xi < 0, \ \ \forall\, \xi \in \mathbf{G} \tag{21}$$

where $\mathbf{G}$ is defined in (13) with $\mathbf{S}$ given by (17). By Lemma 5, the set $\mathbf{S}$ has Properties (a) and (b). Hence, from the generalized S-procedure (Theorem 1), condition (21) is equivalent to the existence of $S \in \mathbf{S}$ satisfying $\Theta + S < 0$, or equivalently, the existence of $P, Q \in \mathbf{H}_n$ such that $Q \geq 0$ and (19) hold. Since the inequality in (19) is strict, we can strengthen the positivity of Q as $Q > 0$ without loss of generality. This completes the proof. ∎

In view of the developments in Section 3.2, the term $\Phi \otimes P + \Psi \otimes Q$ in statement (ii) of Theorem 2 can be specialized to the low/middle/high frequency ranges in the continuous/discrete-time settings as follows:

	LF	MF	HF
Cont.	$\begin{bmatrix} -Q & P \\ P & \varpi_\ell^2 Q \end{bmatrix}$	$\begin{bmatrix} -Q & P + j\varpi_c Q \\ P - j\varpi_c Q & -\varpi_1\varpi_2 Q \end{bmatrix}$	$\begin{bmatrix} Q & P \\ P & -\varpi_h^2 Q \end{bmatrix}$
Disc.	$\begin{bmatrix} -P & Q \\ Q & P - (2\cos\vartheta_\ell)Q \end{bmatrix}$	$\begin{bmatrix} -P & e^{j\vartheta_c}Q \\ e^{-j\vartheta_c}Q & P - (2\cos\vartheta)Q \end{bmatrix}$	$\begin{bmatrix} -P & -Q \\ -Q & P + (2\cos\vartheta_h)Q \end{bmatrix}$

It should be noted that Φ and Ψ are real if the region $\mathbf{\Lambda}$ on the complex plane is symmetric about the real axis. When F, Θ, Φ, and Ψ are all real matrices, P and Q in statement (ii) of Theorem 2 can be restricted to be real without loss of generality. This follows basically from the fact that the real part of a complex Hermitian positive definite matrix is positive definite. Note that, if the frequency region $\mathbf{\Lambda}$ is not symmetric about the real axis, then one needs to search for *complex* P and Q even when F and Θ are real. In such case, the LMI in complex variables can be converted to an LMI of larger dimension in real variables through the following equivalence:

$$X + jY = (X + jY)^* > 0 \quad \Leftrightarrow \quad \begin{bmatrix} X & Y \\ -Y & X \end{bmatrix} = \begin{bmatrix} X & Y \\ -Y & X \end{bmatrix}^{\mathsf{T}} > 0$$

where X and Y are real square matrices.

4.3 Specific extensions of the KYP lemma

In this section, we shall specialize the general result in Theorem 2 to obtain several versions of the KYP lemma. There are two aspects of specialization: the frequency range and the underlying system. The former has been briefly discussed at the end of the previous section. We will devote this section to the latter. In particular, we consider FDIs for descriptor and state space systems as well as for polynomial functions.

Theorem 3 (Descriptor, strict inequality version) *Let matrices* $A, E \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{n\times m}$, $\Theta \in \mathbf{H}_{n+m}$, *and* $\Phi, \Psi \in \mathbf{H}$ *be given and define* $\mathbf{\Lambda}$ *and* $\bar{\mathbf{\Lambda}}$ *by (7) and (16), respectively. Suppose that (a)* $\mathbf{\Lambda}$ *represents curves on the complex plane, (b)* $\det(\lambda E - A) \neq 0$ *for all* $\lambda \in \mathbf{\Lambda}$*, and (c) either* E *is nonsingular or* $\mathbf{\Lambda}$ *is bounded. Then, the following statements are equivalent.*

(i) $\begin{bmatrix} (\lambda E - A)^{-1}B \\ I \end{bmatrix}^* \Theta \begin{bmatrix} (\lambda E - A)^{-1}B \\ I \end{bmatrix} < 0 \quad \forall\, \lambda \in \bar{\mathbf{\Lambda}}.$

(ii) There exist $P, Q \in \mathbf{H}_n$ *such that* $Q > 0$ *and*

$$\begin{bmatrix} A & B \\ E & 0 \end{bmatrix}^* (\Phi \otimes P + \Psi \otimes Q) \begin{bmatrix} A & B \\ E & 0 \end{bmatrix} + \Theta < 0. \tag{22}$$

Proof. The result follows from Theorem 2 by choosing

$$F := \begin{bmatrix} A & B \\ E & 0 \end{bmatrix}, \quad \Gamma_\lambda F = \begin{cases} \begin{bmatrix} A - \lambda E & B \end{bmatrix} & \text{if} \quad \lambda \in \mathbf{\Lambda}, \\ \begin{bmatrix} -E & 0 \end{bmatrix} & \text{if} \quad \lambda = \infty. \end{cases} \tag{23}$$

If $\mathbf{\Lambda}$ is bounded, then the result is immediate by noting that the null space of $\Gamma_\lambda F$ is given by

$$N_\lambda = \begin{bmatrix} (\lambda E - A)^{-1}B \\ I \end{bmatrix} \tag{24}$$

for all $\lambda \in \mathbf{\Lambda}$ under the invertibility assumption on $\lambda E - A$. If $\mathbf{\Lambda}$ is unbounded, then we need to make sure that the above N_λ coincides with the null space of $\Gamma_\lambda F$ at $\lambda = \infty$ as well. This is indeed the case if E is nonsingular. ∎

In Theorem 3, statement (ii) implies that $\det(\lambda E - A) \neq 0$ for all $\lambda \in \mathbf{\Lambda}$, provided that the upper left $n \times n$ block of Θ, denoted by Θ_{11}, is positive semidefinite. To see this, let us assume (ii) and $\det(\lambda E - A) = 0$ for some $\lambda \in \mathbf{\Lambda}$. Let $x \neq 0$ be a vector in the null space of $\lambda E - A$, i.e., $\lambda E x = Ax$. Denote by W the matrix on the left hand side of (22). Then we have

$$\begin{bmatrix} x \\ 0 \end{bmatrix}^* W \begin{bmatrix} x \\ 0 \end{bmatrix} = (Ex)^*(\sigma(\lambda, \Phi)P + \sigma(\lambda, \Psi)Q)(Ex) + x^*\Theta_{11}x < 0$$

which cannot be true as every terms are nonnegative. By contradiction, we conclude that (ii) implies $\det(\lambda E - A) \neq 0$ for any $\lambda \in \mathbf{\Lambda}$. Based on this observation, Theorem 3 can be modified as follows: Replace supposition (b) with $\Theta_{11} \geq 0$, and add the condition $\det(\lambda E - A) \neq 0$ for all $\lambda \in \mathbf{\Lambda}$ to statement (i).

Theorem 4 (State space, nonstrict inequality version) *Let matrices $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{n\times m}$, $\Theta \in \mathbf{H}_{n+m}$, and $\Phi, \Psi \in \mathbf{H}$ be given and define $\mathbf{\Lambda}$ and $\bar{\mathbf{\Lambda}}$ by (7) and (16), respectively. Suppose that (a) $\mathbf{\Lambda}$ represents curves on the complex plane, and (b) the pair (A, B) is controllable. Let $\mathbf{\Omega}$ be the set of eigenvalues of A in $\mathbf{\Lambda}$. Then, the following statements are equivalent.*

(i) $\begin{bmatrix} (\lambda I - A)^{-1}B \\ I \end{bmatrix}^* \Theta \begin{bmatrix} (\lambda I - A)^{-1}B \\ I \end{bmatrix} \leq 0 \quad \forall\, \lambda \in \bar{\mathbf{\Lambda}} \backslash \mathbf{\Omega}.$

(ii) There exist $P, Q \in \mathbf{H}_n$ such that $Q \geq 0$ and

$$\begin{bmatrix} A & B \\ I & 0 \end{bmatrix}^* (\Phi \otimes P + \Psi \otimes Q) \begin{bmatrix} A & B \\ I & 0 \end{bmatrix} + \Theta \leq 0. \tag{25}$$

Proof. The result basically follows from the nonstrict inequality case of Theorem 2 by choosing F as in (23) with $E = I$. Note that the rank condition in (18) translates to controllability of (A, B) and nonsingularity of E. The only crucial step in this proof is to show that statement (i) is equivalent to

$$N_\lambda^* \Theta N_\lambda \leq 0, \quad \forall\, \lambda \in \mathbf{\Lambda} \tag{26}$$

where N_λ is the null space of $\Gamma_\lambda F$. As discussed in the proof of Theorem 3, N_λ can be given by (24) if $\lambda \in \mathbf{\Lambda} \backslash \mathbf{\Omega}$. Hence (i) is equivalent to

$$N_\lambda^* \Theta N_\lambda \leq 0, \quad \forall\, \lambda \in \mathbf{\Lambda} \backslash \mathbf{\Omega}.$$

By Lemma 12 in Appendix A, this condition is equivalent to (26) under the controllability assumption. ∎

Theorems 3 and 4 capture certain properties of dynamical systems expressed in terms of state space, possibly descriptor, realizations. The results include a finite frequency version [3] as well as the standard version [2] of the KYP lemma as special cases. A version of Theorem 3 with nonsingular E may also be obtained through an application of a recent, independent result by Scherer [18].

The following result characterizes certain properties of polynomials, and is particularly useful for the design of digital filters as discussed later.

Theorem 5 (Polynomial version) *Let matrices $G_k \in \mathbb{C}^{p\times m}$ ($k \in \mathbf{Z}_N$), $\Pi \in \mathbf{H}_{p+m}$, and $\Phi, \Psi \in \mathbf{H}$ be given and define $\mathbf{\Lambda}$ by (7). Suppose that $\mathbf{\Lambda}$ represents curves on the complex plane. Define*

$$G(\lambda) := \sum_{k=0}^{N} \lambda^k G_k, \quad \mathcal{G} := \begin{bmatrix} G_0 & \cdots & G_N \end{bmatrix}. \tag{27}$$

Then the following statements are equivalent.

(i) $\sigma(G(\lambda),\Pi) \leq 0$ *holds for all* $\lambda \in \mathbf{\Lambda}$.

(ii) There exist $P, Q \in \mathbf{H}_n$ *such that* $Q \geq 0$ *and*

$$F^*(\Phi \otimes P + \Psi \otimes Q)F + \begin{bmatrix} \mathcal{G} \\ J \end{bmatrix}^* \Pi \begin{bmatrix} \mathcal{G} \\ J \end{bmatrix} \leq 0 \tag{28}$$

$$J := \begin{bmatrix} I_m & 0 \end{bmatrix}, \quad U := \begin{bmatrix} 0 & I_n \end{bmatrix}, \quad V := \begin{bmatrix} I_n & 0 \end{bmatrix}, \quad F := \begin{bmatrix} U \\ V \end{bmatrix} \tag{29}$$

where $U, V \in \mathbb{R}^{n \times (n+m)}$, $J \in \mathbb{R}^{m \times (n+m)}$, *and* $n := mN$.

The equivalence also holds when the three inequalities are replaced by strict inequalities with the modification that condition $G_N^* \Pi_{11} G_N < 0$ *is added to statement (i) if* $\mathbf{\Lambda}$ *is unbounded, where* $\Pi_{11} \in \mathbf{H}_p$ *is the upper left block of* Π.

Proof. For a vector $u \in \mathbb{C}^m$, define $x_k := \lambda^k u$ for $k \in \mathbf{Z}_N$. Then $x_0 = u$ and $\Gamma_\lambda F x = 0$ hold where $F^\top := [\ U^\top \ \ V^\top\ ]$ and $x^* := [\ x_0^* \ \cdots \ x_N^*\ ]$. Noting that $y = G(\lambda)u = \mathcal{G}x$, we see that, for a fixed $\lambda \in \mathbf{\Lambda}$, condition $\sigma(\mathcal{G}(\lambda), \Pi) \leq 0$ holds if and only if

$$x^* \begin{bmatrix} \mathcal{G} \\ J \end{bmatrix}^* \Pi \begin{bmatrix} \mathcal{G} \\ J \end{bmatrix} x \leq 0, \quad \text{for all } x \text{ such that} \quad \Gamma_\lambda F x = 0.$$

If $\mathbf{\Lambda}$ is bounded, the result then directly follows from Theorem 2. When $\mathbf{\Lambda}$ is unbounded, Theorem 2 infers equivalence of (i) and (ii) if condition

$$x^* \begin{bmatrix} \mathcal{G} \\ J \end{bmatrix}^* \Pi \begin{bmatrix} \mathcal{G} \\ J \end{bmatrix} x \leq 0, \quad \text{for all } x \text{ such that} \quad \Gamma_\infty F x = 0.$$

is added to statement (i). This additional condition is equivalent to $G_N \Pi_{11} G_N \leq 0$. We claim that this is already implied by the original statement (i) when $\mathbf{\Lambda}$ is unbounded. To see this, note that

$$\lim_{|\lambda| \to \infty} \frac{\sigma(G(\lambda), \Pi)}{|\lambda|^{2N}} = G_N^* \Pi_{11} G_N.$$

This implies that $\sigma(G(\lambda), \Pi)$ has a positive eigenvalue for $\lambda \in \mathbf{\Lambda}$ with a sufficiently large magnitude, whenever $G_N^* \Pi_{11} G_N \not\leq 0$. This completes the proof for the nonstrict inequality case. Finally the strict inequality case can be proved in the same manner. Note, however, that condition $G_N^* \Pi G_N < 0$ is not implied by $\sigma(G(\lambda), \Pi) < 0$ for all $\lambda \in \mathbf{\Lambda}$ in general, and hence has to be added to statement (i). ∎

Corollary 1 (Positive quasi-polynomial on the unit circle) *Let* $\Phi, \Psi \in \mathbf{H}$ *and* $G_k \in \mathbb{C}^{m \times m}$ *(*$|k| \in \mathbf{Z}_N$*) be given such that* $G_{-k} = G_k^*$ *for all* $k \in \mathbf{Z}_N$*. Define* $n := mN$*. Suppose* $\mathbf{\Lambda}$ *given by (7) defines an arc on the unit circle. Define* J, F, *and* $\mathcal{G}$ *by (29) and (27) with* G_0 *replaced by* $G_0/2$*. The following statements are equivalent.*

(i) $\displaystyle\sum_{k=-N}^{N} \lambda^k G_k \geq 0$ *holds for all* $\lambda \in \mathbf{\Lambda}$.

(ii) There exist $P, Q \in \mathbf{H}_n$ such that $Q \geq 0$ and

$$F^*(\Phi \otimes P + \Psi \otimes Q)F \leq \mathcal{G}^* J + J^* \mathcal{G}. \tag{30}$$

(iii) There exist $X \in \mathbf{H}_{n+m}$ and $Q \in \mathbf{H}_n$ such that, for each $k \in \mathbf{Z}_N$,

$$G_k = \sum_{j-i=k} Y_{ij}, \quad Y := X + F^*(\Psi \otimes Q)F, \quad X \geq 0, \quad Q \geq 0$$

where the summation is over all pair (i,j) such that $j - i = k$ and $i, j \in \mathbf{Z}_N$, and $Y_{ij} \in \mathbb{C}^{m \times m}$ $(i, j \in \mathbf{Z}_N)$ are obtained by equal partition of matrix of Y.

The equivalence also holds when every inequalities are replaced by strict inequalities.

Proof. The equivalence (i) $\Leftrightarrow$ (ii) directly follows from Theorem 5 by choosing

$$\Pi := \begin{bmatrix} 0 & -I \\ -I & 0 \end{bmatrix}, \quad \Phi := \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

It remains to show (ii) $\Leftrightarrow$ (iii). We consider the case $N = 3$. It should be obvious how to extend the idea for the general case. Let

$$P := \begin{bmatrix} P_{00} & P_{01} & P_{02} \\ & P_{11} & P_{12} \\ & & P_{22} \end{bmatrix}, \quad Y := \begin{bmatrix} Y_{00} & Y_{01} & Y_{02} & Y_{03} \\ & Y_{11} & Y_{12} & Y_{13} \\ & & Y_{22} & Y_{23} \\ & & & Y_{33} \end{bmatrix}$$

where the entries in the blank space are implied by the symmetry. We may replace Φ by $-\Phi$ without changing the set $\mathbf{\Lambda}$. Then

$$J^*\mathcal{G} + \mathcal{G}^* J - F^*(\Phi \otimes P)F = \begin{bmatrix} G_0 - P_{00} & G_1 - P_{01} & G_2 - P_{02} & G_3 \\ & P_{00} - P_{11} & P_{01} - P_{12} & P_{02} \\ & & P_{11} - P_{22} & P_{12} \\ & & & P_{22} \end{bmatrix} = Y,$$

from which we obtain

$$\begin{bmatrix} G_3 \\ P_{02} \\ P_{12} \\ P_{22} \end{bmatrix} = \begin{bmatrix} Y_{03} \\ Y_{02} \\ Y_{12} \\ Y_{22} \end{bmatrix}, \quad \begin{bmatrix} G_2 \\ P_{01} \\ P_{11} \end{bmatrix} = \begin{bmatrix} Y_{02} \\ Y_{12} \\ Y_{22} \end{bmatrix} + \begin{bmatrix} P_{02} \\ P_{12} \\ P_{22} \end{bmatrix} = \begin{bmatrix} Y_{02} + Y_{13} \\ Y_{12} + Y_{23} \\ Y_{22} + Y_{33} \end{bmatrix},$$

$$\begin{bmatrix} G_1 \\ P_{00} \end{bmatrix} = \begin{bmatrix} Y_{01} \\ Y_{11} \end{bmatrix} + \begin{bmatrix} P_{01} \\ P_{11} \end{bmatrix} = \begin{bmatrix} Y_{01} + Y_{12} + Y_{23} \\ Y_{11} + Y_{22} + Y_{33} \end{bmatrix},$$

$$G_0 = Y_{00} + P_{00} = Y_{00} + Y_{11} + Y_{22} + Y_{33}.$$

■

Corollary 2 (Positive polynomial on the real axis) *Let matrices $R_k \in \mathbf{H}_m$ ($k \in \mathbf{Z}_{2N}$) and $\Phi, \Psi \in \mathbf{H}$ be given and define $\mathbf{\Lambda}$ by (7). Suppose that $\mathbf{\Lambda}$ represents some portion(s) of the real axis or the entire real axis. Let $\mathcal{R} \in \mathbf{H}_{n+m}$ with $n := mN$ be the matrix whose (i,j) block is given by*

$$\mathcal{R}_{ij} := \frac{R_k}{\alpha_k}, \quad k := i+j, \quad \alpha_k := \min(k+1, 2N+1-k)$$

for $i,j \in \mathbf{Z}_N$. Define $F \in \mathbb{R}^{2n\times(n+m)}$ by (29). Then the following statements are equivalent.

(i) $\sum_{k=0}^{2N} \lambda^k R_k \geq 0$ *holds for all $\lambda \in \mathbf{\Lambda}$.*

(ii) There exist $P, Q \in \mathbf{H}_n$ such that $Q \geq 0$ and

$$F^*(\Phi \otimes P + \Psi \otimes Q)F \leq \mathcal{R}. \tag{31}$$

(iii) There exist $X \in \mathbf{H}_{n+m}$ and $Q \in \mathbf{H}_n$ such that

$$R_k = \sum_{i+j=k} Y_{ij}, \quad Y := X + F^*(\Psi \otimes Q)F, \quad X \geq 0, \quad Q \geq 0$$

where $Y_{ij} \in \mathbb{C}^{m\times m}$ ($i,j \in \mathbf{Z}_N$) are the partitioned block matrices of Y, and the summation is taken over all (i,j) such that $i+j=k$ and $i,j \in \mathbf{Z}_N$.

The equivalence also holds when all the inequalities are replaced by strict inequalities with the modification that condition $R_{2N} > 0$ is added to statement (i) if $\mathbf{\Lambda}$ is unbounded.

Proof. We show (i) $\Leftrightarrow$ (ii) first. Notice that, for a real λ,

$$\sum_{k=0}^{2N} \lambda^k R_k = N_\lambda^{\mathsf{T}} \mathcal{R} N_\lambda, \quad N_\lambda := \begin{bmatrix} I & \lambda I & \cdots & \lambda^N I \end{bmatrix}^{\mathsf{T}}. \tag{32}$$

Then, if $\mathbf{\Lambda}$ is bounded, the result follows from Theorem 2 by noting that the range space of N_λ coincides with the null space of $U - \lambda V$ where U and V are defined in (29). The case where $\mathbf{\Lambda}$ is unbounded can be treated in a similar manner to the proof of Theorem 5.

Next we prove (ii) $\Leftrightarrow$ (iii). Since $\mathbf{\Lambda}$ is a subset of real numbers, Φ can be given by

$$\Phi = j\Phi_r, \quad \Phi_r := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Let $\mathcal{P} := F^*(\Phi_r \otimes P)F$ and partition $\mathcal{P}$ into $\mathcal{P}_{ij} \in \mathbb{C}^{m\times m}$ for $i,j \in \mathbf{Z}_N$. Then (31) can be written as

$$\mathcal{R} - j\mathcal{P} - F^*(\Psi \otimes Q)F =: X \geq 0 \quad \text{or} \quad \mathcal{R} = Y + j\mathcal{P}.$$

Taking the summation of certain blocks on both sides, we have

$$R_k = \sum_{i+j=k} \mathcal{R}_{ij} = \sum_{i+j=k} Y_{ij} + j\mathcal{P}_{ij} = \sum_{i+j=k} Y_{ij}$$

where we noted that $\sum_{i+j=k} \mathcal{P}_{ij} = 0$ due to the structure of $\mathcal{P}$. For instance, when $N = 3$, we have

$$\mathcal{R} = \begin{bmatrix} R_o & R_1/2 & R_2/3 & R_3/4 \\ R_1/2 & R_2/3 & R_3/4 & R_4/3 \\ R_2/3 & R_3/4 & R_4/3 & R_5/2 \\ R_3/4 & R_4/3 & R_5/2 & R_6 \end{bmatrix}, \quad \mathcal{P} = \begin{bmatrix} 0 & P_{00} & P_{01} & P_{02} \\ -P_{00} & P_{01}^* - P_{01} & P_{11} - P_{02} & P_{12} \\ -P_{01}^* & P_{02}^* - P_{11} & P_{12}^* - P_{12} & P_{22} \\ -P_{02}^* & -P_{12}^* & -P_{22} & 0 \end{bmatrix}.$$

This proves necessity of (iii). Sufficiency can be verified once we notice that

$$\sum_{k=0}^{2N} \lambda^k R_k = \sum_{k=0}^{2N} \sum_{i+j=k} \lambda^k Y_{ij} = N_\lambda^\mathsf{T} Y N_\lambda = N_\lambda^\mathsf{T} X N_\lambda + \sigma(\lambda, \Psi)(V N_\lambda)^\mathsf{T} Q (V N_\lambda) \geq 0$$

for $\lambda \in \mathbf{\Lambda}$, where N_λ is defined in (32). Finally, the nonstrict inequality case can be proved in exactly the same manner by replacing all the nonstrict inequalities with strict inequalities. ■

Versions of positive-real lemma for polynomials have been obtained in the standard (unrestricted) frequency setting; References [19–21] proved the equivalence (i) ⇔ (iii) in Corollaries 1 and 2 with nonstrict inequalities for matrix-valued polynomials. Our technique for proving statement (iii) is adopted from [19]. On the other hand, results in the restricted frequency setting are limited. References [5, 6] showed the equivalence (i) ⇔ (iii) in Corollary 1 for scalar-valued polynomials (in the discrete-time setting) with nonstrict inequalities. Our results extend and unify these prior results to the case of matrix-valued positive definite polynomials in both continuous-time and discrete-time settings with both strict and nonstrict inequality cases. Finally, we remark that if G_k and Ψ are real, X and Q in statement (iii) of Corollaries 1 and 2 can be restricted to real matrices without loss of generality.

5 Low frequency properties

In this section, we focus on the FDIs in the low frequency range, explicitly show their state space characterizations, and give physical interpretations.

5.1 State space characterizations

The following is the continuous-time result which has been obtained in [4]. It can also be obtained as a special case of Theorem 4 as shown in the previous section.

Corollary 3 *Let real matrices A, B, a real symmetric matrix Θ, and a positive scalar ϖ be given. Suppose (A, B) is controllable. The following statements are equivalent.*

(i) The frequency domain inequality

$$\begin{bmatrix} (sI - A)^{-1}B \\ I \end{bmatrix}^* \Theta \begin{bmatrix} (sI - A)^{-1}B \\ I \end{bmatrix} \leq 0 \tag{33}$$

holds for all $s = j\omega \in \jmath\mathbb{R}$ such that $|\omega| \leq \varpi$.

(ii) There exist real symmetric P and $Q \geq 0$ satisfying

$$\begin{bmatrix} A & B \\ I & 0 \end{bmatrix}^\mathsf{T} \begin{bmatrix} -Q & P \\ P & \varpi^2 Q \end{bmatrix} \begin{bmatrix} A & B \\ I & 0 \end{bmatrix} + \Theta \leq 0. \tag{34}$$

It can be shown that when ϖ approaches infinity, the parameter Q must approach zero and we recover the standard KYP lemma [2, 22, 23].

Next, we shall present a discrete-time counterpart of the above result.

Corollary 4 *Let real matrices A, B, a real symmetric matrix Θ, and a scalar $\vartheta \in (0, \pi)$ be given. Suppose (A, B) is controllable. The following statements are equivalent.*

(i) The finite frequency inequality

$$\begin{bmatrix} (zI-A)^{-1}B \\ I \end{bmatrix}^* \Theta \begin{bmatrix} (zI-A)^{-1}B \\ I \end{bmatrix} \leq 0 \tag{35}$$

holds for all $z = e^{j\theta} \in \mathbb{C}$ such that $|\theta| \leq \vartheta$,

(ii) There exist real symmetric P and $Q \geq 0$ such that

$$\begin{bmatrix} A & B \\ I & 0 \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} -P & Q \\ Q & P-\gamma Q \end{bmatrix} \begin{bmatrix} A & B \\ I & 0 \end{bmatrix} + \Theta \leq 0 \tag{36}$$

where $\gamma := 2\cos\vartheta$.

If we let $\vartheta = \pi$ in Corollary 4, we recover the standard discrete-time KYP lemma [2]. In particular, we have $\gamma = -2$ and the inequality in statement (ii) can be written as

$$\begin{bmatrix} A & B \\ I & 0 \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} -\tilde{P}+Q & Q \\ Q & \tilde{P}+Q \end{bmatrix} \begin{bmatrix} A & B \\ I & 0 \end{bmatrix} + \Theta \leq 0$$

where $\tilde{P} := P + Q$. Since $Q \geq 0$, the term associated with Q is positive semidefinite. Hence, matrix $\tilde{P}$ satisfies this inequality for some $Q \geq 0$ if and only if it does so for $Q = 0$. The condition obtained by letting $Q = 0$ is exactly the one appeared in the standard discrete-time KYP lemma in [2].

5.2 Time domain interpretations

The following theorem provides a physical interpretation for the low frequency condition in Corollary 3 in terms of a time domain characterization of the input-to-state behavior.

Theorem 6 ([4]) *Consider a real symmetric matrix Θ and a stable controllable system*

$$\dot{x} = Ax + Bu.$$

The following statements are equivalent.

(i) The statements (i) and (ii) in Corollary 3 hold.

(ii) The property

$$\int_{-\infty}^{\infty} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^{\mathsf{T}} \Theta \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} dt \leq 0 \tag{37}$$

holds for all square integrable inputs u such that

$$\int_{-\infty}^{\infty} \dot{x}(t)\dot{x}(t)^{\mathsf{T}} dt \leq \varpi^2 \int_{-\infty}^{\infty} x(t)x(t)^{\mathsf{T}} dt. \tag{38}$$

Roughly speaking, Theorem 6 states that the finite frequency condition means that the system possesses the property (37) with respect to the input signal u that does not drive the states too quickly where the bound on the "quickness" is given by ϖ in the sense of (38). It should be noted that condition (38) is coordinate free, i.e. if (38) holds for some minimal realization, then it holds for all other minimal realizations.

The discrete-time result follows.

Theorem 7 *Consider a real symmetric matrix* Θ *and a stable controllable system*

$$x_{k+1} = Ax_k + Bu_k.$$

The following statements are equivalent.

(i) The statements (i) and (ii) in Corollary 4 hold.

(ii) The property

$$\sum_{k=-\infty}^{\infty} \begin{bmatrix} x_k \\ u_k \end{bmatrix}^{\mathsf{T}} \Theta \begin{bmatrix} x_k \\ u_k \end{bmatrix} \leq 0 \tag{39}$$

holds for all square summable inputs u *such that*

$$\sum_{k=-\infty}^{\infty} r_k r_k^{\mathsf{T}} \leq (2-\gamma) \sum_{k=-\infty}^{\infty} x_k x_k^{\mathsf{T}} \tag{40}$$

where $r_k := x_{k+1} - x_k$.

Proof. Suppose (i) holds. Multiplying the inequality in (ii) of Corollary 4 by $[\, x_k^{\mathsf{T}} \;\; u_k^{\mathsf{T}} \,]$ from the left and by its transpose from the right, we have

$$\begin{bmatrix} x_{k+1} \\ x_k \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} -Z-\alpha Q & Q \\ Q & Z-\alpha Q \end{bmatrix} \begin{bmatrix} x_{k+1} \\ x_k \end{bmatrix} + \phi_k \leq 0$$

$$\phi_k := \begin{bmatrix} x_k \\ u_k \end{bmatrix}^{\mathsf{T}} \Theta \begin{bmatrix} x_k \\ u_k \end{bmatrix}$$

or equivalently,

$$x_k^{\mathsf{T}} Z x_k - x_{k+1}^{\mathsf{T}} Z x_{k+1} + \phi_k + \mathrm{tr}[Q((1-\alpha)(x_{k+1}x_{k+1}^{\mathsf{T}} + x_k x_k^{\mathsf{T}}) - r_k r_k^{\mathsf{T}})] \leq 0$$

where $\alpha := \gamma/2$ and $Z := P - \alpha Q$. Taking the summation from $k = -\infty$ to ∞, using the stability property, and completing the square, we have

$$\sum_{k=-\infty}^{\infty} \left(\mathrm{tr}[Q((2-\gamma)x_k x_k^{\mathsf{T}} - r_k r_k^{\mathsf{T}})] + \phi_k \right) \leq 0.$$

Since $Q \geq 0$, the first term on the left hand side is nonnegative and hence we have statement (ii).

To show (ii) $\Rightarrow$ (i), suppose (i) is not true. Denote by $H(z)$ the left hand side of (35). Then there exists $\theta_\star \in \mathbb{R}$ such that $|\theta_\star| \leq \vartheta$ and the maximum eigenvalue of $H(e^{j\theta_\star})$ is strictly positive. Since A, B, and Θ are all real, such $\theta_\star$ can always be chosen to be nonnegative. Thus we have

$$\lambda_{\max}[H(e^{j\theta_\star})] > 0, \quad 0 \leq \theta_\star \leq \vartheta.$$

Now, due to stability of the system, $\lambda_{\max}[H(e^{j\theta})]$ is a continuous function of θ and hence a small perturbation of $\theta_\star$ does not alter the sign of $\lambda_{\max}[H(e^{j\theta})]$. Therefore we can assume $0 < \theta_\star < \vartheta$ without loss of generality.

Let $u_\star$ be the unit eigenvector of $H(e^{j\theta_\star})$ corresponding to the maximum eigenvalue, and define the input signal $u := \eta$ by its z-transform $\hat{\eta}$ as follows:[1]

$$\hat{\eta}(e^{j\theta}) := \begin{cases} u_\star & (\text{when } |\theta - \theta_\star| \le \epsilon) \\ \bar{u}_\star & (\text{when } |\theta + \theta_\star| \le \epsilon) \\ 0 & (\text{otherwise}) \end{cases}$$

where $\bar{u}_\star$ is the complex conjugate of $u_\star$. Now, let $\Phi(z)$ be the transfer function from u to x. Then

$$\sum_{k=-\infty}^{\infty} (2-\gamma)x_k x_k^\mathsf{T} - r_k r_k^\mathsf{T} = \frac{1}{\pi}\int_{-\pi}^{\pi} (\cos\theta - \cos\vartheta)\Phi(e^{j\theta})\hat{\eta}(e^{j\theta})\hat{\eta}(e^{j\theta})^*\Phi(e^{j\theta})^* d\theta$$

which is positive (semi)definite for small $\epsilon > 0$ such that $\theta_\star + \epsilon < \vartheta$. On the other hand, the left hand side of (39) is equal to

$$\sum_{k=-\infty}^{\infty} \begin{bmatrix} x_k \\ u_k \end{bmatrix}^\mathsf{T} \Theta \begin{bmatrix} x_k \\ u_k \end{bmatrix} = \frac{1}{\pi}\int_{\theta_\star-\epsilon}^{\theta_\star+\epsilon} u_\star^* H(e^{j\theta})u_\star d\theta.$$

Since $\psi(\theta) := u_\star^* H(e^{j\theta})u_\star$ is a real-valued, continuous function of θ and $\psi(\theta_\star) > 0$, there exists a sufficiently small $\epsilon > 0$ such that $\psi(\theta) > 0$ for all $\theta \in [\theta_\star - \epsilon, \theta_\star + \epsilon]$. For such ϵ, the above integral takes a positive value, implying that statement (ii) is false. Hence, if (i) is false, (ii) must be false. ■

A comment similar to its continuous-time counterpart applies to Theorem 7. In particular, it shows that the finite frequency condition in statement (i) is equivalent to requiring the property in (39) for any input u such that the rate of change of the state r is bounded as in (40). As the frequency bound ϑ reduces from π to 0, the rate bound $2-\gamma$ reduces from 4 to 0. It can readily be shown by a completion of the square that condition (40) holds for any square summable sequence x when $\vartheta = \pi$. On the other hand, condition (40) becomes $x_k = x_{k+1}$ when $\vartheta = 0$.

6 Frequency response specifications

Consider a transfer function $H(\lambda)$ that depends on the design parameter vector ϱ. Examples of $H(\lambda)$ include the open-loop transfer function $L(s)$, the closed-loop sensitivity functions $S(s)$ and $T(s)$, and the digital filter $F(z)$ as discussed in Section 2. The design problem is to find ϱ such that

$$H(\lambda) \in \Xi, \quad \forall\, \lambda \in \mathbf{\Lambda}$$

where Ξ is a prescribed subset of the complex matrices. More generally, ϱ may be required to satisfy multiple constraints of this type, with different sets Ξ_k in different frequency ranges $\mathbf{\Lambda}_k$ where $k = 1, \ldots, \ell$. We shall address such design problems in the next section. The objective of this section is to show what classes of Ξ can be treated within the framework of our generalized KYP lemma.

Recall that the FDIs arising in various versions of the KYP lemma (theorems in Section 4) are given in terms of quadratic forms. Thus, a set Ξ can be completely captured within our framework if it can be described by a quadratic form. The idea is to construct, from given $H(\lambda)$ and Ξ, a transfer function $G(\lambda)$ and a Hermitian matrix Π such that

$$H \in \Xi \quad \Leftrightarrow \quad \sigma(G, \Pi) \le 0 \tag{41}$$

[1] A simple calculation shows that η_k is given by

$$\eta_k = \frac{1}{k\pi}\Im\left[u_\star(e^{j(\theta_\star+\epsilon)k} - e^{j(\theta_\star-\epsilon)k})\right].$$

for each fixed frequency. If $G(\lambda)$ is a polynomial, then the FDI $\sigma(G,\Pi) \le 0$ in a given frequency range can be characterized by an LMI using Theorem 5. If it is a transfer function $G(\lambda) := C(\lambda E - A)^{-1}B + D$, then the FDI can be expressed as

$$\sigma(G(\lambda),\Pi) = \begin{bmatrix} (\lambda E - A)^{-1}B \\ I \end{bmatrix}^* \Theta \begin{bmatrix} (\lambda E - A)^{-1}B \\ I \end{bmatrix} \le 0,$$

$$\Theta := \begin{bmatrix} C & D \\ 0 & I \end{bmatrix}^* \Pi \begin{bmatrix} C & D \\ 0 & I \end{bmatrix} \tag{42}$$

Thus, the appropriate choice of Θ allows us to treat the FDI within the framework of Theorem 3 (or Theorem 4 when $E = I$). Hence, the question is how to construct G and Π so that (41) holds. We shall answer this for different classes of Ξ in the following subsections.

6.1 Half plane and circular regions

The simplest choice of G is $G := H$. In this case, the set Ξ characterized by the FDI in (41) is

$$\Xi = \{\, H \mid \sigma(H,\Pi) \le 0 \,\}. \tag{43}$$

This set captures some fundamental properties. For instance, the small gain condition is given by

$$\sigma_{\max}[H(\lambda)] \le \gamma : \quad \Pi = \begin{bmatrix} I & 0 \\ 0 & -\gamma^2 I \end{bmatrix}$$

and the positive-real condition is given by

$$H(\lambda)^* + H(\lambda) \ge 0 : \quad \Pi = \begin{bmatrix} 0 & -I \\ -I & 0 \end{bmatrix}.$$

To discuss a more general specification on the Nyquist plot, let us consider the case where $H(\lambda)$ represents an SISO system. In this case, $H(\lambda) \in \Xi$ for all $\lambda \in \mathbf{\Lambda}$ means that the segment of the Nyquist plot specified by the frequency range $\mathbf{\Lambda}$ lies in the region Ξ on the complex plane. From Lemma 1, the region Ξ defined by (43) is nontrivial if and only if $\det(\Pi) < 0$, in which case, it is given by a half plane or a region inside or outside of a circle. Specifically, we have

$$\sigma(G,\Pi) = |G - G_c|^2 - r^2 \le 0, \quad \Pi := \begin{bmatrix} 1 & -G_c \\ -G_c^* & |G_c|^2 - r^2 \end{bmatrix}$$

for a disk of radius r with center at G_c, and

$$\sigma(G,\Pi) = a\Re(G) + b\Im(G) + c, \quad \Pi := \begin{bmatrix} 0 & a + jb \\ a - jb & 2c \end{bmatrix}$$

for a half plane with normal vector $a + jb$. Clearly, the small gain (disk) and the positive-real (right half plane) conditions are characterized by special cases of such regions.

6.2 General conic sections

Next we consider the case where $H(\lambda)$ is a scalar (SISO) transfer function and Ξ is a general conic section on the complex plane[2]. Recall that a conic section can be characterized by

$$\Xi = \left\{ \xi \in \mathbb{C} \; : \; \zeta := \begin{bmatrix} \Re(\xi) \\ \Im(\xi) \end{bmatrix}, \; \sigma(\zeta, \Delta) \le 0 \right\} \tag{44}$$

for some real matrix $\Delta \in \mathbf{H}_3$. The following lemma shows that this characterization can be converted to another which is compatible with our description of the frequency domain inequality.

Lemma 6 *Let a real matrix $\Delta \in \mathbf{H}_3$ be given and consider the set of complex numbers defined in (44). This set can also be characterized by*

$$\Xi = \left\{ \xi \in \mathbb{C} \; : \; \varphi := \begin{bmatrix} \xi^* \\ \xi \end{bmatrix}, \;\; \sigma(\varphi, \Pi) \le 0 \right\}$$

where $\Pi \in \mathbf{H}_3$ is defined by

$$\Pi := \begin{bmatrix} T & 0 \\ 0 & 1 \end{bmatrix}^* \Delta \begin{bmatrix} T & 0 \\ 0 & 1 \end{bmatrix}, \quad T := \frac{1}{2} \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix}.$$

Proof. Note that

$$T\varphi = T \begin{bmatrix} \xi^* \\ \xi \end{bmatrix} = \begin{bmatrix} (\xi + \xi^*)/2 \\ (\xi - \xi^*)/(2j) \end{bmatrix} = \zeta.$$

Then it is straightforward to see that

$$\sigma(\varphi, \Pi) = \sigma(T^{-1}\zeta, \Pi) = \sigma(\zeta, \Delta)$$

from which the result follows. ■

Lemma 6 allows us to establish the equivalence in (41) when Ξ is a conic section by choosing appropriate G and Π. In particular, the transfer function $H(\lambda)$ belongs to the conic section Ξ defined in (44) if and only if Π is chosen as shown in Lemma 6 and $G(\lambda)$ is specified as

$$G(\lambda) := \begin{bmatrix} H^{\sim}(\lambda) \\ H(\lambda) \end{bmatrix}$$

where $H^{\sim}(\lambda)$ is defined by condition $H^{\sim}(\lambda) = H(\lambda)^*$ for $\lambda \in \mathbf{\Lambda}$. A realization of $G(\lambda)$ is given by

$$\left(\begin{array}{cc|c} A^{\sim} & 0 & B^{\sim} \\ 0 & A & B \\ \hline C^{\sim} & 0 & D^{\sim} \\ 0 & C & D \end{array} \right), \qquad \begin{array}{l} H(\lambda) := \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right) \\ H^{\sim}(\lambda) := \left(\begin{array}{c|c} A^{\sim} & B^{\sim} \\ \hline C^{\sim} & D^{\sim} \end{array} \right) \end{array}$$

where $H^{\sim}(\lambda)$ is defined by

$$H^{\sim}(s) := H(-s)^{\mathsf{T}} = \left(\begin{array}{c|c} -A^{\mathsf{T}} & C^{\mathsf{T}} \\ \hline -B^{\mathsf{T}} & D^{\mathsf{T}} \end{array} \right)$$

[2]The developments below can be extended for MIMO square plants, but the physical meaning is not completely clear.

for the continuous-time case and

$$H^{\sim}(z) := H(1/z)^{\mathsf{T}} = \left(\begin{array}{c|c} A^{-\mathsf{T}} & -A^{-\mathsf{T}}C^{\mathsf{T}} \\ \hline B^{\mathsf{T}}A^{-\mathsf{T}} & D^{\mathsf{T}} - B^{\mathsf{T}}A^{-\mathsf{T}}C^{\mathsf{T}} \end{array}\right)$$

for the discrete-time case with nonsingular A. Thus, for each frequency $\lambda \in \mathbb{C}$, the conic section condition $H \in \boldsymbol{\Xi}$ can be converted to $\sigma(G, \Pi) \leq 0$ with the above augmentation, and the latter can further be converted to an LMI via theorems in Section 4. The conic sections captured by (44) are summarized below.

Proposition 3 *Let a real matrix $\Delta \in \mathbf{H}_3$ be given and consider the set of complex numbers $\boldsymbol{\Xi}$ in (44). Partition Δ as*

$$\begin{bmatrix} W & q \\ q^{\mathsf{T}} & r \end{bmatrix} := \Delta, \quad W = W^{\mathsf{T}} \in \mathbb{R}^{2\times 2}, \quad q \in \mathbb{R}^2, \quad r \in \mathbb{R}$$

and denote the spectral decomposition of W by $U\Lambda U^{\mathsf{T}}$ where $UU^{\mathsf{T}} = I$ and $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2)$. Then the set $\boldsymbol{\Xi}$ is nontrivial and has an interior if and only if

$$\lambda_{\max}(\Delta)\lambda_{\min}(\Delta) < 0. \tag{45}$$

In this case, $\boldsymbol{\Xi}$ is given by the set of $\xi \in \mathbb{C}$ such that

$$\begin{bmatrix} \Re(\xi) \\ \Im(\xi) \end{bmatrix} = Ux - W^{\dagger}q$$

which is parametrized by a vector $x \in \mathbb{R}^2$ satisfying

$$x^{\mathsf{T}}\Lambda x + 2\eta^{\mathsf{T}}x + \rho \leq 0, \qquad \begin{array}{l} \eta := (I - \Lambda\Lambda^{\dagger})U^{\mathsf{T}}q, \\ \rho := r - q^{\mathsf{T}}W^{\dagger}q. \end{array}$$

In particular, it is obtained from one of the following 10 regions in the x-plane (see Fig. 2) through a rotation by U and a shift by $-W^{\dagger}q$:

#	η	λ_1	λ_2	ρ	Boundary	Region
1	0	$+$	$+$	$-$	*ellipse*	*inside*
2	0	$-$	$-$	$+$	*ellipse*	*outside*
3	0	$-$	$+$	0	*lines*	*sector*
4	0	$-$	$+$	$-$	*hyperbolas*	*connected*
5	0	$-$	$+$	$+$	*hyperbolas*	*disconnected*
6	0	0	$+$	$-$	*lines*	*convex*
7	0	$-$	0	$+$	*lines*	*nonconvex*
8	$\neq 0$	0	$+$	*any*	*parabola*	*convex*
9	$\neq 0$	$-$	0	*any*	*parabola*	*nonconvex*
10	$\neq 0$	0	0	*any*	*line*	*half plane*

Proof. From Lemma 6, we have

$$\boldsymbol{\Xi} = \left\{ \xi \in \mathbb{C} : \begin{bmatrix} \Re(\xi) \\ \Im(\xi) \end{bmatrix} \in \mathbf{Z} \right\}, \quad \mathbf{Z} := \{ z \in \mathbb{R}^2 : \begin{bmatrix} z \\ 1 \end{bmatrix}^{\mathsf{T}} \Delta \begin{bmatrix} z \\ 1 \end{bmatrix} \leq 0 \}.$$

If $\Delta \leq 0$, then clearly we have $\boldsymbol{\Xi} = \mathbb{C}$. If $\Delta \geq 0$, then

$$\mathbf{Z} = \{ z \in \mathbb{R}^2 : \ Wz + q = 0, \ \ q^{\mathsf{T}}z + r = 0 \}.$$

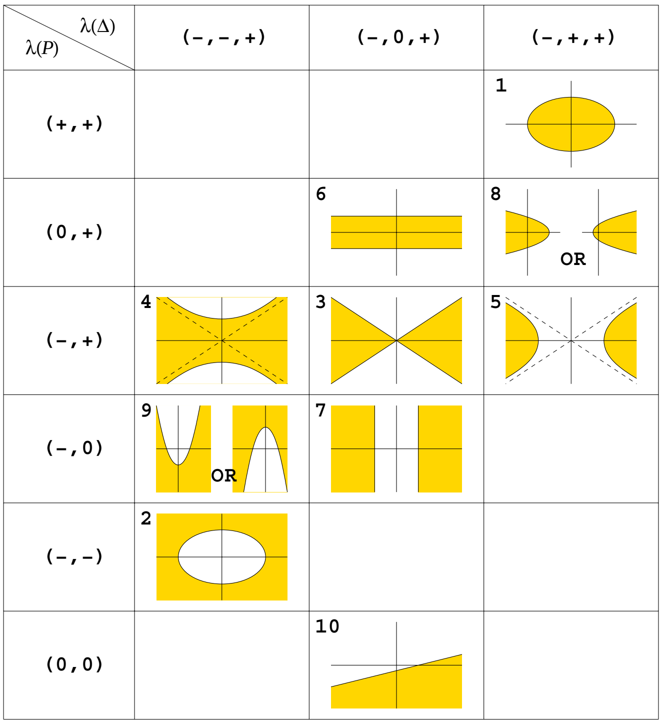


Figure 2: Essential regions for Ξ

The set $\mathbf{Z}$ is nonempty if and only if

$$(I - WW^\dagger)q = 0, \quad r = q^\mathsf{T} W^\dagger q. \tag{46}$$

In this case, the set is characterized as $\mathbf{Z} = \mathbf{Z}_o$ where

$$\mathbf{Z}_o = \{ \, z \in \mathbb{R}^2 : \ \exists \, x \in \mathbb{R}^2 \text{ such that } z = Ux - W^\dagger q, \ \ \Lambda x = 0 \, \}. \tag{47}$$

The necessity of (46) can be seen if we notice that

$$\begin{aligned} &0 = (I - WW^\dagger)(Wz + q) = (I - W^\dagger)q, \\ &0 = q^\mathsf{T} z + r = (WW^\dagger q)^\mathsf{T} z + r = q^\mathsf{T} W^\dagger (Wz) + r = q^\mathsf{T} W^\dagger (-q) + r. \end{aligned}$$

The sufficiency follows if we show $\mathbf{Z} = \mathbf{Z}_o$ under condition (46). If $z \in \mathbf{Z}_o$, we have

$$\begin{aligned} &Wz + q = W(Ux - W^\dagger q) + q = U\Lambda x + (I - WW^\dagger)q = 0, \\ &q^\mathsf{T} z + r = q^\mathsf{T}(Ux - W^\dagger q) + r = q^\mathsf{T} Ux + r - q^\mathsf{T} W^\dagger q = (-Wz)^\mathsf{T} Ux = -z^\mathsf{T} U \Lambda x = 0 \end{aligned}$$

and therefore $z \in \mathbf{Z}$. Conversely, suppose $z \in \mathbf{Z}$. Then, for $x := (I - \Lambda^\dagger \Lambda) U^\mathsf{T} z$, we have

$$\Lambda x = 0, \quad Ux - W^\dagger q = (I - W^\dagger W)z - W^\dagger q = z - W^\dagger (Wz + q) = z$$

and hence $z \in \mathbf{Z}_o$. Note that $\mathbf{Z}_o$ is a point if $\Lambda > 0$, a line if $\Lambda \geq 0$ and $\mathrm{rank}(\Lambda) = 1$, and the entire plane $\mathbb{R}^2$ if $\Lambda = 0$. Thus we conclude that Ξ is trivial if $\Delta \geq 0$ or $\Delta \leq 0$, proving the necessity of (45).

Now, suppose Δ is not definite, i.e., it has positive and negative eigenvalues. The negative eigenvalue implies the existence of a vector $v := [\, v_1 \ v_2 \ v_3 \,]^\mathsf{T} \in \mathbb{R}^3$ such that $v^\mathsf{T} \Delta v < 0$ and $v_3 \neq 0$. Clearly, $z := [\, v_1 \ v_2 \,]/v_3$ defines an interior point of Ξ. Moreover, the positive eigenvalue implies the existence of v such that $v^\mathsf{T} \Delta v > 0$, which in turn implies that $\Xi \neq \mathbb{C}$. Thus Ξ is nontrivial when (45) holds.

To prove the table, note that if $\eta = 0$, then $(I - WW^\dagger)q = 0$ and we have

$$\Delta = \begin{bmatrix} I & 0 \\ q^\mathsf{T} W^\dagger & 1 \end{bmatrix} \begin{bmatrix} W & 0 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} I & W^\dagger q \\ 0 & 1 \end{bmatrix}.$$

Thus every possibilities under condition (45) can be categorized into Cases 1–7 in the table, based on the signs of ρ and the eigenvalues of W. If $\eta \neq 0$, then $(I - WW^\dagger)q \neq 0$ and hence W must be singular. Based on the eigenvalues of W, we have Cases 8–10 in the table. Finally, the claims regarding the various regions can be verified once we notice that $\xi \in \Xi$ if and only if

$$\begin{bmatrix} z \\ 1 \end{bmatrix}^\mathsf{T} \begin{bmatrix} W & q \\ q^\mathsf{T} & r \end{bmatrix} \begin{bmatrix} z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ 1 \end{bmatrix}^\mathsf{T} \begin{bmatrix} \Lambda & \eta \\ \eta^\mathsf{T} & \rho \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} \leq 0$$

where $x := U^\mathsf{T}(z + W^\dagger q)$ and $z := [\, \Re(\xi) \;\; \Im(\xi) \,]^\mathsf{T}$. ■

7 Applications to systems design

This section is devoted to synthesis problems in engineering applications that can be solved using the LMI characterization of the FDI in the generalized KYP lemma.

7.1 Basic idea for synthesis

From the discussion in the previous section, a general synthesis problem can be formulated as follows. For each $k \in \mathbf{Z}_\ell$, let transfer function $G_k(\lambda)$, Hermitian matrix Π_k, and frequency range $\mathbf{\Lambda}_k \subset \mathbb{C}$ be given where $G_k(\lambda)$ depends on the parameter vector ϱ to be designed, and $\mathbf{\Lambda}_k$ is specified by (7) for some $\Phi_k, \Psi_k \in \mathbf{H}$. The problem is to find ϱ such that

$$\sigma(G_k(\lambda), \Pi_k) \leq 0 \quad \forall\, \lambda \in \mathbf{\Lambda}_k \tag{48}$$

for all $k \in \mathbf{Z}_\ell$. We show how the problem can be solved exactly under certain assumptions. For brevity of exposition, let us consider the case $\ell = 1$, i.e. we have just one specification. The case with multiple specifications can be handled similarly. In the sequel, we shall drop the subscript "k" in equation (48).

A standing assumption is that the design parameter vector ϱ appears affinely in the numerator of $G(\lambda)$. If $G(\lambda)$ is a polynomial, then its coefficients are affine functions of ϱ. If $G(\lambda)$ is a proper rational transfer function, then it has a state space realization $G(\lambda) = C(\lambda I - A)^{-1}B + D$ where $C(\varrho)$ and $D(\varrho)$ are affine functions of ϱ while A and B are constant matrices. Another assumption is that the weighting matrix $\Pi \in \mathbf{H}_{m+p}$ is such that

$$\Pi = \begin{bmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{12}^* & \Pi_{22} \end{bmatrix}, \quad \Pi_{11} \in \mathbf{H}_p, \quad \Pi_{11} \geq 0. \tag{49}$$

where m and p are the numbers of inputs and outputs of $G(\lambda)$. This condition makes the set of G satisfying $\sigma(G, \Pi) \leq 0$ convex. Note in particular that, when $m = p = 1$, the design specification (48) defines a convex region in the complex plane in which $G(\lambda)$ is desired to lie. This framework captures such standard specifications as the bounded-real and positive-real properties.

Under these assumptions, the condition in (48) is clearly convex in terms of the design parameter ϱ. Furthermore, theorems in Section 4 can be used to reduce the problem to an LMI. We will explicitly show how this can be done using Theorem 4 for the case where $G(\lambda)$ is a proper transfer function. Other cases where $G(\lambda)$ is a polynomial or a nonproper (descriptor) transfer function can be treated similarly using Theorems 5 and 3, except that the inequality in (48) has to be made strict in the latter case.

Lemma 7 *For a transfer function $G(\lambda) := C(\lambda I - A)^{-1}B + D$, the condition $\sigma(G(\lambda), \Pi) \leq 0,\ \forall \lambda \in \mathbf{\Lambda}$ holds if and only if there exist Hermitian matrices P and $Q \geq 0$ such that*

$$\begin{bmatrix} \Gamma(P,Q,C,D) & \begin{bmatrix} C & D \end{bmatrix}^* \Pi_{11} \\ \Pi_{11} \begin{bmatrix} C & D \end{bmatrix} & -\Pi_{11} \end{bmatrix} \leq 0 \tag{50}$$

holds, where

$$\Gamma(P,Q,C,D) := \begin{bmatrix} A & B \\ I & 0 \end{bmatrix}^* (\Phi \otimes P + \Psi \otimes Q) \begin{bmatrix} A & B \\ I & 0 \end{bmatrix} + \begin{bmatrix} 0 & C^*\Pi_{12} \\ \Pi_{12}^* C & D^*\Pi_{12} + \Pi_{12}^* D + \Pi_{22} \end{bmatrix}$$

provided the assumption (49) holds.

Proof. This result follows directly from Theorem 4 with Θ in (42) and the Schur complement. ∎

When C and D are affine functions of the design parameter vector ϱ, equation (50) defines an LMI in terms of the variables P, Q, and ϱ. Therefore, the design parameter ϱ can be computed via convex programming. The reduction of the synthesis problem (48) to LMIs is rather straightforward once the proper assumptions are imposed. Attention must be paid, however, to when the assumptions are satisfied, and how they can be enforced through re-parametrization of the design parameters. We shall discuss these points for several engineering problems in the sequel.

Structure/control design integration: The specification is given by the FFPR condition in (5), which can be converted to an LMI via Theorem 4. If the transfer function $G(s)$ is an affine function of the design parameter ϱ, then ϱ appears linearly in the LMI, so that the resulting problem can be solved via semidefinite programming. This situation occurs in some practical design problems [4, 10]. The design parameter would appear only in the numerator of $G(s)$ for certain problem settings, including actuator or sensor placements. The design space may then be re-parametrized such that the numerator coefficients are affine functions of the new design parameter. In this case, the design process is reduced to an LMI problem.

Open-loop shaping: The control design via open-loop shaping discussed in Section 2 falls into the above framework if the poles of the controller are fixed *a priori*. An important practical problem that naturally fits in here is the PID control design. The controller transfer function is described by

$$K(s) = k_p + \frac{k_i}{s} + \frac{k_d s}{1 + \epsilon s}, \tag{51}$$

where $\epsilon > 0$ is a small parameter introduced to approximate the differentiator by a proper transfer function. If ϵ is fixed, all the design parameters (k_p, k_i, and k_d) appear linearly in the numerator of the open-loop transfer function $L(s) := K(s)P(s)$. Various specifications on the Nyquist plot of L may be expressed as constraints of the form $L(j\omega) \in \mathbf{\Xi}_k$ for all $\omega \in \mathbf{\Omega}_k$ ($k \in \mathbf{Z}_\ell$) where $\mathbf{\Omega}_k$ is a frequency interval and $\mathbf{\Xi}_k$ is the corresponding convex region on the complex plane such as a disk, half plane, and a conic section. The idea described in Section 6 can be used to represent $L \in \mathbf{\Xi}_k$ in terms of quadratic forms of some (possibly augmented) transfer function G_k as in (48). The design parameters will enter the numerator of G_k linearly, and hence the PID controller can be designed by solving an LMI as in Lemma 7.

Closed-loop control design: This is a fundamental problem of designing a controller to satisfy various constraints on the closed-loop transfer functions $G_k(s)$, including the sensitivity-shaping problem discussed in Section 2 as a special case. With the Youla parametrization of all stabilizing controllers [24], each closed-loop transfer function depends on the free parameter $Q(s)$ in an affine manner. Thus the search for $Q(s)$ is a convex problem, provided the constraints are convex in terms of the closed-loop transfer functions [25]. Since the infinite dimensionality of the parameter space is problematic, one often approximates the space by a finite dimensional space spanned by a selected set of basis functions. The problem then becomes the search for the coefficients of the basis functions, which can be converted to an LMI problem within the above framework if the specifications are given by (48) with each Π_k satisfying (49).

Digital filter design: This is another problem discussed in Section 2. Let us first consider the finite impulse response (FIR) filter synthesis problem. Since the denominator of the FIR filter $F(z) = n(z)/d(z)$ is fixed to $d(z) := z^n$, a state space realization of

$$F(z) = \frac{n_0 + n_1 z + \cdots + n_n z^n}{z^n} \tag{52}$$

may be given by the pair (A, B) in the controllable canonical form and

$$C = \left[\begin{array}{cccc} n_0 & n_1 & \cdots & n_{n-1} \end{array} \right], \; D = n_n$$

where n_i ($i \in \mathbf{Z}_n$) are the coefficients of $n(z)$. Note that A and B are fixed matrices and the design parameters n_i appear linearly in C and D. Moreover, it is not difficult to see that the constraints in (6) are convex in terms of $n(z)$. Therefore, Lemma 7 can be directly applied.

Next we consider the infinite impulse response (IIR) filter design where both $d(z)$ and $n(z)$ are design variables. The problem is fundamentally more difficult than the FIR case because the above specifications are not convex in terms of $d(z)$, and the simple-minded approach described above does not yield LMI problems due to the dependence of A on d_k ($k \in \mathbf{Z}_n$), the coefficients of $d(z)$. However, using the idea of [5,26], the problem can be made tractable when the specifications are given by convex constraints on the gain $|F(e^{j\theta})|$ only. For instance, one can handle the specifications (6) if the third constraint is replaced by

$$1 - \epsilon_p \leq |F(e^{j\theta})| \leq 1 + \epsilon_p, \quad : \quad \vartheta_1 \leq |\theta| \leq \vartheta_2.$$

Such replacement can be reasonable in certain applications, but the general problem with phase constraints still remains open for the IIR case. We shall briefly describe the idea below.

Suppose that each specification has the form $|F(e^{j\theta})| \leq \gamma$ or $|F(e^{j\theta})| \geq \gamma$ for a given $\gamma \in \mathbb{R}$. We explain how the former condition can be reduced to LMIs. The latter, and multiple combinations of these, can be treated similarly. First note that the condition is equivalent to

$$|n(z)|^2 \leq \gamma^2 |d(z)|^2 \tag{53}$$

with z on (some arc of) the unit circle $\mathbf{C}_o$ on the complex plane. This condition is convex in n, but not in d, indicating the difficulty. The idea is to re-parametrize the variables so that the constraint becomes convex in terms of the new variables. The following result due to Fejer and Riesz ([27], Theorem 8.4.5) has been used by [5,6] in a similar context, and is useful for our purpose as well. Below, we denote by $\mathbf{P}_n$ the set of polynomials with real coefficients and degree n.

Lemma 8 *Given $q \in \mathbf{P}_n$, there exists $p \in \mathbf{P}_n$ such that*

$$p(z) + p(1/z) = |q(z)|^2, \quad \forall\, z \in \mathbf{C}_o \tag{54}$$

$$p(z) + p(1/z) \geq 0, \quad \forall\, z \in \mathbf{C}_o \tag{55}$$

where $\mathbf{C}_o$ is the unit circle on the complex plane. Conversely, given $p \in \mathbf{P}_n$ satisfying (55), there exists $q \in \mathbf{P}_n$ such that (54) holds.

This result suggests to replace the variable $d(z)$ by $D(z)$ subject to the constraint $|d(z)|^2 = D(z)+D(1/z) \geq 0$ for $z \in \mathbf{C}_o$, and similarly for $n(z)$. With this change of variables, (53) becomes

$$\begin{array}{l} N(z) + N(1/z) \leq \gamma(D(z) + D(1/z)), \\ D(z) + D(1/z) \geq 0, \quad N(z) + N(1/z) \geq 0. \end{array} \tag{56}$$

Clearly, the condition is now convex (in fact linear) in the coefficients of N and D. These conditions naturally fit into the framework of Corollary 1. Polynomials $D(z)$ and $N(z)$ satisfying the positive-realness condition can be explicitly parametrized by positive semidefinite matrices (X_D, Q_D) and (X_N, Q_N), respectively, as in statement (iii) of Corollary 1. The first inequality condition in (56) can then be converted to an inequality condition via statement (ii) of Corollary 1. Note that variables X_D, Q_D, X_N, and Q_N appear linearly in $\mathcal{G}$ and hence we have an LMI which can be solved numerically. If there are more constraints on the magnitude of the filter, we would simply have more LMIs with variables X_D, Q_D, X_N, Q_N and additional P and Q matrices. Finally, once the LMIs are solved, a minimum phase filter $F(z)$ with the desired properties can be found by stable factorizations of $D(z) + D(1/z)$ and $N(z) + N(1/z)$, which are always possible due to Lemma 8. We remark that, without loss of generality, the coefficient of z^n in $D(z)$ can be set to unity to ensure that $d(z)$ is of n^{th} order and the resulting filter $F(z)$ is proper.

7.2 Design examples

7.2.1 PID controller

We consider the design of PID controllers (51) for a mechanical plant with a lightly damped flexible mode:

$$P(s) = \frac{10}{(s+1)(s^2+s+10)}.$$

The design objective is to have a good tracking performance with a reasonable stability margin and robustness against unmodeled dynamics. Our approach is to shape the Nyquist plot of the open-loop transfer function $L(s) := K(s)P(s)$. In particular, we consider the following specifications:

$$\begin{array}{lll} \text{(a)} & |L(j\omega)| < \gamma_h, & \forall\, |\omega| \ge \varpi_h \\ \text{(b)} & \alpha\Re[L(j\omega)] + \beta > \Im[L(j\omega)], & \forall\, \varpi_\epsilon \le \omega \\ \text{(c)} & \Im[L(j\omega)] < -\gamma_\ell, & \forall\, \varpi_\epsilon \le \omega \le \varpi_\ell. \end{array}$$

Spec.(a) with small γ_h ensures robustness against unmodeled dynamics which typically exists in the high frequency range. Spec.(b) is meant to guarantee a certain stability margin. It may be more natural to require the Nyquist plot to be outside of a circle with its center at the point $-1 + j0$. However, such requirement leads to a nonconvex region on the complex plane and our design method cannot be applied as condition (49) is not satisfied. On the other hand, the above Spec.(b) is a half-plane requirement which can be handled within our framework. The frequency range corresponding to Spec.(b) would be the entire imaginary axis, but a small lower bound $\varpi_\epsilon > 0$ has been introduced to exclude $\omega = 0$ from the range and to avoid a possible numerical difficulty due to the controller pole at the origin. Spec.(c) with a large $\gamma_\ell > 0$ ensures sensitivity reduction in the low frequency range by making $L(s)$ high gain. Again, $|L| > \gamma_\ell$ would be more direct for the purpose but does not define a convex constraint on L. Spec.(c) is a reasonable alternative, given that the phase angle of $L(j\omega)$ approaches $-\pi/2$ as ω approaches 0 from above.

We have designed a PID controller by minimizing β subject to Specs (a), (b), and (c) where the other parameters are fixed as

$$\epsilon = 0.05, \quad \alpha = 3, \quad \gamma_\ell = 2, \quad \gamma_h = 0.1, \quad \varpi_\epsilon = 0.05, \quad \varpi_\ell = 0.45, \quad \varpi_h = 7.$$

The resulting optimal PID controller is found to be

$$k_p = 0.2532, \quad k_i = 1.0604, \quad k_d = 0.1626$$

with the optimal value $\beta = 0.7204$. The Nyquist plot, sensitivity functions, and the step response are depicted in Figs. 3, 4, and 5, respectively. Also plotted for comparison are the result of a popular heuristic PID tuning rule (Ziegler-Nichols ultimate sensitivity method) [28] where the PID parameters are given by

$$k_p = 0.7200, \quad k_i = 0.7601, \quad k_d = 0.1705.$$

The stars and the (small) circles in Fig. 3 indicate the points on the Nyquist plots at frequencies $\omega = \varpi_h$ and ϖ_ℓ, respectively. We see from Fig. 3 that the constraints in Specs (b) and (c) are active for the optimal PID control, and the optimization clearly has improved the stability margin and the sensitivity reduction in the low frequency range. The latter effect is transparent in Fig. 4, which shows that the optimization added more damping to the closed-loop poles than the Ziegler-Nichols design. This can also be seen directly from the oscillatory nature of the step responses plotted in Fig. 5. Overall, this example illustrates that our method allows for the tuning of the PID parameters to achieve desired frequency responses in prescribed frequency ranges.

7.2.2 FIR filter

We consider the design of low pass FIR filters to meet the following specifications:

(a) $|F(e^{j\theta}) - e^{-jd\theta}| \le \epsilon_p \quad \forall\, |\theta| \le \vartheta_p$
(b) $|F(e^{j\theta})| \le \epsilon_s \quad \forall\, |\theta| \ge \vartheta_s$

where $F(z)$ is the filter transfer function to be designed. Spec.(a) requires, within the pass-band $|\theta| \le \vartheta_p$, that the filter $F(z)$ has frequency characteristics similar to the desired function z^{-d} having unity gain and the linear phase with group delay d, where d is a positive integer.[3] Spec.(b) ensures that $F(z)$ has a low gain in the stop-band $|\theta| \ge \vartheta_s$. Let n be the order of the FIR filter to be designed. For fixed n, d, ϵ_s, ϑ_s, and ϑ_p, we design an FIR filter by minimizing ϵ_p subject to the above constraints. This is a Chebyshev approximation problem where the peak value of the frequency response is used as the design criteria.

Most of the currently available methods for solving the Chebyshev approximation problem are based on the frequency gridding [11, 12, 14]. Some recent methods avoid frequency gridding using KYP lemmas [5, 6, 26], but these methods can deal only with a subclass of the problem where all the FDIs are given in terms of the gain of $F(z)$. No methods have been known, to our knowledge, to solve the FIR Chebyshev approximation problem with phase constraints as formulated above, without gridding the frequency axis. Our method naturally handles such problem.

A standard approach in the literature to achieve the linear phase for FIR filters is to make the FIR coefficients symmetric [11, 29]. For instance, when n is even, it can be shown that the constraint on the FIR coefficients $n_{n-i} = n_i$ ($i = 0, 1, \ldots, n/2$) in (52) enforces the linear phase with group delay $n/2$. With this symmetry constraint, Spec.(a) can be equivalently written as a condition on the gain only:

(a') $1 - \epsilon_p \le |F(e^{j\theta})| \le 1 + \epsilon_p \quad \forall\, |\theta| \le \vartheta_p$

where we used the fact that $F(z) = |F(z)|z^{-d}$ holds on the unit circle for a symmetric FIR filter $F(z)$ with $d := n/2$. Thus the design problem is exactly solvable by existing methods [5, 6] under the symmetry constraint.

A drawback of this approach is that some design freedom is wasted to enforce the unnecessary constraint of the linear phase property in the stop-band. By removing the unnecessary constraint, the design freedom may be used to improve the overall performance. We will illustrate this point by numerical examples in the sequel. We fix the frequency ranges and the gain bound in the stop-band as follows:

$$\epsilon_s = 0.01, \quad \vartheta_p = 0.3\pi, \quad \vartheta_s = 0.4\pi$$

and design several FIR filters for different choices of n and d.

Fig. 6(a) shows the gain response of a typical symmetric FIR filter with $n = 30$ and $d = n/2 = 15$, where the frequency is normalized by 2π. This filter is obtained as the optimal solution to the above Chebyshev approximation problem. The optimal performance value is found to be $\epsilon_p = 0.0569$. Now, consider a situation

[3] A transfer function $H(z)$ is said to have a linear phase if it can be represented as $H(e^{j\theta}) = r(\theta)e^{-j\delta\theta}$ for some real-valued function r and for a real positive scalar δ. The parameter δ is called the group delay.

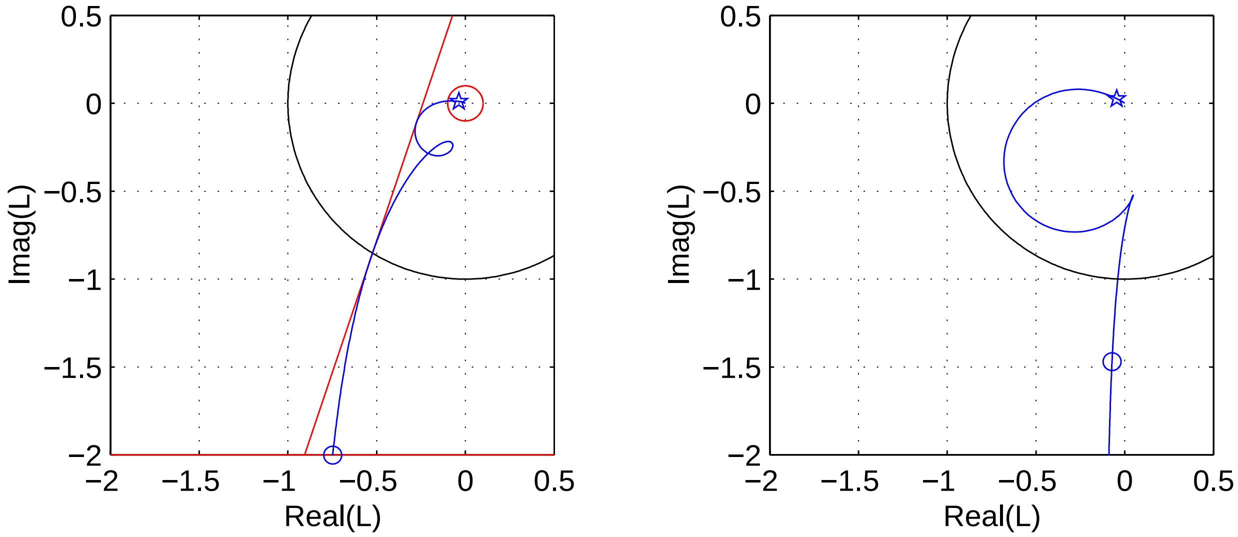


Figure 3: Nyquist plots (Left: optimized, Right: Ziegler-Nichols)

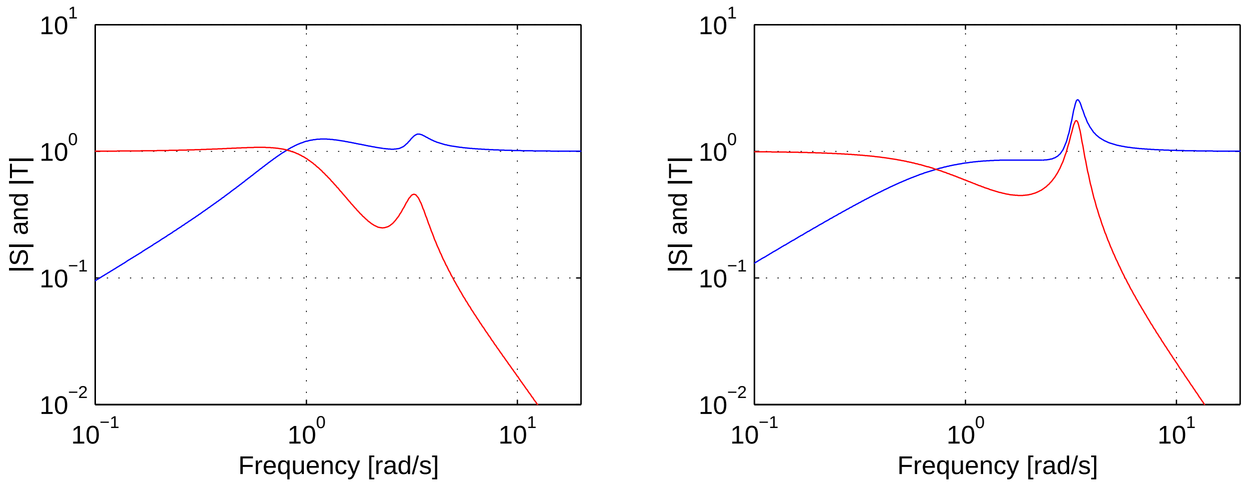


Figure 4: Sensitivity functions (Left: optimized, Right: Ziegler-Nichols)

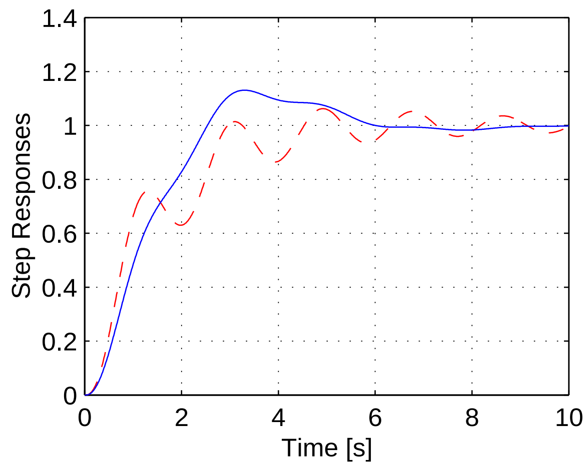
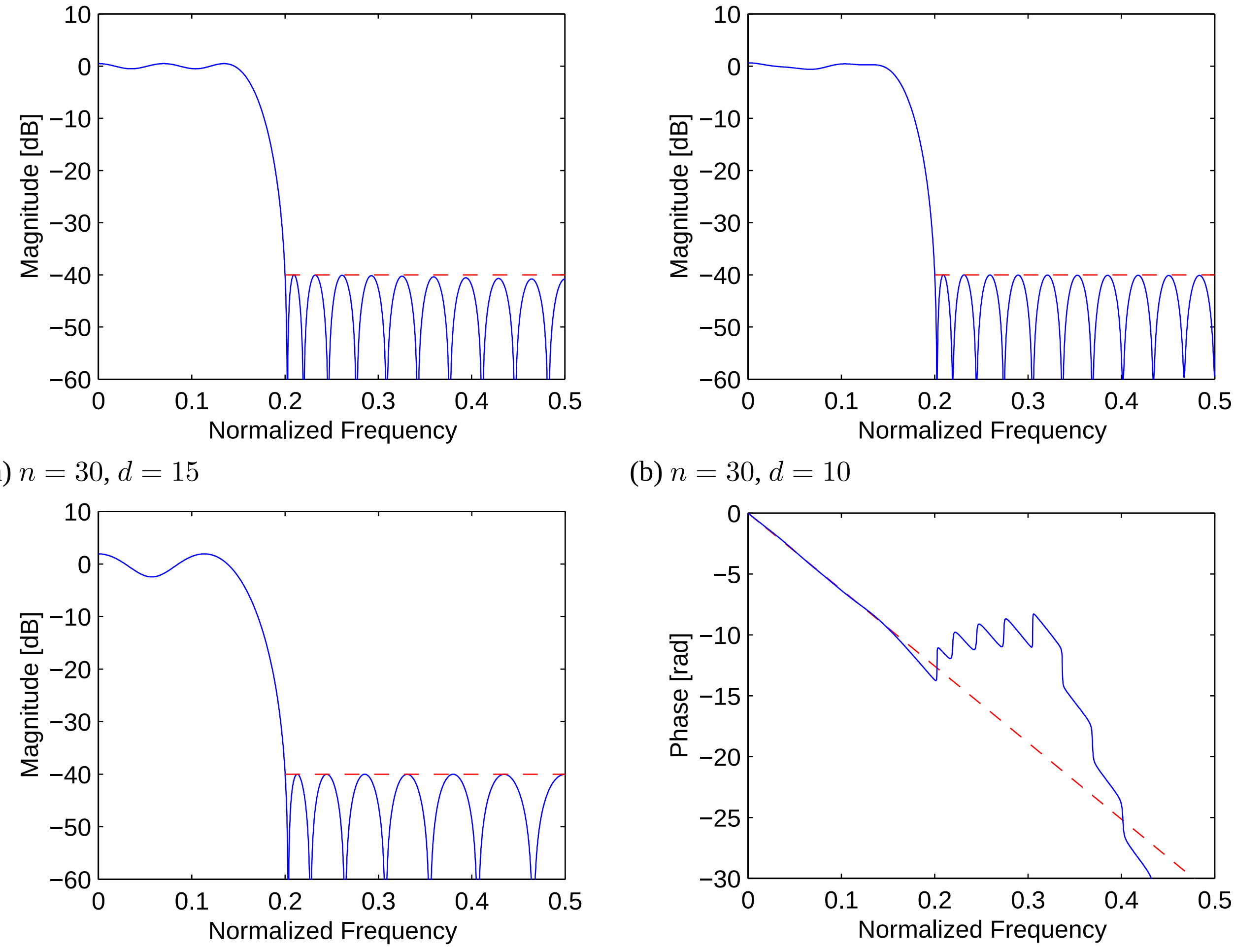


Figure 5: Step responses (Solid: optimized, Dashed: Ziegler-Nichols)

where the group delay of $d = 15$ time steps is too large and is not acceptable. Such small delay requirement would be important for certain real-time applications [14, 30]. Suppose that $d = 10$ is the desired group delay. Then the standard framework of symmetric FIR filters requires that the filter order is $n = 2d = 20$. The gain response of the optimal FIR filter for this case is shown in Fig. 6(c). The optimal performance value is $\epsilon_p = 0.2468$ which is significantly larger than the previous design, as seen by the nonflatness of the gain response in the pass-band. On the other hand, our design method is free from the constraint $d = n/2$ and, for instance, the group delay of $d = 10$ may be (approximately) achieved with a filter of order equal to the original design, i.e., $n = 30$, as shown in Fig. 6(b) and (b'). We see that, within the pass-band, the linear phase requirement is practically met and the gain response is almost as good as the original ($\epsilon_p = 0.0750$). This example shows the advantage of our method that allows us to treat the linear phase constraint in a finite frequency range, as opposed to the existing methods that enforce the constraint for all frequencies by symmetry of the FIR coefficients.



(a) $n = 30$, $d = 15$

(b) $n = 30$, $d = 10$

(c) $n = 20$, $d = 10$

(b') $n = 30$, $d = 10$

Figure 6: FIR responses

8 Conclusion

We have generalized the KYP lemma to allow for more flexibility in various frequency ranges and system classes. In particular, the frequency range is specified by two quadratic forms that define segment(s) of a circle or a line on the complex plane. In this way, an FDI condition can be considered in the low/middle/high frequency ranges for both continuous-time and discrete-time settings in a unified manner. The system class is described by the null space of a certain frequency-dependent matrix, capturing descriptor systems and polynomials in addition to the standard state space systems.

The generalized KYP lemma converts a certain FDI in a finite frequency range to a numerically tractable LMI condition. When system design specifications are expressed in terms of such FDIs, the generalized KYP lemma may be used to reduce the problem to an LMI optimization. We have given some technical conditions under which this reduction is possible, and discussed various engineering applications that fit into the framework, including the design of digital filters, feedback controllers, and mechanical structures. General control problems with state or output feedback, however, still remain open for further research.

Acknowledgments: The authors gratefully acknowledge I. Yamada and H. Hasegawa at Tokyo Institute of Technology (TIT) for their helpful comments on filtering design, and Y. Iwatani at TIT for his help with the PID design example. They are also thankful to C. Scherer at Delft University of Technology for providing them with a preprint of [18] accompanied by his insightful comments. This research is supported in part by The Ministry of Education, Science, Sport and Culture, Japan, under Grant 14550439, and by the 21st Century COE Program on Information Science and Technology Strategic Core.

Appendix

A Miscellaneous results

Lemma 9 *Let $\Phi, \Psi \in \mathbf{H}$ be given. Suppose $\det(\Phi) < 0$. Then the following statements are equivalent.*

(i) $x^*\Psi x \leq 0$ for all $x \in \mathbb{C}^2$ such that $x^*\Phi x = 0$.

(ii) There exists $\tau \in \mathbb{R}$ such that $\Psi + \tau\Phi \leq 0$.

Proof. (ii) ⇒ (i) is obvious. To show the converse, suppose (i) holds. Since $\det(\Phi) < 0$, there exists a nonsingular matrix M such that $\Phi = M^* \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right] M$. Define

$$\Omega := M^{-*}\Psi M^{-1} = \begin{bmatrix} \alpha & \beta e^{j\gamma} \\ \beta e^{-j\gamma} & \delta \end{bmatrix}$$

where $\alpha, \beta, \gamma, \delta \in \mathbb{R}$. Then (i) implies the following equivalent statements:

$$z^*\Omega z \leq 0 \quad \forall z \in \mathbb{C}^2 \quad \text{s.t.} \quad z^* \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} z = 0$$

$$\Leftrightarrow \quad z^*\Omega z \leq 0 \quad \forall z := \begin{bmatrix} e^{j\theta} \\ 1 \end{bmatrix} z_o, \quad z_o \in \mathbb{C},\ \theta \in \mathbb{R}$$

$$\Leftrightarrow \quad \alpha + \delta + 2\beta\cos(\gamma - \theta) \leq 0 \quad \forall \theta \in \mathbb{R}$$

$$\Leftrightarrow \quad \alpha + \delta \leq -2|\beta|$$

Now, letting $\tau := (\delta - \alpha)/2$ and using the last equation, it is straightforward to verify that

$$M^{-*}(\Psi + \tau\Phi)M^{-1} = \begin{bmatrix} (\delta+\alpha)/2 & \beta e^{j\gamma} \\ \beta e^{-j\gamma} & (\delta+\alpha)/2 \end{bmatrix}$$

is negative semidefinite. ■

We remark two things. First, (ii) ⇒ (i) is always true without assuming $\det(\Phi) < 0$. Second, (i) ⇒ (ii) is not true in general if $\det(\Phi) \not< 0$. A counter example is

$$\Phi = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \Psi = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Lemma 10 *Let $\Theta \in \mathbf{H}_\ell$ and $\mathbf{S} \subset \mathbf{H}_\ell$ be given. Suppose that $\mathbf{S}$ possesses Properties (a) and (c) of Definition 3. Then the following statements are equivalent.*

(i) There exists $S \in \mathbf{S}$ such that $\Theta + S \leq 0$.

(ii) For each $\epsilon > 0$, there exists $S \in \mathbf{S}$ such that $\Theta + S \leq \epsilon I$.

Proof. The implication (i) ⇒ (ii) is obvious. To show the converse, consider

(iii) There exists $(\tau, S) \in \mathbf{T}$ such that $\tau\Theta + S \leq 0$, where

$$\mathbf{T} := \{\ (\tau, S) \in \mathbb{R} \times \mathbf{S} \mid \tau \geq 0,\ \tau + \|S\| = 1\ \}.$$

We shall show (ii) ⇒ (iii) ⇒ (i). Suppose (iii) does not hold. Then $\lambda_{\max}(\tau\Theta + S) > 0$ for all $(\tau, S) \in \mathbf{T}$. Since $\mathbf{S}$ is closed, $\mathbf{T}$ is compact and hence there exists $\epsilon \in \mathbb{R}$ such that

$$\lambda_{\max}(\tau\Theta + S) > \epsilon > 0, \quad \forall\ (\tau, S) \in \mathbf{T}.$$

Let $S_o \in \mathbf{S}$ be arbitrary and define

$$\tau := (1 + \|S_o\|)^{-1} \leq 1, \quad S_1 := \tau S_o.$$

Then, we have $S_1 \in \mathbf{S}$ since $\mathbf{S}$ is a cone due to Property (a), and it can readily be verified that $(\tau, S_1) \in \mathbf{T}$. Therefore,

$$\lambda_{\max}(\Theta + S_o) = \frac{1}{\tau}\lambda_{\max}(\tau\Theta + S_1) > \frac{\epsilon}{\tau} \geq \epsilon > 0.$$

Since $S_o \in \mathbf{S}$ is arbitrary, we see that (ii) does not hold. Thus we have shown (ii) ⇒ (iii). Now, suppose (iii) holds and let $(\tau, S_1) \in \mathbf{T}$ be such that $\tau\Theta + S_1 \leq 0$. If $\tau = 0$, then $S_1 \leq 0$ and Property (c) implies $S_1 = 0$, contradicting $(\tau, S_1) \in \mathbf{T}$. So $\tau \neq 0$, in which case, $S_2 := S_1/\tau$ is an element of $\mathbf{S}$ and satisfies $\Theta + S_2 \leq 0$, proving (i). ■

Lemma 11 *[31, 32] Let $\mathbf{X}$ be a convex subset of $\mathbb{C}^m$, and $F : \mathbf{X} \to \mathbb{C}^{n\times n}$ be a Hermitian-valued affine function. The following statements are equivalent.*

(i) The set $\{\ x \mid x \in \mathbf{X},\ F(x) < 0\ \}$ is empty.

(ii) $^\exists$ nonzero $H = H^* \geq 0$ s.t. $\mathrm{tr}(F(x)H) \geq 0\ \ ^\forall x \in \mathbf{X}$.

Lemma 12 *Let matrices* $F \in \mathbb{C}^{2n\times(n+m)}$ *and* $\Theta \in \mathbf{H}_{n+m}$*, and sets of complex numbers* $\mathbf{\Lambda}$ *and* $\mathbf{\Omega}$ *be given such that* $\mathbf{\Omega} \subset \mathbf{\Lambda}$ *and the closure of* $\mathbf{\Lambda}\backslash\mathbf{\Omega}$ *coincides with* $\mathbf{\Lambda}$*. Denote by* N_λ *the null space of* $\Gamma_\lambda F$ *where* Γ_λ *is defined in (15). Suppose*

$$\text{rank}(\Gamma_\lambda F) = n, \quad \forall\, \lambda \in \mathbb{C}. \tag{57}$$

Then $N_\lambda^* \Theta N_\lambda \le 0$ *holds for all* $\lambda \in \mathbf{\Lambda}$ *if and only if it holds for* $\lambda \in \mathbf{\Lambda}\backslash\mathbf{\Omega}$*.*

Proof. The necessity is obvious. To prove the sufficiency, suppose that there exists $\lambda_o \in \mathbf{\Omega}$ such that $N_{\lambda_o}^* \Theta N_{\lambda_o} \le 0$ does not hold. Then there exists a vector ξ_o such that $\xi_o^* \Theta \xi_o > 0$ and $\Gamma_{\lambda_o} F \xi_o = 0$. Let

$$\xi(\lambda) := [I - (\Gamma_\lambda F)^\dagger (\Gamma_\lambda F)]\xi_o.$$

Then we have $\xi(\lambda_o) = \xi_o$ and $\Gamma_\lambda F \xi(\lambda) = 0$ holds for all $\lambda \in \mathbb{C}$. Note that $\xi(\lambda)$ is continuous on $\mathbb{C}$ due to the rank assumption (57), and hence $\xi(\lambda)^* \Theta \xi(\lambda) > 0$ holds if λ is sufficiently close to λ_o. Such λ can be chosen from $\mathbf{\Lambda}\backslash\mathbf{\Omega}$ because $\lambda_o \in \mathbf{\Omega} \subset \mathbf{\Lambda}$ is an element of the closure of $\mathbf{\Lambda}\backslash\mathbf{\Omega}$. Thus we have shown that, if $N_\lambda^* \Theta N_\lambda \le 0$ holds for all $\lambda \in \mathbf{\Lambda}\backslash\mathbf{\Omega}$, then it must also hold for all $\lambda \in \mathbf{\Omega}$. This completes the proof. ∎

B Simultaneous matrix factorization

Lemma 13 *Let* $x, y \in \mathbb{C}$ *be given. Then* $\Im(\bar{x}y) = 0$ *if and only if there exist* $z \in \mathbb{C}$ *and* $x_o, y_o \in \mathbb{R}$ *such that* $x = x_o z$ *and* $y = y_o z$*.*

Proof. Sufficiency is trivially seen. On the other hand, if $x = 0$, we have $z := y$, $x_o := 0$, and $y_o := 1$. If $x \ne 0$, then let $y_o := y/x$, $x_o := 1$, and $z := x$. Note that y_o is real because $\bar{x}y - x\bar{y} = 0$ implies $y/x = \bar{y}/\bar{x}$. ∎

Lemma 14 *Let* $N \in \mathbb{C}^{2\times 2}$ *be given and define*

$$\mathbf{N} := \{\, J^* F J e^{j\theta} \mid F \in \mathbb{R}^{2\times 2},\ \theta \in \mathbb{R},\ \det(F) = 1 \,\}, \quad J := \begin{bmatrix} 1 & 0 \\ 0 & j \end{bmatrix}.$$

Then

$$N^* \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} N = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \Leftrightarrow \quad N \in \mathbf{N}. \tag{58}$$

Proof. Let

$$\begin{bmatrix} e & f \\ g & h \end{bmatrix} := JNJ^*.$$

Then we have

$$N^* \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} N = \begin{bmatrix} j(\bar{g}e - \bar{e}g) & \bar{e}h - \bar{g}f \\ \bar{h}e - \bar{f}g & j(\bar{h}f - \bar{f}h) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

or

$$\Im(\bar{g}e) = 0, \quad \Im(\bar{h}f) = 0, \quad \Im(\bar{h}e - \bar{f}g) = 0, \quad \Re(\bar{h}e - \bar{f}g) = 1.$$

From Lemma 13, the first two conditions are equivalent to the existence of $v, w \in \mathbb{C}$ and $e_o, f_o, g_o, h_o \in \mathbb{R}$ such that

$$e = e_o v, \quad g = g_o v, \quad f = f_o w, \quad h = h_o w.$$

The third condition then becomes

$$(e_o h_o - f_o g_o)\Im(\bar{w}v) = 0$$

which in turn is equivalent to $\Im(\bar{w}v) = 0$ due to $\det(N) \neq 0$. Hence Lemma 13 again implies the existence of $u \in \mathbb{C}$ and $v_o, w_o \in \mathbb{R}$ such that $v = v_o u$ and $w = w_o u$. Thus, letting $u = u_o e^{j\theta}$, we see that the first three conditions hold if and only if

$$\exists\, F \in \mathbb{R}^{2\times 2},\ \theta \in \mathbb{R} \quad \text{such that} \quad \begin{bmatrix} e & f \\ g & h \end{bmatrix} = F e^{j\theta}.$$

Finally, the fourth condition is now equivalent to $1 = e\,\bar{h} - \bar{f}g = \det(F)$. ■

Lemma 15 *Let $\Upsilon \in \mathbf{H}$ be given. Then Υ admits the following factorization:*

$$\Upsilon = N^* \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix} N \tag{59}$$

where $\alpha, \beta, \gamma \in \mathbb{R}$ and $N \in \mathbf{N}_o$ with

$$\mathbf{N}_o := \{\, J^* F J \mid F \in \mathbb{R}^{2\times 2},\ \det(F) = 1 \,\}.$$

In particular, α and γ are the eigenvalues of

$$\Upsilon_o := \Re[J \Upsilon J^*].$$

Proof. Let the entries of Υ be defined by

$$\Upsilon = \begin{bmatrix} x & \beta + jy \\ \beta - jy & z \end{bmatrix}, \quad \beta, x, y, z \in \mathbb{R}.$$

Then

$$\Upsilon_o = \begin{bmatrix} x & y \\ y & z \end{bmatrix} = J\left(\Upsilon - \begin{bmatrix} 0 & \beta \\ \beta & 0 \end{bmatrix}\right) J^*.$$

The spectral factorization of Υ_o gives

$$\Upsilon_o = F^{\mathsf{T}} \begin{bmatrix} \alpha & 0 \\ 0 & \gamma \end{bmatrix} F$$

where the columns of F^{T} are eigenvectors and α, γ are eigenvalues. Note that F, α, and γ are all real. Moreover, since Υ_o is 2×2 real symmetric, F can be chosen to satisfy $\det(F) = 1$. Thus $N := J^* F J$ belongs to $\mathbf{N}_o$. Now, equating the two expressions for Υ_o, we have

$$\Upsilon = N^* \begin{bmatrix} \alpha & 0 \\ 0 & \gamma \end{bmatrix} N + \beta \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Finally, since $N \in \mathbf{N}_o \subset \mathbf{N}$, Lemma 14 can be applied to the second term to obtain the result. ■

C Proofs of the results related to the class $\mathbf{\Lambda}$

Proof of Lemma 1. Suppose $\det(\Delta) = pr - |q|^2 < 0$. If $p = 0$, then $q \neq 0$ and

$$\sigma(\lambda, \Delta) = 2\Re(q^*\lambda) + r \leq 0.$$

Thus both $\mathbf{\Lambda}_\Phi$ and $\mathbf{\Lambda}_\Psi$ are clearly nontrivial, and are given by a straight line and a half plane, respectively. If $p \neq 0$, then

$$\sigma(\lambda, \Delta) = p\left(\left|\lambda + \frac{q}{p}\right|^2 - \frac{|q|^2 - pr}{p^2}\right).$$

Hence $\mathbf{\Lambda}_\Phi$ is the circle $\mathbf{C}(-q/p, \gamma)$, and $\mathbf{\Lambda}_\Psi$ is either the inside or the outside of the circle depending upon the sign of p. Note that the circle is well defined because $\det(\Delta)$ is negative. Thus $\mathbf{\Lambda}_\Phi$ and $\mathbf{\Lambda}_\Psi$ are nontrivial.

Suppose $\det(\Phi) \geq 0$. If $\Phi < 0$ or $\Phi > 0$, then clearly $\mathbf{\Lambda}_\Phi = \emptyset$. If $\Phi = 0$, then $\mathbf{\Lambda}_\Phi = \mathbb{C}$. If $\Phi \geq 0$ or $\Phi \leq 0$ with rank one, then it can be expressed as $\Phi = \pm\phi\phi^*$ for some nonzero vector $\phi := [\ \phi_1\ \ \phi_2\ ]^* \in \mathbb{C}^2$. In this case, $\mathbf{\Lambda}_\Phi = \emptyset$ if $\phi_1 = 0$, and $\mathbf{\Lambda}_\Phi$ has a single element if $\phi_1 \neq 0$.

Suppose $\det(\Psi) \geq 0$. Then Ψ is either positive semidefinite, negative semidefinite with rank one, or negative definite. If $\Psi < 0$, then $\mathbf{\Lambda}_\Psi = \emptyset$. If $\Psi \geq 0$, then $\mathbf{\Lambda}_\Psi = \mathbb{C}$. If $\Psi \leq 0$ with rank one, it can be expressed as $\Psi = -\psi\psi^*$ for some nonzero vector $\psi := [\ \psi_1\ \ \psi_2\ ]^* \in \mathbb{C}^2$. In this case, $\mathbf{\Lambda}_\Psi = \emptyset$ if $\psi_1 = 0$, and $\mathbf{\Lambda}_\Psi$ has a single element if $\psi_1 \neq 0$. ∎

Proof of Lemma 2. Suppose (i) holds. If $\Psi \geq 0$ or $\Psi = k\Phi$ for some $k \in \mathbb{R}$, then $\mathbf{\Lambda}$ is equal to $\mathbf{\Lambda}_\Phi$ and is given by a line or a circle, and hence is nontrivial. So $\Psi \not\geq 0$ and $\Psi \neq k\Phi$ for any $k \in \mathbb{R}$. If $\Psi \leq 0$ and $\Psi \neq 0$, then $\Psi \leq \tau\Phi$ holds for $\tau = 0$. Thus we have (ii). It remains to show that (ii) holds when $\det(\Psi) < 0$.

When $\det(\Phi)$ and $\det(\Psi)$ are both negative, $\mathbf{\Lambda}$ is given by the intersection of a circle or a line $\mathbf{\Lambda}_\Phi$ and a region with a circular or straight boundary $\mathbf{\Lambda}_\Psi$. Now, since $\mathbf{\Lambda}$ is trivial, there is no intersection between $\mathbf{\Lambda}_\Phi$ and the interior of $\mathbf{\Lambda}_\Psi$:

$$\exists \text{ no } \lambda \in \mathbb{C} \text{ s.t. } \sigma(\lambda, \Phi) = 0, \quad \sigma(\lambda, \Psi) > 0,$$

or equivalently,

$$\sigma(\lambda, \Psi) \leq 0 \quad \forall\, \lambda \in \mathbb{C} \text{ s.t. } \sigma(\lambda, \Phi) = 0. \tag{60}$$

We claim that this condition is further equivalent to

$$x^*\Psi x \leq 0 \quad \forall\, s \in \mathbb{C}^2 \text{ s.t. } x^*\Phi x = 0. \tag{61}$$

Once this claim is proven, (61) is equivalent to (ii) by the S-procedure (Lemma 9) in Appendix A, and we complete the proof of (i) ⇒ (ii).

Clearly, (61) implies (60). To show the converse, suppose (60) holds. Let Φ_{11} and Ψ_{11} be the upper left corner entries of Φ and Ψ, respectively. If $\Phi_{11} = 0$, then $\sigma(\lambda, \Phi) = 0$ is a line and the set $\mathbf{\Lambda}_\Phi$ is unbounded. Hence $\Psi_{11} \leq 0$ must be true because otherwise $\sigma(\lambda, \Psi) > 0$ would hold for a $\lambda \in \mathbf{\Lambda}_\Phi$ such that $|\lambda|$ is sufficiently large, contradicting (60). Now, let $x \in \mathbb{C}^2$ be such that $x^*\Phi x = 0$. If $x = 0$, then clearly $x^*\Psi x \leq 0$. So let us assume $x \neq 0$. Partition x as $x = [\ x_1\ \ x_2\ ]^\mathsf{T}$. If $x_2 = 0$, then $x^*\Phi x = |x_1|^2\Phi_{11} = 0$, implying $\Phi_{11} = 0$, and hence $\Psi_{11} \leq 0$ due to the previous discussion. This further implies that $x^*\Psi x = |x_1|^2\Psi_{11} \leq 0$, proving (61). If $x_2 \neq 0$, then for $\lambda := x_1/x_2$, we have $\sigma(\lambda, \Phi) = 0$. Hence (60) implies $\sigma(\lambda, \Psi) \leq 0$ which in turn implies $x^*\Psi x \leq 0$, proving (60) again. Thus (60) ⇒ (61).

We now prove the converse (ii) ⇒ (i). Suppose (ii) holds and fix $\tau \in \mathbb{R}$ such that $\Psi + \tau\Phi \leq 0$. For $\lambda \in \mathbf{\Lambda}$, we have

$$0 \leq \sigma(\lambda, \Psi) = \sigma(\lambda, \Psi + \tau\Phi) \leq 0 \quad \Rightarrow \quad \sigma(\lambda, \Psi + \tau\Phi) = 0.$$

If there were more than one elements in $\mathbf{\Lambda}$, then there exist $\lambda_i \in \mathbb{C}$ ($i = 1, 2$) such that

$$\sigma(\lambda_i, \Psi + \tau\Phi) = 0 \quad \lambda_1 \neq \lambda_2.$$

Since $\Psi + \tau\Phi \leq 0$, this implies

$$(\Psi + \tau\Phi) \left[\begin{array}{cc} \lambda_1 & \lambda_2 \\ 1 & 1 \end{array} \right] = 0$$

and hence we have $\Psi + \tau\Phi = 0$, violating the first condition in (ii). Hence $\mathbf{\Lambda}$ cannot have more than one elements and we conclude that $\mathbf{\Lambda}$ is trivial. ■

Proof of Lemma 3. (a) We first show "⇐". By a congruence transformation,

$$(JK)^{-*}\Psi(JK)^{-1} \leq \tau(JK)^{-*}\Phi(JK)^{-1} = \tau \left[\begin{array}{cc} 0 & -j \\ j & 0 \end{array} \right].$$

Since the real part of a Hermitian negative semidefinite matrix is also negative semidefinite, we obtain $\Psi_\phi \leq 0$ by taking the real part of the both sides. To show "⇒", suppose there is no $\tau \in \mathbb{R}$ such that $\Psi \leq \tau\Phi$. Then

$$\lambda_{\max}(\Psi - \tau\Phi) > 0 \quad \forall\, \tau \in \mathbb{R},$$

where $\lambda_{\max}(\cdot)$ denotes the maximum eigenvalue. By a congruence transformation,

$$\lambda_{\max}\left((JK)^{-*}\Psi(JK)^{-1} - \tau \left[\begin{array}{cc} 0 & -j \\ j & 0 \end{array} \right] \right) > 0 \quad \forall\, \tau \in \mathbb{R}.$$

One can choose τ to kill the imaginary part of $(JK)^{-*}\Psi(JK)^{-1}$, in which case, the matrix becomes Ψ_ϕ and we have $\lambda_{\max}(\Psi_\phi) > 0$. Thus $\Psi_\phi \not\leq 0$.
(b) The result is obtained by replacing Ψ by $-\Psi$ in Part (a). ■

D Proof of Lemma 4

We first state a few results that are useful for the proofs.

Lemma 16 *Consider the set $\mathbf{\Lambda}$ in (7). Suppose it represents curves on the complex plane.* **(This assumption may not be necessary.)** *Then the set $\mathbf{\Lambda}$ is unbounded if and only if $\Phi_{11} = 0$ and $\Psi_{11} \geq 0$.*

Proof. Suppose $\Phi_{11} = 0$ and $\Psi_{11} \geq 0$. Then $\mathbf{\Lambda}_\Phi$ is a straight line, while $\mathbf{\Lambda}_\Psi$ is a half plane or the outside of a circle. Hence, their intersection, $\mathbf{\Lambda}$, is clearly unbounded. This proves the sufficiency. The necessity can be proved by contradiction. If $\Phi_{11} \neq 0$, then $\mathbf{\Lambda}_\Phi$ is a circle and its subset $\mathbf{\Lambda}$ must be bounded. If $\Psi_{11} < 0$, then $\sigma(\lambda, \Psi) < 0$ for λ with sufficiently large magnitude and hence $\mathbf{\Lambda}_\Psi$ is bounded. Therefore, its subset $\mathbf{\Lambda}$ must be bounded. ■

Lemma 17 *Let $f, g \in \mathbb{C}^n$ be given. The following statements are equivalent.*

(i) $fg^* + gf^* = 0$

(ii) There exist $\sigma, \omega \in \mathbb{R}$ and $h \in \mathbb{C}^n$ such that $f = j\omega h$ and $g = \sigma h$.

Proof. The implication (ii) $\Rightarrow$ (i) can be verified by direct substitution. To show the converse, suppose (i) holds. If $g = 0$, then $h := jf$, $\omega := -1$, and $\sigma := 0$ proves that (ii) holds. If $g \neq 0$, then

$$(fg^* + gf^*)g = 0 \quad \Rightarrow \quad f = -\frac{f^*g}{g^*g} \cdot g.$$

Choosing $h := g$, $\sigma := 1$, and $\omega := j(f^*g)/(g^*g)$ proves (ii) where we note that ω is real because $2\mathrm{Re}(f^*g) = f^*g + g^*f = \mathrm{tr}(gf^* + fg^*) = 0$. ■

We are now ready to prove Lemma 4.

Proof of Lemma 4. Let $\xi \in \bigcup_{\lambda \in \bar{\mathbf{\Lambda}}} \mathbf{G}_\lambda$. Then $\xi \neq 0$ and there exists $\lambda \in \bar{\mathbf{\Lambda}}$ such that $\Gamma_\lambda F\xi = 0$. We shall show $\xi \in \mathbf{G}$ to conclude that $\mathbf{G} \supseteq \bigcup_{\lambda \in \bar{\mathbf{\Lambda}}} \mathbf{G}_\lambda$ for each of the following two cases.

Case: $\lambda \in \mathbf{\Lambda}$ Note that

$$\Gamma_\lambda F\xi = 0 \quad \Rightarrow \quad F\xi = \begin{bmatrix} \lambda I \\ I \end{bmatrix} \eta \quad \text{for some } \eta \in \mathbb{C}^n.$$

Hence we have

$$\xi^* F^* L(P, Q) F\xi = \eta^* \begin{bmatrix} \lambda I \\ I \end{bmatrix}^* L(P, Q) \begin{bmatrix} \lambda I \\ I \end{bmatrix} \eta = \eta^* \left(\sigma(\lambda, \Phi)P + \sigma(\lambda, \Psi)Q\right) \eta \geq 0$$

for all $P = P^*$ and $Q = Q^* \geq 0$. Thus $\xi \in \mathbf{G}$.

Case: $\lambda = \infty$ This is the case only if $\mathbf{\Lambda}$ is unbounded. By Lemma 16, we have $\Phi_{11} = 0$ and $\Psi_{11} \geq 0$. Note that

$$\Gamma_\lambda F\xi = 0 \quad \Rightarrow \quad F\xi = \begin{bmatrix} f \\ 0 \end{bmatrix} \quad \text{for some } f \in \mathbb{C}^n.$$

Hence

$$\xi^* F^* L(P, Q) F\xi = \Phi_{11}(f^* P f) + \Psi_{11}(f^* Q f) \geq 0 \tag{62}$$

for all $P = P^*$ and $Q = Q^* \geq 0$. Thus we have $\xi \in \mathbf{G}$.

Next we show $\mathbf{G} \subseteq \bigcup_{\lambda \in \bar{\mathbf{\Lambda}}} \mathbf{G}_\lambda$. Let $\xi \in \mathbf{G}$ and define

$$\begin{bmatrix} f \\ g \end{bmatrix} := F\xi.$$

Case: $g = 0$ If $f = 0$, then

$$\Gamma_\lambda \xi = f - \lambda g = 0$$

for all $\lambda \in \mathbb{C}$ and hence $\xi \in \mathbf{G}_\lambda$ for all $\lambda \in \mathbf{\Lambda}$. Suppose $f \neq 0$. Note that $\xi \in \mathbf{G}$ implies (62) for all $P = P^*$ and $Q = Q^* \geq 0$, which in turn implies $\Phi_{11} = 0$ and $\Psi_{11} \geq 0$. By Lemma 16, we see that $\mathbf{\Lambda}$ is unbounded. Then $\lambda \in \bar{\mathbf{\Lambda}}$ can be chosen to be infinity, and we have $\Gamma_\infty F\xi = -g = 0$, proving that $\xi \in \mathbf{G}_\infty$.

Case: $g \neq 0$ Define

$$\begin{bmatrix} \hat{f} \\ \hat{g} \end{bmatrix} := (M \otimes I)F\xi = \begin{bmatrix} aI & bI \\ cI & dI \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix}.$$

Then, by the simultaneous factorizations of Φ and Ψ in Proposition 2, we have

$$\begin{aligned}(F\xi)^*L(P,Q)(F\xi) &= \begin{bmatrix} \hat{f} \\ \hat{g} \end{bmatrix}^* \begin{bmatrix} \alpha Q & P+\beta Q \\ P+\beta Q & \gamma Q \end{bmatrix} \begin{bmatrix} \hat{f} \\ \hat{g} \end{bmatrix} \\ &= \text{tr}\,[(\hat{g}\hat{f}^* + \hat{f}\hat{g}^*)P + (\alpha\hat{f}\hat{f}^* + \beta\hat{g}\hat{f}^* + \beta\hat{f}\hat{g}^* + \gamma\hat{g}\hat{g}^*)Q] \geq 0\end{aligned}$$

for all $P = P^*$ and $Q = Q^* \geq 0$, implying that

$$\hat{g}\hat{f}^* + \hat{f}\hat{g}^* = 0, \quad \begin{bmatrix} \hat{f} & \hat{g} \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix} \begin{bmatrix} \hat{f}^* \\ \hat{g}^* \end{bmatrix} \geq 0.$$

The first equation implies, by Lemma 17, that there exist $\sigma, \omega \in \mathbb{R}$ and $h \in \mathbb{C}^n$ such that

$$\hat{f} = -j\omega h, \quad \hat{g} = \sigma h.$$

The second equation then becomes

$$\begin{bmatrix} j\omega \\ \sigma \end{bmatrix}^* \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix} \begin{bmatrix} j\omega \\ \sigma \end{bmatrix} hh^* \geq 0 \tag{63}$$

Since g is nonzero, so is h, and hence the coefficient of hh^* must be nonnegative. Note that

$$\begin{bmatrix} f \\ g \end{bmatrix} = (M \otimes I)^{-1} \begin{bmatrix} \hat{f} \\ \hat{g} \end{bmatrix} = \begin{bmatrix} (j\omega\tilde{a} + \sigma\tilde{b})h \\ (j\omega\tilde{c} + \sigma\tilde{d})h \end{bmatrix}, \quad \begin{bmatrix} \tilde{a} & \tilde{b} \\ \tilde{c} & \tilde{d} \end{bmatrix} := M^{-1}.$$

Since $g \neq 0$, we have $j\omega\tilde{c} + \sigma\tilde{d} \neq 0$ and hence

$$\lambda := \frac{j\omega\tilde{a} + \sigma\tilde{b}}{j\omega\tilde{c} + \sigma\tilde{d}}$$

is well defined. Clearly,

$$\Gamma_\lambda \xi = f - \lambda g = 0$$

holds and hence $\xi \in \mathbf{G}_\lambda$. It now remains to show that $\lambda \in \mathbf{\Lambda}$. Noting that

$$M \begin{bmatrix} \lambda \\ 1 \end{bmatrix} (j\omega\tilde{c} + \sigma\tilde{d}) = M \begin{bmatrix} \tilde{a} & \tilde{b} \\ \tilde{c} & \tilde{d} \end{bmatrix} \begin{bmatrix} j\omega \\ \sigma \end{bmatrix} = \begin{bmatrix} j\omega \\ \sigma \end{bmatrix},$$

it is straightforward to verify that

$$|j\omega\tilde{c} + \sigma\tilde{d}|^2 \begin{bmatrix} \lambda \\ 1 \end{bmatrix}^* \Phi \begin{bmatrix} \lambda \\ 1 \end{bmatrix} = \begin{bmatrix} j\omega \\ \sigma \end{bmatrix}^* \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} j\omega \\ \sigma \end{bmatrix} = 0,$$

$$|j\omega\tilde{c} + \sigma\tilde{d}|^2 \begin{bmatrix} \lambda \\ 1 \end{bmatrix}^* \Psi \begin{bmatrix} \lambda \\ 1 \end{bmatrix} = \begin{bmatrix} j\omega \\ \sigma \end{bmatrix}^* \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix} \begin{bmatrix} j\omega \\ \sigma \end{bmatrix} \geq 0$$

where the last inequality follows from (63) and $h \neq 0$. Thus we have $\lambda \in \mathbf{\Lambda}$. ■

E Proof of Lemma 5

We first state several results that are useful for the proofs.

Lemma 18 *Let $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{n\times m}$, and*

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{C}^{2\times 2}$$

be given. Suppose M is nonsingular and A has no eigenvalues λ such that $\sigma(\lambda, M^*\Phi_o M) = 0$ where Φ_o is defined in (11). Then

$$\det(dI + cA) \neq 0$$

and the following matrices are well defined:

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} := \begin{bmatrix} (bI + aA)\aleph & (ad - bc)\aleph B \\ \aleph & -c\aleph B \end{bmatrix} \tag{64}$$

$$\aleph := (dI + cA)^{-1}. \tag{65}$$

In this case, we have

$$(a) \quad \det(sI - \mathcal{A}) \neq 0, \quad \forall\, s \in j\mathbb{R},$$

$$(b) \quad (A, B) \text{ controllable} \quad \Rightarrow \quad (\mathcal{A}, \mathcal{B}) \text{ controllable}.$$

Moreover, we have

$$\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ I & 0 \end{bmatrix} \begin{bmatrix} \mathcal{C} & \mathcal{D} \\ 0 & I \end{bmatrix}^{-1} = \begin{bmatrix} aI & bI \\ cI & dI \end{bmatrix} \begin{bmatrix} A & B \\ I & 0 \end{bmatrix}. \tag{66}$$

Proof. Suppose $\det(dI + cA) = 0$. If $c = 0$, this implies $d = 0$, contradicting $\det(M) \neq 0$. Hence $c \neq 0$ and A should have an eigenvalue at $\lambda_1 := -d/c$. However, it can be verified that $\sigma(\lambda_1, M^*\Phi_o M) = 0$. By contradiction, $\det(dI + cA) \neq 0$.

Suppose $\det(sI - \mathcal{A}) = 0$ for some $s \in j\mathbb{R}$. Then

$$\det(sI - (bI + aA)\aleph) = 0 \quad \Rightarrow \quad \det((ds - b)I + (cs - a)A) = 0.$$

If $cs = a$, then $ds = b$ and hence $c(ds - b) = \det(M) = 0$ which cannot be true because M is nonsingular. Hence $cs \neq a$. In this case, A has an eigenvalue at $\lambda_2 := (b - ds)/(cs - a)$ which satisfies $\sigma(\lambda_2, M^*\Phi_o M) = 0$. By contradiction, $\det(sI - \mathcal{A}) \neq 0$ for all $s \in j\mathbb{R}$, i.e., (a) is true.

The statement (b) can be also shown by contradiction. Suppose $(\mathcal{A}, \mathcal{B})$ is uncontrollable. Then there exists $s \in \mathbb{C}$ such that

$$\text{rank}(W) < n, \quad W := \aleph^{-1} \begin{bmatrix} \mathcal{A} - sI & \mathcal{B} \end{bmatrix} = \begin{bmatrix} (b - ds)I + (a - cs)A & \det(M)B \end{bmatrix}.$$

If $a = cs$ were true, we have $ds \neq b$ as before, implying $\text{rank}(W) = n$. Hence $a \neq cs$. Thus $(\mathcal{A}, \mathcal{B})$ has an uncontrollable mode at $\lambda_2 = (b - ds)/(cs - a)$. Finally, it is tedious but straightforward to verify the identity in (66). ∎

We note that the matrix M in the above lemma corresponds to a bilinear transform

$$\tilde{s} = \frac{b - ds}{cs - a},$$

which transform (A, B) into $(\mathcal{A}, \mathcal{B})$.

Proof of Lemma 5.
(i) Define

$$L(P, Q) := \Phi \otimes P + \Psi \otimes Q. \tag{67}$$

From Proposition 2, we have

$$L(P, Q) = (M \otimes I)^* \left[\begin{array}{cc} \alpha Q & P + \beta Q \\ P + \beta Q & \gamma Q \end{array} \right] (M \otimes I)$$

where $\alpha < 0 < \gamma$ or $0 \leq \alpha \leq \gamma$ can be assumed. In the former case, defining

$$W := \text{diag}(\sqrt{-\alpha}, \sqrt{\gamma})(M \otimes I)F, \quad X := (P + \beta Q)/\sqrt{-\alpha\gamma}, \quad Y := Q,$$

the set $\mathbf{S}$ can be characterized as

$$\mathbf{S} = \{ \, W^* \left[\begin{array}{cc} -Y & X \\ X & Y \end{array} \right] W \mid X = X^*, \; Y = Y^* \geq 0 \, \}.$$

It then follows from Lemma 3 of [3] that this set $\mathbf{S}$ has Properties (a) and (b). We now consider the latter case, i.e., $0 \leq \alpha \leq \gamma$. It can be verified that a matrix $H \in \mathbf{H}_\ell$ ($\ell := n + m$) satisfies $\text{tr}(SH) \geq 0$ for all $S \in \mathbf{S}$ if and only if

$$\left\{ \begin{array}{l} \alpha U + \gamma W \geq 0, \\ V + V^* = 0 \end{array} \right. \qquad \left[\begin{array}{cc} U & V \\ V^* & W \end{array} \right] := (M \otimes I)FHF^*(M \otimes I)^*.$$

For $H \geq 0$ with rank $r > 0$, Lemma 5 of [2] implies that there exist ξ_i ($i = 1, \ldots, r$) such that

$$H = \sum_{k=1}^{r} \xi_i \xi_i^*, \quad V_i + V_i^* = 0$$

where V_i is defined in the same way as V except that H is replaced by $H_i := \xi_i \xi_i^*$. It is then straightforward to verify that $\xi_i^* S \xi_i \geq 0$ for all $S \in \mathbf{S}$.
(ii) We will first show the "only if" part of the proof. Suppose, for contradiction, that

$$F^* L(P, Q) F \leq 0, \quad P, Q \in \mathbf{H}_n, \quad Q \geq 0 \quad \Rightarrow \quad P = Q = 0 \tag{68}$$

but there exists $\lambda \in \mathbb{C} \cup \{\infty\}$ such that $\text{rank}(\Gamma_\lambda F) < n$. Partition F as $F^* = [\, F_1^* \;\; F_2^* \,]$ so that each of F_1 and F_2 has n rows. If $\lambda = \infty$, then there exists a nonzero vector $v_\infty \in \mathbb{C}^n$ satisfying $v_\infty^* F_2 = 0$. Choose a nonzero scalar $p_\infty \in \mathbb{R}$ such that $p_\infty \Phi_{11} \leq 0$ and let $P_\infty := p_\infty v_\infty v_\infty^*$ and $Q_\infty := 0$. Then

$$S_\infty := F^* L(P_\infty, Q_\infty) F = (p_\infty \Phi_{11}) F_1^* v_\infty v_\infty^* F_1 \leq 0,$$

contradicting (68). Hence $\lambda \neq \infty$ and $\lambda \in \mathbb{C}$. In this case, there exists a nonzero vector $v \in \mathbb{C}^n$ satisfying $v^* F_1 = \lambda v^* F_2$. Choose a nonzero scalar $p \in \mathbb{R}$ such that $p\sigma(\lambda, \Phi) \leq 0$ and let $P := pvv^*$ and $Q := 0$. Then

$$S := F^* L(P, QF = p\sigma(\lambda, \Phi) F_2^* vv^* F_2 \leq 0,$$

which again contradicts (68). This completes the proof of necessity.

To prove the converse, suppose that the rank condition (18) is satisfied. Then F_2 has full row rank and hence there exists a nonsingular matrix T that transforms F into the following form:

$$FT = \begin{bmatrix} A_o & B \\ I_n & 0 \end{bmatrix}, \quad A_o \in \mathbb{C}^{n\times n}, \quad B \in \mathbb{C}^{n\times m}.$$

Note that (A_o, B) is controllable because

$$\text{rank} \begin{bmatrix} A_o - \lambda I & B \end{bmatrix} = \text{rank}(\Gamma_\lambda FT) = n, \quad \forall\, \lambda \in \mathbb{C}.$$

Let $P, Q \in \mathbf{H}_n$ be such that $S := F^* L(P,Q) F \le 0$ and $Q \ge 0$. We will show that $P = Q = 0$. Since (A_o, B) is controllable, there exists K such that $A := A_o + BK$ has no eigenvalues in $\mathbf{\Lambda}_\Phi$. Then

$$\begin{bmatrix} I & 0 \\ K & I \end{bmatrix}^* T^* S T \begin{bmatrix} I & 0 \\ K & I \end{bmatrix} = \begin{bmatrix} A & B \\ I & 0 \end{bmatrix}^* L(P,Q) \begin{bmatrix} A & B \\ I & 0 \end{bmatrix} \le 0.$$

Note from Proposition 2 that $L(P,Q)$ can be expressed as

$$L(P,Q) = (M \otimes I)^* (\Phi_o \otimes P + \Psi_o \otimes Q)(M \otimes I).$$

Then, in view of Lemma 18, we see that a congruence transformation yields

$$\Sigma := \mathcal{F}^* (\Phi_o \otimes P + \Psi_o \otimes Q) \mathcal{F} \le 0, \quad \mathcal{F} := \begin{bmatrix} \mathcal{A} & \mathcal{B} \\ I & 0 \end{bmatrix}$$

where $(\mathcal{A}, \mathcal{B})$ is controllable and $\det(sI - \mathcal{A}) \neq 0$ for all $s \in j\mathbb{R}$. Let

$$\mathbf{\Omega} := \{\, \omega \in \mathbb{R} \mid \sigma(j\omega, \Psi_o) > 0 \,\}.$$

Note that the set $\mathbf{\Omega}$ is nonempty since $\mathbf{\Lambda}$ represents curves. Now, for $R(s) := (sI - \mathcal{A})^{-1}\mathcal{B}$, $s \in j\mathbf{\Omega}$, we have

$$\sigma(R(s), \Sigma) = \sigma(s, \Psi_o) R(s)^* Q R(s) \le 0.$$

Note that $\sigma(s, \Psi_o) > 0$ due to $s \in j\mathbf{\Omega}$. This means that each numerator polynomials of the transfer function $QR(s)$ has an infinitely many roots in $j\mathbf{\Omega}$, which in turn implies that $QR(s)$ is identically zero. In this case, we have $Q\mathcal{A}^k\mathcal{B} = 0$ for all nonnegative integers k. Hence, by controllability of $(\mathcal{A}, \mathcal{B})$, we conclude that $Q = 0$. Finally, it follows from [2, 33] that $\Sigma \le 0$ with $Q = 0$ implies $P = 0$ under controllability of $(\mathcal{A}, \mathcal{B})$. ■

References

[1] B. D. O. Anderson. A system theory criterion for positive real matrices. *SIAM J. Contr.*, 5:171–182, 1967.

[2] A. Rantzer. On the Kalman-Yakubovich-Popov lemma. *Sys. Contr. Lett.*, 28(1):7–10, 1996.

[3] T. Iwasaki, G. Meinsma, and M. Fu. Generalized S-procedure and finite frequency KYP lemma. *Mathematical Problems in Engineering*, 6:305–320, 2000.

[4] T. Iwasaki, S. Hara, and H. Yamauchi. Dynamical system design from a control perspective: Finite frequency positive-realness approach. *IEEE Trans. Auto. Contr.*, 48(8), August 2003. (To appear).

[5] Y. Genin, Y. Hachez, Yu. Nesterov, and P. Van Dooren. Convex optimization over positive polynomials and filter design. *Proc. 2000 UKACC Int. Conf. Contr.*, 2000.

[6] T. Davidson, Z. Luo, and J. Sturm. Linear matrix inequality formulation of spectral mask constraints. *Proc. ICASSP*, 2001.

[7] A. Messac and K. Malec. Control-structure integrated design. *AIAA Journal*, 30(8):2124–2131, 1992.

[8] P. Maghami, S. Joshi, and D. Price. Integrated controls-structures design methodology for flexible space craft. *J. Spacecrafts and Rockets*, 32(5):839–844, 1995.

[9] K. M. Grigoriadis, G. Zhu, and R. E. Skelton. Optimal redesign of linear systems. *ASME J. Dyn. Syst. Meas. Contr.*, 118:596–605, 1996.

[10] S. Hara, T. Iwasaki, and F. Shimizu. Finite frequency characterization of easily controllable mechanical systems under control effort constraint. *Proc. IFAC World Congress*, 2002.

[11] T.W. Parks and J.H. McClellan. Chebyshev approximation for nonrecursive digital filters with linear phase. *IEEE Trans. Circuit Theory*, 19(2):189–194, 1972.

[12] X. Chen and T.W. Parks. Design of FIR filters in the complex domain. *IEEE Trans. Acoustics, Speech, and Signal Processing*, ASSP-35(2):144–153, February 1987.

[13] K. Preuss. On the design of FIR filters by complex Chebyshev approximation. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 37(5):702–712, 1989.

[14] L.J. Karam and J.H. McClellan. Chebyshev digital FIR filter design. *Signal Processing*, 76:17–36, 1999.

[15] K. Zhou, J. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice Hall, 1996.

[16] T. Glad and L. Ljung. *Control Theory — Multivariable and Nonlinear Methods*. Taylor & Francis, 2000.

[17] V. A. Yakubovič. S-procedure in nonlinear control theory. *Vestnik Leningrad Univ.*, 1:62–77, 1971.

[18] C. Scherer. When are multiplier relaxations exact? *IFAC Sympo. Robust Contr.*, 2003. (Submitted preprint).

[19] Y. Genin, Yu. Nesterov, and P. Van Dooren. Positive transfer functions and convex optimization. *Proc. Euro. Contr. Conf.*, 1999.

[20] Y. Genin, Y. Hachez, Yu. Nesterov, and P. Van Dooren. Optimization problems over positive pseudo-polynomial matrices. *SIAM J. Mat. Anal. Appl.*, 2001. (Submitted).

[21] Y. Genin, Yu. Nesterov, R. Stefan, P. Van Dooren, and S. Xu. Positivity and linear matrix inequalities. *Euro. J. Contr.*, 8(3):275–298, 2002.

[22] B. D. O. Anderson and S. Vongpanitlerd. *Network Analysis and Synthesis*. Prentice Hall, 1973.

[23] J. C. Willems. Least squares stationary optimal control and the algebraic Riccati equation. *IEEE Trans. Auto. Contr.*, 16:621–634, 1971.

[24] D. C. Youla, H. A. Jabr, and J. J. Bongiorno. Modern Wiener-Hopf design of optimal controllers: Part 2. *IEEE Trans. Auto. Contr.*, 21:319–338, 1976.

[25] S Boyd and C. H. Barratt. *Linear Controller Design: Limits of Performance*. Prentice Hall, Englewood Cliffs, 1991.

[26] J. Tuqan and P.P. Vaidyanathan. A state space approach to the design of globally optimal FIR energy compaction filters. *IEEE Trans. Signal Processing*, 48(10):2822–2838, October 2000.

[27] W. Rudin. *Fourier analysis on groups*. Interscience, New York and London, 1962.

[28] K.J. Astrom and Bjorn Wittenmark. *Computer-controlled Systems — Theory and Design*. Prentice Hall, 1997.

[29] L.R. Rabiner. Techniques for designing finite-duration impulse-response digital filters. *IEEE Trans. Communication Technology*, 19(2):188–195, 1971.

[30] X. Zhang and T. Yoshikawa. Design of FIR Nyquist filters with low group delay. *IEEE Trans. Signal Processing*, 47(5):1454–1458, 1999.

[31] D. G. Luenberger. *Optimization by Vector Space Methods*. John Wiley, 1968.

[32] G. Meinsma, Y. Shrivastava, and M. Fu. A dual formulation of mixed μ and on the losslessness of (D, G) scaling. *IEEE Trans. Auto. Contr.*, 42(7):1032–1036, 1997.

[33] V. Balakrishnan and L. Vandenberghe. Semidefinite programming duality and linear time-invariant systems. *IEEE Trans. Auto. Contr.*, 48(1):30–41, 2003.