Positivity and linear matrix inequalities

Y. Genin, Y. Hachez, Yu. Nesterov, R. Ştefan, P. Van Dooren*, S. Xu
CESAME – Université catholique de Louvain
Av. G.Lemaître, 4, B-1348 Louvain-la-Neuve, Belgium
{genin,hachez,nesterov,stefan,vdooren,xu}@csam.ucl.ac.be

Revision 1 – November 22, 2001

Keywords: positivity, linear matrix inequalities, convex optimization.

Abstract

In this paper we first recall the general theory of Popov realizations of parahermitian transfer functions in the context of generalized state space systems. We then use this general framework to derive linear matrix inequalities for some particular applications in systems and control. Finally, we indicate how these problems can be solved numerically and what specific numerical difficulties can be encountered in these applications.

1 Introduction

Parahermitian transfer functions play a fundamental role in systems and control theory : they represent e.g. spectral density functions of stochastic processes, show up in spectral factorizations, and are also related to the algebraic Riccati equation. If these transfer functions are positive, they also form a convex set. This property has lead in systems and control theory to the extensive use of convex optimization techniques in this area, especially for so-called linear matrix inequalities [5]. We treat the case of positive transfer functions defined on the unit circle (i.e. the discrete-time case) as well as on the imaginary axis (i.e. the continuous-time case) since these are the most relevant to the area of systems and control.

We start by briefly recalling in Section 2 the basic theory of parahermitian transfer functions. The novelty of this section lies in the treatment of the most general rational transfer functions for which we propose realizations that reflect the parahermitian structure of the transfer function. The Kalman-Yakubovich-Popov (KYP) Lemma is re-derived in this general framework, and we also make the link to the linear matrix inequalities (LMI) which constitute the main thread of this paper. The KYP Lemma is also presented as an algebraic tool for problems of *spectral factorization* of rational matrices.

The rest of the paper contains two parts. Section 3 surveys some recent formulations of LMI solutions to some specific problems in systems and control. We treat e.g. the particular problems of positive real systems (including dissipativity), of non-negative matrix polynomials (for which we derive new necessary and sufficient conditions), of integral quadratic constraints and of model

*corresponding author.

reduction (possibly with passivity constraints). These problems either make explicit use of the generalized state-space formulation of LMI or illustrate some of their particular algorithmic aspects such as complexity and feasibility. Sections 4 and 5 then survey the basic numerical tools for solving these LMI problems. The basic approach is the interior-point method for conic problems for which the basic theory is developed in some detail. We also discuss issues of feasibility, complexity and speed of convergence. The last section shows that level set and cutting plane methods can efficiently solve a particular subset of these problems.

We end by proving in appendix some of the basic ideas underlying this paper, which can not be found in this generality in the literature.

2 Positive parahermitian matrix functions

A matrix function $\Phi : \mathbb{C} \to \mathbb{C}^{n\times n}$ maps the complex variable λ to a complex matrix $\Phi(\lambda)$. In this paper we restrict ourselves to *rational* matrix functions, i.e. where the elements of $\Phi(\lambda)$ are rational functions of λ. We introduce the *paraconjugate transpose* of this function with respect to a particular curve Γ in the complex plane. For convenience we use a different symbol for the variable λ for each choice of Γ. These three case are (a) the real axis (including the point at infinity), (b) the imaginary axis (relevant to continuous-time dynamical systems) and (c) the unit circle (relevant to discrete-time dynamical systems) :

$$\Phi_{\mathbb{R}}(x) \doteq \Phi^*(x) \qquad \text{for } \Gamma = \mathbb{R}, \tag{2.1a}$$

$$\Phi_{\mathbb{C}}(s) \doteq \Phi^*(-s) \qquad \text{for } \Gamma = j\,\mathbb{R}, \tag{2.1b}$$

$$\Phi_{\mathbb{D}}(z) \doteq \Phi^*(1/z) \qquad \text{for } \Gamma = e^{j\,\mathbb{R}}, \tag{2.1c}$$

where $\Phi^*(.)$ is just the conjugate transpose $\Phi(.)$. In this paper we will see that these three particular cases can be treated in much the same way. For this reason, we use the generic notation $\Phi_*(.)$ for $\Phi_{\mathbb{R}}(.)$, $\Phi_{\mathbb{C}}(.)$ and $\Phi_{\mathbb{D}}(.)$ and λ for the variables x, s and z.

If $\Phi(\lambda)$ is rational and analytic in λ, so is its paraconjugate transpose as it can be easily verified. We now say that $\Phi(\lambda)$ is a parahermitian matrix function, if it is its own paraconjugate transpose, i.e. if

$$\Phi_*(\lambda) = \Phi(\lambda) \tag{2.2}$$

for the three cases (2.1a, 2.1b, 2.1c). It is a well known fact that there exist first order conformable transformations of the complex plane that link the above cases with each other. Since substituting the real axis for the imaginary axis is a trivial operation ($jx = s$), only the cases (2.1b)and (2.1c) will be emphasized in the sequel, in view of the importance of their role in continuous-time and discrete-time control systems, respectively. Also, results specific to the above two situations will be identified by using the so-called Laplace operator ($\lambda = s$) for continuous-time systems and the delay operator ($\lambda = z$) for discrete-time systems.

2.1 Generalized state-space realization

Rational transfer matrices can always be represented as simple expressions involving first order polynomial matrix functions (i.e. pencils). Indeed, every rational transfer matrix $G(\lambda)$ of dimension $p \times m$ is known from realization theory to admit a generalized state-space model [54] of the form

$$G(\lambda) = (C - \lambda\, F)(\lambda\, E - A)^{-1}\, B + D, \tag{2.3}$$

which is the Schur complement of the so-called system matrix $S(\lambda)$ of dimension $(n+p)\times(n+m)$

$$S(\lambda) = \begin{bmatrix} A-\lambda E & B \\ C-\lambda F & D \end{bmatrix} \tag{2.4}$$

with respect to its top left block entry. The minimum dimension n of the invertible pencil $(A-\lambda E)$ is the McMillan degree of $G(\lambda)$ [54], and the generalized eigenvalues of $A-\lambda E$ are then the poles of $G(\lambda)$ [54, 29]. A test for the minimality of the realization $S(\lambda)$ is the following set of conditions [43]:

$$\begin{array}{ll} (i) & \text{rank}\begin{bmatrix} A-\lambda E & B \end{bmatrix} = n, \quad \forall\,\lambda\in\mathbb{C},\ |\lambda|<\infty; \\ (ii) & \text{rank}\begin{bmatrix} E & B \end{bmatrix} = n; \\ (iii) & \text{rank}\begin{bmatrix} A-\lambda E \\ C-\lambda F \end{bmatrix} = n, \quad \forall\,\lambda\in\mathbb{C},\ |\lambda|<\infty; \\ (iv) & \text{rank}\begin{bmatrix} E \\ F \end{bmatrix} = n. \end{array} \tag{2.5}$$

If these conditions are not all satisfied, then the system matrix (2.4) is not minimal and the state space dimension can always be reduced so as to achieve minimality [54]. Minimal realizations are not unique, even though their dimension n is. A simple class of transformations acting on (2.4) and preserving minimality is defined by

$$\hat{S}(\lambda) = \begin{bmatrix} \hat{A}-\lambda\hat{E} & \hat{B} \\ \hat{C}-\lambda\hat{F} & \hat{D} \end{bmatrix} = \begin{bmatrix} P & 0 \\ Y & I_p \end{bmatrix} \begin{bmatrix} A-\lambda E & B \\ C-\lambda F & D \end{bmatrix} \begin{bmatrix} Q & X \\ 0 & I_m \end{bmatrix}, \tag{2.6}$$

with P, Q invertible matrices and where X is chosen so as to have $E\,X = F\,X = 0$. Indeed, the Schur complements of $S(\lambda)$ and $\hat{S}(\lambda)$ are easily checked to be identical so that they are both well defined realizations of $G(\lambda)$. The minimality conditions (2.5) for $\hat{S}(\lambda)$ are easily seen to hold as well since left and right invertible transformation of the relevant pencils do not change their rank. Let us finally point out that there is no unique definition of minimality for generalized state-space systems.

The zeros of $G(\lambda)$ can also be computed as generalized eigenvalues of a smaller pencil, derived from $S(\lambda)$. Let M be any invertible row transformation such that

$$\begin{bmatrix} \tilde{E} \\ 0 \end{bmatrix} \doteq M \begin{bmatrix} E \\ F \end{bmatrix}, \tag{2.7}$$

where $\tilde{E}$ is $n\times n$ (one can choose M to be unitary to achieve this). Applying this transformation to $S(\lambda)$ defines the matrices $\tilde{A},\tilde{B},\tilde{C}$ and $\tilde{D}$ as follows

$$\tilde{S}(\lambda) \doteq M\,S(\lambda) = \begin{bmatrix} \tilde{A}-\lambda\tilde{E} & \tilde{B} \\ \tilde{C} & \tilde{D} \end{bmatrix}. \tag{2.8}$$

If $\tilde{D}$ is invertible, it follows from [29] that the Schur complement

$$(\tilde{A}-\lambda\tilde{E}) - \tilde{B}\,\tilde{D}^{-1}\tilde{C} \tag{2.9}$$

is a so-called zero pencil of the system : its generalized eigenvalues are the zeros of $G(\lambda)$. If $\tilde{D}$ is not invertible, it is shown in [29] that one can still derive a zero pencil but we will not further elaborate on this here.

2.2 Parahermitian realizations

It is natural to expect that parahermitian transfer matrices admit realizations (and hence system matrices), that reflect this property in some structural way. In fact, one will consider in this paper generalized state space realizations of Popov type

$$S(s) = \left[\begin{array}{ccc} 0 & A^* + s\,E^* & C^* + s\,F^* \\ A - s\,E & H_{11} & H_{12} \\ C - s\,F & H_{21} & H_{22} \end{array}\right] \qquad \text{for } \Gamma = j\,\mathbb{R}, \tag{2.10a}$$

$$S(z) = \left[\begin{array}{ccc} 0 & z\,A^* - E^* & z\,C^* - F^* \\ A - z\,E & H_{11} & H_{12} \\ C - z\,F & H_{21} & H_{22} \end{array}\right] \qquad \text{for } \Gamma = e^{j\,\mathbb{R}}, \tag{2.10b}$$

where

$$H = \left[\begin{array}{cc} H_{11} & H_{12} \\ H_{21} & H_{22} \end{array}\right] \in \mathbb{C}^{(n+p)\times(n+p)}$$

is hermitian. Setting the matrix $T(\lambda) = (C - \lambda F)(\lambda E - A)^{-1}$, one finds after some algebraic manipulations that the transfer function corresponding to these system matrices is given by

$$\Phi(\lambda) = \left[\begin{array}{cc} T(\lambda) & I \end{array}\right] \left[\begin{array}{cc} H_{11} & H_{12} \\ H_{21} & H_{22} \end{array}\right] \left[\begin{array}{c} T_*(\lambda) \\ I \end{array}\right] \tag{2.11}$$

which is clearly parahermitian. Note that the above realizations are not necessarily minimal. If it is the case, then the poles of the transfer function are the eigenvalues of, respectively,

$$\left[\begin{array}{cc} 0 & A^* + s\,E^* \\ A - s\,E & H_{11} \end{array}\right] \quad \text{and} \quad \left[\begin{array}{cc} 0 & z\,A^* - E^* \\ A - z\,E & H_{11} \end{array}\right], \tag{2.12}$$

which are symmetric with respect to Γ (i.e. $s_i, -\bar{s}_i$ or $z_i, 1/\bar{z}_i$). The zeros of the transfer function are those of the respective system matrices, which clearly exhibit the same form of symmetry. If we define a transformation matrix M as in (2.7) and then define accordingly :

$$\left[\begin{array}{c} \tilde{A} \\ \tilde{C} \end{array}\right] \doteq M \left[\begin{array}{c} A \\ C \end{array}\right], \quad \left[\begin{array}{cc} \tilde{H}_{11} & \tilde{H}_{12} \\ \tilde{H}_{21} & \tilde{H}_{22} \end{array}\right] \doteq M \left[\begin{array}{cc} H_{11} & H_{12} \\ H_{21} & H_{22} \end{array}\right] M^*, \tag{2.13}$$

this yields a new system matrix $\tilde{S}(\lambda)$ which is similar to $S(\lambda)$ but with $\tilde{F} = 0$, and which has the same zeros. If now $\tilde{H}_{22}$ is invertible, then the zeros are also the generalized eigenvalues of the respective Schur complements of the system matrices :

$$\left[\begin{array}{cc} 0 & \tilde{A}^* + s\,\tilde{E}^* \\ \tilde{A} - s\,\tilde{E} & \tilde{H}_{11} \end{array}\right] - \left[\begin{array}{c} \tilde{C}^* \\ \tilde{H}_{12} \end{array}\right] \tilde{H}_{22}^{-1} \left[\begin{array}{cc} \tilde{C} & \tilde{H}_{21} \end{array}\right], \tag{2.14a}$$

$$\left[\begin{array}{cc} 0 & z\,\tilde{A}^* - \tilde{E}^* \\ \tilde{A} - z\,\tilde{E} & \tilde{H}_{11} \end{array}\right] - \left[\begin{array}{c} z\,\tilde{C}^* \\ \tilde{H}_{12} \end{array}\right] \tilde{H}_{22}^{-1} \left[\begin{array}{cc} \tilde{C} & \tilde{H}_{21} \end{array}\right], \tag{2.14b}$$

which are known in the literature as the Hamiltonian and the symplectic pencils describing the zeros of the respective parahermitian transfer functions. This could suggest that parahermitian transfer functions would be always of even degree. This is actually not the case, as exemplified by the first degree scalar parahermitian transfer function $G(s) = 2j/s$, which has a pole at $s = 0$ and a zero at $s = \infty$; note that odd degree parahermitian transfer functions must have at least one pole and zero on the curve Γ. However, if the transfer function, as considered in the sequel, is not only parahermitian but also non-negative definite on the contour Γ, then it has always a minimal realization of type (2.10); this is easily established from the well known spectral factorization property $\Phi(\lambda) = G(\lambda)\,G_*(\lambda)$ of any such transfer function [64]. We will also show that nonminimal realizations of the above type always exist.

2.3 Linear matrix forms and the Kalman-Yakubovich-Popov Lemma

Realizations of the type (2.10) are obviously not unique, since they are invariant under transformations of the form (2.6). We now introduce the following transformations which leave (E, A, C, F) invariant :

$$\begin{bmatrix} I & 0 & 0 \\ E\,X & I & 0 \\ F\,X & 0 & I \end{bmatrix} S(s) \begin{bmatrix} I & X\,E^* & X\,F^* \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} = \begin{bmatrix} 0 & A^* + s\,E^* & C^* + s\,F^* \\ A - s\,E & H_{11}(X) & H_{12}(X) \\ C - s\,F & H_{21}(X) & H_{22}(X) \end{bmatrix}, \quad (2.15a)$$

$$\begin{bmatrix} I & 0 & 0 \\ E\,X & I & 0 \\ F\,X & 0 & I \end{bmatrix} S(z) \begin{bmatrix} I & X\,A^* & X\,C^* \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} = \begin{bmatrix} 0 & z\,A^* - E^* & z\,C^* - F^* \\ A - z\,E & H_{11}(X) & H_{12}(X) \\ C - z\,F & H_{21}(X) & H_{22}(X) \end{bmatrix}, \quad (2.15b)$$

where the respective matrices $H(X)$ are given by

$$H(X) \doteq \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} + \begin{bmatrix} E \\ F \end{bmatrix} X \begin{bmatrix} A^* & C^* \end{bmatrix} + \begin{bmatrix} A \\ C \end{bmatrix} X \begin{bmatrix} E^* & F^* \end{bmatrix}, \quad (2.16a)$$

$$H(X) \doteq \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} + \begin{bmatrix} A \\ C \end{bmatrix} X \begin{bmatrix} A^* & C^* \end{bmatrix} - \begin{bmatrix} E \\ F \end{bmatrix} X \begin{bmatrix} E^* & F^* \end{bmatrix}. \quad (2.16b)$$

Let us emphasize that the Schur complement (i.e. the transfer function) of these realizations does not change under these transformations.

It turns out that the celebrated Kalman-Yakubovich-Popov Lemma [31, 62, 48] allows one to express the non-negative definiteness of $\Phi(\lambda)$ on the curve Γ in terms of conditions on $H(X)$.

Theorem 2.1 (KYP Lemma). *Every parahermitian transfer function $\Phi(\lambda)$ has a realization as in (2.10), with condition (2.5) satisfied. Moreover, it is non-negative definite on Γ, i.e.*

$$\Phi(\lambda) \succeq 0, \qquad \textit{for almost all } \lambda \in \Gamma, \quad (2.17)$$

if and only if there exists a hermitian matrix X such that

$$H(X) \succeq 0. \quad (2.18)$$

Proof. The basic ideas of the proof are only given. Every parahermitian transfer function admits an additive decomposition

$$\Phi(\lambda) = G(\lambda) + G_*(\lambda)$$

(take e.g. $G(\lambda) = \Phi(\lambda)/2$). Using a realization $G(\lambda) = (C - \lambda F)(\lambda E - A)^{-1} B + D$ which is minimal in the sense of (2.5), we clearly have the identity

$$\Phi(\lambda) = \begin{bmatrix} T(\lambda) & I \end{bmatrix} \begin{bmatrix} 0 & B \\ B^* & D + D^* \end{bmatrix} \begin{bmatrix} T_*(\lambda) \\ I \end{bmatrix} \quad (2.19)$$

with $T(\lambda) = (C - \lambda F)(\lambda E - A)^{-1}$, which is a realization of the type (2.10) because of (2.11). The corresponding realization is then given by (2.10), where one has

$$H = \begin{bmatrix} 0 & B \\ B^* & D + D^* \end{bmatrix}. \quad (2.20)$$

In appendix we show that with condition (2.5) satisfied for the subsystem $T(\lambda)$, the transfer matrix given in (2.19) satisfies $\Phi(\lambda) \succeq 0$ for almost all $\lambda \in \Gamma$ if and only if there exists a hermitian matrix X such that $H(X) \succeq 0$. □

The above result is linked to that of *spectral factorization*. If $H(X) \succeq 0$ it can be factorized as

$$H(X) = \begin{bmatrix} L \\ W \end{bmatrix} \begin{bmatrix} L^* & W^* \end{bmatrix}, \tag{2.21}$$

so that $\hat{S}(\lambda)$ is easily seen to be the system matrix of

$$\Phi(\lambda) = \Xi(\lambda)\, \Xi_*(\lambda) \tag{2.22}$$

with the so-called spectral factor $\Xi(\lambda)$ defined by

$$\Xi(\lambda) = (C - \lambda F)\,(\lambda E - A)^{-1}\, L + W. \tag{2.23}$$

For more details on spectral factorizations of rational matrices, see [43, 64].

Closely related to the above result is the so-called *Positive Real Lemma*; see, for instance, Theorem 3.2 in Section 3.1 and the references herein.

Another related result is the *Bounded Real Lemma*. A transfer function $G(\lambda)$, which is bounded almost everywhere on Γ, has $L_\infty-$norm lower than $\gamma > 0$ if and only if it satisfies the constraint

$$G(\lambda)\, G_*(\lambda) \preceq \gamma^2\, I, \qquad \text{for almost all } \lambda \in \Gamma, \tag{2.24}$$

which can be recast in term of the non-negativity condition

$$\gamma^2\, I - G(\lambda)\, G_*(\lambda) \succeq 0, \qquad \text{for almost all } \lambda \in \Gamma. \tag{2.25}$$

If $G(\lambda) = (C - \lambda F)\,(\lambda E - A)^{-1}\, B + D$, then a realization is trivially obtained via

$$\gamma^2\, I - G(\lambda)\, G_*(\lambda) = \begin{bmatrix} T(\lambda) & I \end{bmatrix} H \begin{bmatrix} T_*(\lambda) \\ I \end{bmatrix} \tag{2.26}$$

with T introduced before and H defined as

$$H = \begin{bmatrix} 0 & 0 \\ 0 & \gamma^2 I \end{bmatrix} - \begin{bmatrix} B \\ D \end{bmatrix} [B^* \; D^*],$$

which again is of the type (2.10).

Remark 2.2.

1. We have not assumed anywhere that $G(\lambda)$ is stable (i.e. that the generalized eigenvalues of $\lambda E - A$ are in the stable region of the complex plane, hence $G(\lambda)$ is analytic in $\mathrm{Re}\,\lambda > 0$). If $\Phi(\lambda)$ has no poles on Γ then $G(\lambda)$ in the above theorem can be chosen stable. This implies e.g. that the realization for $\Phi(\lambda)$ is minimal and that the solution X for the KYP Lemma is negative definite.

2. As pointed out in the proof of Theorem 2.1, every parahermitian transfer function can be realized as in (2.10) when minimality is given up. The first order transfer function $\Phi(s) = 2j/s$ has e.g. a realization (2.19) with $sE - A = s$, $B = j$, $C = 1$ and $D = F = 0$.

3. Alternatively, one can consider realizations of the form

$$\tilde{G}(\lambda) = \tilde{C}(\lambda\tilde{E} - \tilde{A})^{-1}(\tilde{B} - \lambda\tilde{K}) + \tilde{D}, \tag{2.27}$$

   which are "dual" to the generalized state-space representation in (2.3), in the sense that $\tilde{G}_*(\lambda)$ has precisely a realization of the form (2.3). Then, the state-space representation (2.11) or the expressions of the matrices $H(X)$ showing up in (2.16) can be rewritten accordingly.

2.4 Linear and Riccati matrix inequalities

As shown before, a transfer function $\Phi(\lambda)$, non-negative definite on Γ, necessarily implies the existence of a Hermitian matrix X such that $H(X) \succeq 0$. Since the entries of this matrix inequalities are linear in the elements of the unknown matrix X, they are called linear matrix inequalities.

Let us now further assume that $H_{22}(X)$ is positive definite. Then, the Schur complement of $H(X)$ with respect to $H_{22}(X)$ must be non-negative definite. It is easy to check that this amounts to the constraints

$$\begin{aligned} &H_{11} + E\,X\,A^* + A\,X\,E^* - (H_{12} + E\,X\,C^* + A\,X\,F^*) \\ &\quad (H_{22} + F\,X\,C^* + C\,X\,F^*)^{-1}\,(H_{21} + F\,X\,A^* + C\,X\,E^*) \succeq 0, \end{aligned} \tag{2.28a}$$

$$\begin{aligned} &H_{11} + A\,X\,A^* - E\,X\,E^* - (H_{12} + A\,X\,C^* - E\,X\,F^*) \\ &\quad (H_{22} + C\,X\,C^* - F\,X\,F^*)^{-1}\,(H_{21} + C\,X\,A^* - F\,X\,E^*) \succeq 0, \end{aligned} \tag{2.28b}$$

in the continuous-time and discrete-time cases respectively. These are the so-called Riccati matrix inequalities introduced in [58].

For appropriate choices of X, one has that rank $H_{22}(X) =$ rank $H(X)$ so that its Schur complement in $H(X)$ must be zero. The above inequalities then become equalities, which have the form of the celebrated Riccati equations for which X appears therefore to be a solution. These equations are rewritten below for the case that $F = 0$ since this can always be obtained under a transformation (2.7)–(2.13) of the system :

$$H_{11} + E\,X\,A^* + A\,X\,E^* - (H_{12} + E\,X\,C^*)(H_{22})^{-1}\,(H_{21} + C\,X\,E^*) = 0, \tag{2.29a}$$

$$H_{11} + A\,X\,A^* - E\,X\,E^* - (H_{12} + A\,X\,C^*)(H_{22} + C\,X\,C^*)^{-1}\,(H_{21} + C\,X\,A^*) = 0. \tag{2.29b}$$

The solution X of these equations is obtained from the calculation of eigenspaces of the zero pencils (2.14) [35]. If there are no repeated eigenvalues in these pencils (this is the generic case) then the number of solutions X to these equations are finite, whereas the inequalities have typically an infinite solution set, as shown in a later section.

3 Selected applications

3.1 Positive Realness and Dissipativity

In the study and control of general dynamical systems the energy stored in the system is an important information. This idea, borrowed from thermodynamics, has been clearly established in the influential paper of Willems [59], where *dissipativity* has been introduced as a basic concept in system theory.

Energy related concepts like passivity and (later) dissipativity have been widely used as control tools, since the breakthrough by Popov in the early 1960's. Until the work of Popov [48], passivity was a basic network theory concept, dealing with rational transfer functions that are *positive real* and can be realized with passive elements of a circuit (RLC networks). Popov's main contribution consisted in introducing passivity as one of the fundamental feedback properties. He developed a comprehensive theory of so-called *hyperstability*. A typical situation is that of a controlled dynamical system and an associated quadratic index

$$\dot{x} = Ax + Bu, \;\; \eta(0,T) = x^*(t)Jx(t) + \int_0^T \begin{bmatrix} x^*(t) & u^*(t) \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} \mathrm{d}t, \tag{3.1}$$

which is called *hyperstable* if there exist $s_0 > 0$, $s_1 > 0$ such that

$$s_1\|x(T)\|^2 \leq \eta(0,T) + s_0\|x(0)\|^2, \ \ \forall T > 0 \tag{3.2}$$

along the solution of the differential equation. But (3.2) is precisely the *dissipation inequality* (see [60]) satisfied by the linear dynamical system $\dot{x} = Ax + Bu$, with the quadratic supply rate $w(x,u) = \eta(0,T)$.

Under certain assumptions ($B \neq 0$, "minimal stability"), Popov obtains necessary and sufficient frequency-domain conditions for hyperstability, showing that the system (3.1) is hyperstable if and only if

$$\Pi(j\omega) = \left[\begin{array}{cc} B^*(-j\omega I - A^*)^{-1} & I \end{array}\right] \left[\begin{array}{cc} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{array}\right] \left[\begin{array}{c} (j\omega I - A)^{-1}B \\ I \end{array}\right] \geq 0,$$

for every $\omega \in \mathbb{R}$, $\det(j\omega I - A) \neq 0$. In other words, hyperstability is equivalent to the non-negativeness (on $j\mathbb{R}$) of a particular parahermitian transfer function (also called Popov function). In [60], Willems then gives an equivalent condition for dissipativity, expressed by the *positive realness* of the function $\Pi(s)$, thus linking the concepts of dissipativity and hyperstability to that of *positivity* (see [49]). Passivity can be regarded as a special case of dissipativity for the system $\dot{x} = Ax + Bu$, $y = Cx + Du$, with quadratic supply rate

$$w(x,u) = \left[\begin{array}{cc} x^* & u^* \end{array}\right] \left[\begin{array}{cc} 0 & \frac{1}{2}C \\ \frac{1}{2}C^* & D + D^* \end{array}\right] \left[\begin{array}{c} x \\ u \end{array}\right]. \tag{3.3}$$

For more details on this relationship, see [28]. The KYP Lemma now expresses that a system is dissipative if and only if an LMI of the form (2.18) is satisfied by a (negative definite) hermitian matrix X. It is well-known that the feedback interconnection of two dissipative (or passive) systems is dissipative (or passive) as well. Actually, LMI show up in most stability analysis problems involving these concepts. Applications of dissipativity (passivity) can be found in many areas: for instance, in adaptive control (see [6, 7]), in robot control [2] or in stability analysis and robust stabilization (see [24, 50, 57] and in the references therein).

Another important problem is the positive real synthesis for linear time-invariant systems, which has been considered in [53]. Its objective is the design of controllers such that the resulting closed-loop system is stable and the closed-loop transfer function is positive real. As shown in [53], a solution to this problem involves solving a pair of Riccati inequalities. These results have been extended to linear systems with time-invariant uncertainty in [52, 61]. Moreover, observer-based stabilizing controllers which achieve positive realness have been designed in [36].

Some positive realness results have also been extended to generalized state-space systems in [55] by using generalized Lyapunov theory, while [65] obtained similar results for both continuous and discrete generalized state-space systems in terms of LMI. Consider a transfer matrix given in generalized state-space form with E nonsingular :

$$G(\lambda) = C\left(\lambda E - A\right)^{-1} B + D. \tag{3.4}$$

Definition 3.1. [25, 26, 55]
a) For $\lambda = s$, the rational matrix function (3.4) is said to be *positive real* if $G(s)$ is analytic in $\mathrm{Re}(s) > 0$ and satisfies $G(s) + G(s)^* \geq 0$ for $\mathrm{Re}(s) > 0$. It is said to be *strongly positive real* if $G(s)$ is analytic in $\mathrm{Re}(s) \geq 0$ and satisfies $G(j\omega) + G(j\omega)^* > 0$ for $\omega \in [0, \infty]$.
b) For $\lambda = z$, the rational matrix function (3.4) is said to be *positive real* if $G(z)$ is analytic in $|z| > 1$ and satisfies $G(z) + G(z)^* \geq 0$ for $|z| > 1$. It is said to be *strongly positive real* if $G(z)$ is analytic in $|z| > 1$ and satisfies $G(e^{j\theta}) + G(e^{j\theta})^* > 0$ for $\theta \in [0, 2\pi]$.

LMI conditions for positive realness are now given in the following theorem [25, 53, 55].

Theorem 3.2. *The transfer function* $G(\lambda) = C\left(\lambda E - A\right)^{-1} B + D$ *is stable and strongly positive real if and only if there exists a* positive definite *matrix* X *such that*

$$\left[\begin{array}{cc} EXA^* + AXE^* & B - EXC^* \\ B^* - CXE^* & -(D + D^*) \end{array}\right] \prec 0 \tag{3.5a}$$

in the continuous-time case ($\lambda = s$)*, and*

$$\left[\begin{array}{cc} AXA^* - EXE^* & B - AXC^* \\ B^* - CXA^* & -\left(D + D^* - CXC^*\right) \end{array}\right] \prec 0 \tag{3.5b}$$

in the discrete-time case ($\lambda = z$)*.*

Remark 3.3. If $\operatorname{rank} E < n$, the results in Theorem 3.2 are extended to index 1 systems in [65]. Necessary and sufficient conditions for singular systems to be regular, impulse-free, stable as well as strongly positive real are obtained in terms of LMI.

3.2 Non-negative matrix polynomials

Positive parahermitian transfer functions obviously form a convex set and were recently studied in the convex optimization literature [5, 41]. The parameterization of pseudo-polynomial matrices proposed in [17] fits naturally into the context of this paper. In particular, it can be obtained as a straightforward application of the celebrated KYP Lemma to an appropriate subset of positive parahermitian transfer functions. Moreover, this particular problem leads to an algorithm of low complexity, as shown in Section 4.4.

Let us start with the case of non-negative transfer functions on the unit circle. It follows from its finite expansion and from its parahermitian character that the corresponding pseudo-polynomial matrix

$$P(z) = \sum_{i=-n}^{n} P_i z^i \succeq 0, \tag{3.6}$$

has $p \times p$ coefficient matrices that satisfy $P_{-i} = P_i^*$. Consider then the set of hermitian matrices

$$Y = \left[\begin{array}{cccc} Y_{0,0} & Y_{0,1} & \cdots & Y_{0,n} \\ Y_{1,0} & Y_{1,1} & \cdots & Y_{2,n} \\ \vdots & \vdots & & \vdots \\ Y_{n,0} & Y_{n,1} & \cdots & Y_{n,n} \end{array}\right],$$

with blocks of dimension $p \times p$ and the block shift operator

$$Z = \left[\begin{array}{ccc} 0 & I_p & \\ & 0 & \ddots \\ & & \ddots \; I_p \\ & & 0 \end{array}\right].$$

If $\Pi(z)$ stands for $\Pi(z) = \left[\begin{array}{cccc} I_p & zI_p & \cdots & z^n I_p \end{array}\right]^T$, the identity

$$\Pi_*(z) Y \Pi(z) = P(z) \tag{3.7}$$

is algebraically equivalent to the equation

$$P_i = \sum_{k-l=i} Y_{k,l}, \tag{3.8}$$

assuming $Y_{k,l} = 0$ for k and l outside their definition range. Clearly, the choice

$$Y_0 = \begin{bmatrix} P_0 & P_1 & \dots & P_n \\ P_1^* & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ P_n^* & 0 & \dots & 0 \end{bmatrix}, \tag{3.9}$$

is an admissible matrix Y. Then the following characterization theorem can be stated [17].

Theorem 3.4. *A hermitian matrix Y satisfies equation (3.7) if and only if it can be expressed as*

$$Y = Y_0 + X - Z^T X Z, \tag{3.10}$$

where X is hermitian, i.e. $X = X^*$, and has the form

$$\left[\begin{array}{c|c} X_0 & 0 \\ \hline 0 & 0 \end{array}\right], \quad X_0 \in \mathbb{C}^{np \times np}. \tag{3.11}$$

The set of non-negative pseudo-polynomial matrices on the unit circle can then be characterized as follows [17].

Theorem 3.5. *A pseudo-polynomial matrix $P(z) = \sum_{i=-n}^{n} P_i z^i$ is non-negative definite on the unit circle if and only if there exists a non-negative definite hermitian matrix Y with blocks $Y_{k,l}, k, l = 0, \dots, n$ such that (assuming $Y_{k,l} = 0$ for k and l outside their definition range) :*

$$P_i = \sum_{k-l=i} Y_{k,l}, \qquad \text{for } i = -n, \dots, 0, \dots, n. \tag{3.12}$$

This characterization of pseudo-polynomials non-negative on the unit circle also extends a result previously obtained by Nesterov [41] for trigonometric polynomials. From a practical viewpoint, it allows us to efficiently solve various filter design problems using semidefinite programming, see Section 4 and [1, 10, 16]. It can alternatively be obtained from the theory of positive parahermitian transfer functions. More precisely, it follows from a straightforward application of the KYP Lemma to the subclass of positive parahermitian transfer functions that have a pseudo-polynomial form. Let us derive this simple proof.

Consider a generalized state space realization of a parahermitian transfer function of the form

$$\Phi(z) = \begin{bmatrix} zF(I - zE)^{-1} & I \end{bmatrix} Y_0 \begin{bmatrix} (zI - E^*)^{-1} F^* \\ I \end{bmatrix} \tag{3.13}$$

with Y_0 some hermitian matrix. The transfer function $\Phi(z)$ is non-negative on the unit circle if and only if the matrix $\Phi(e^{j\theta})$ is non-negative definite definite for all θ in the interval $[0, 2\pi]$. In this setting, the KYP Lemma states that $\Phi(z)$ will be a well defined non-negative parahermitian transfer function if and only if there exits a hermitian matrix $\tilde{X}$ such that

$$Y(\tilde{X}) = Y_0 + \begin{bmatrix} I \\ 0 \end{bmatrix} \tilde{X} \begin{bmatrix} I & 0 \end{bmatrix} - \begin{bmatrix} E \\ F \end{bmatrix} \tilde{X} \begin{bmatrix} E^* & F^* \end{bmatrix} \tag{3.14}$$

is non-negative definite. With Y_0 as in (3.9), $E = Z^*$ and $F = [0, \dots, 0, I_p]$, the following equality holds

$$\Phi(z) = \sum_{i=-n}^{+n} P_i z^i.$$

Therefore, the pseudo-polynomial matrix $P(z)$ is found to be non-negative definite on the unit circle if and only if there exists a hermitian matrix $\tilde{X}$ such that the matrix

$$Y(\tilde{X}) = Y_0 + \begin{bmatrix} \tilde{X} - Z^*\tilde{X}Z & -Z^*\tilde{X}F^* \\ -F\tilde{X}Z & -F\tilde{X}F^* \end{bmatrix}$$

is non-negative definite. This is exactly the characterization proposed in Theorems 3.4 and 3.5 provided one substitutes $\tilde{X}$ for X_0.

The same method can be applied to non-negative transfer functions on the real line or the imaginary axis. It leads to similar results for non-negative matrix polynomials on these curves of the complex plane [17].

Theorem 3.6. *A pseudo-polynomial matrix $P(x) = \sum_{i=0}^{2n} P_i x^i$ is non-negative definite on the real axis if and only if there exists a non-negative definite matrix Y with blocks $Y_{k,l}, k,l = 0,\ldots,n$ such that (assuming $Y_{k,l} = 0$ for k and l outside their definition range) :*

$$P_i = \sum_{k+l=i} Y_{k,l}, \qquad i = 0,\ldots,2n. \tag{3.15}$$

Theorem 3.7. *A pseudo-polynomial matrix $P(s) = \sum_{i=0}^{2n} P_i s^i$ is non-negative on the imaginary axis if and only if there exists a non-negative matrix Y with blocks $Y_{k,l}, k,l = 0,\ldots,n$ such that (assuming $Y_{k,l} = 0$ for k and l outside their definition range) :*

$$P_i = (-j)^i \sum_{k+l=i} Y_{k,l}, \qquad i = 0,\ldots,2n. \tag{3.16}$$

3.3 Integral quadratic constraints

The general concept of *Integral Quadratic Constraints* (IQC) has its origin in the work of Yakubovich [63], and has been used for analyzing system stability of special nonlinear problems. As pointed out recently in [51], IQC have been usually defined in time-domain. Observing that many developments in robust control can be reformulated in a simple IQC framework, Megretski and Rantzer [38] extended the concept to the case where IQC are given in frequency-domain. IQC are used to describe relationships between different components in a complex dynamical system, exploiting structural information about the uncertainty or characterizing properties of external signals.

Two signals $w \in L^{2,m}[0,\infty)$, $v \in L^{2,l}[0,\infty)$ are said to satisfy the IQC defined by Φ, if

$$\int_{-\infty}^{\infty} \begin{bmatrix} \hat{v}(j\omega) \\ \hat{w}(j\omega) \end{bmatrix}^* \Phi(j\omega) \begin{bmatrix} \hat{v}(j\omega) \\ \hat{w}(j\omega) \end{bmatrix} \mathrm{d}\omega \geq 0. \tag{3.17}$$

Here $\hat{w}$ and $\hat{v}$ are the Fourier transforms of the signals w and v, respectively; Φ is typically a parahermitian rational function that is bounded on the imaginary axis. It is also said that $\Delta : L^{2,l}[0,\infty) \to L^{2,m}[0,\infty)$ satisfies the IQC defined by Φ if, for every $v \in L^{2,l}[0,\infty)$, inequality (3.17) holds for $w = \Delta(v)$ and v.

Consider the typical interconnection used in the study of robust control problems (see Figure 1) of "absolute stability" type, which consists of a rational transfer function $G(s) = D + C(sI - A)^{-1}B$ and an uncertainty or perturbation Δ:

$$\begin{aligned} v &= Gw + f \\ w &= \Delta(v) + e \end{aligned}$$

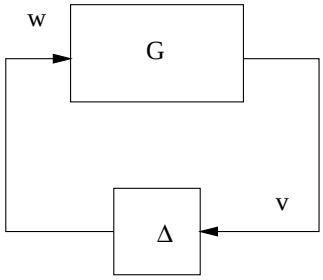

Figure 1: Standard configuration for robustness analysis

The stability theorem given below includes classical dissipativity (passivity) results as particular cases.

Theorem 3.8. *[38] Let $G(s) \in RH_\infty^{l\times m}$ and let Δ be a bounded causal operator. Assume that for every $\tau \in [0,1)$ the interconnection of G and $\tau\Delta$ is well-posed and that the IQC defined by Φ is satisfied by $\tau\Delta$. If there exists $\epsilon > 0$ such that*

$$\left[\begin{array}{c} G(j\omega) \\ I \end{array}\right]^* \Phi(j\omega) \left[\begin{array}{c} G(j\omega) \\ I \end{array}\right] \preceq -\epsilon I, \ \ \forall\, \omega \in \mathbb{R}, \tag{3.18}$$

then the feedback connection of G and Δ is stable.

By choosing appropriate (constant) values for $\Phi(j\omega)$, one can retrieve classical results, as the Small Gain Theorem ($\Phi = \left[\begin{array}{cc} I & 0 \\ 0 & -I \end{array}\right]$) or the Passivity Theorem ($\Phi = \left[\begin{array}{cc} 0 & I \\ I & 0 \end{array}\right]$). The role of the "multipliers" Φ (belonging to a given *set* of Φ's, say $\mathbf{\Phi}_\Delta$) is to summarize information about the uncertainty Δ. In practice, the set $\mathbf{\Phi}_\Delta$, which describes the IQC corresponding to Δ, is defined by a finite number of variables and can be written in the form

$$\Pi(j\omega) = \sum_{k=1}^{N} x_k \Pi_k(j\omega), \ \ x_k \in \mathbb{R}_+ .$$

It can be shown that there exist matrices A and B and symmetric matrices H_k, $1 \le k \le N$, of appropriate dimensions, such that for every $1 \le k \le N$

$$\left[\begin{array}{c} G(j\omega) \\ I \end{array}\right]^* \Pi_k(j\omega) \left[\begin{array}{c} G(j\omega) \\ I \end{array}\right] = \left[\begin{array}{c} (j\omega I - A)^{-1}B \\ I \end{array}\right]^* H_k \left[\begin{array}{c} (j\omega I - A)^{-1}B \\ I \end{array}\right].$$

The frequency domain inequality (3.18) can be eventually rewritten as

$$\left[\begin{array}{c} (j\omega I - A)^{-1}B \\ I \end{array}\right]^* H \left[\begin{array}{c} (j\omega I - A)^{-1}B \\ I \end{array}\right] \preceq -\epsilon I, \ \ \forall\, \omega \in \mathbb{R}, \ \ H = \sum_{k=1}^{N} x_k H_k.$$

By applying now the KYP Lemma it turns out that (3.18) is equivalent to a system of linear matrix inequalities in the variables X and x_k, $1 \le k \le N$. Note that the above frequency domain inequality has been written in its dual form (see point 3. in Remark 2.2).

Large-scale methods for the above problems are usually based on cutting planes methods, which do not require auxiliary variables : they directly work in the frequency domain[46, 32]. These schemes are discussed in Section 5.2. Moreover, they can be extended to transfer functions $G(\lambda)$ given in generalized state-space form.

3.4 Model Reduction

Problems arising in applications give the motivation for studying issues of approximation and model reduction. Some of these problems are: systems described by PDEs, systems arising in circuit simulation, weather prediction, components of mechanical structures (e.g. the space station). Model reduction aims at replacing a system of differential or difference equations of high complexity by one of much lower complexity. At the same time, one tries to preserve certain critical properties of the system (e.g. stability or passivity) and approximate important features (e.g. the system response) appropriately. During the last two decades, a lot of progress has been made in the theory of this approximation problem.

Roughly speaking, the model reduction problem consists of finding a rational matrix function of McMillan degree r, say G_r, such that, for a given transfer function G of McMillan degree n, $G(\lambda)$ and $G_r(\lambda)$ are close in some sense and $r \ll n$. Among the most popular model reduction techniques are the balanced truncation [39, 47], the optimal Hankel norm model reduction [19], the coprime factorization reduction [37] and those based on Krylov methods [23]. Some of these techniques provide guaranteed L_∞ error bounds, expressed in terms of the sum of the "neglected" Hankel singular values of the system.

A problem of great interest has been the lower order approximation in the H_∞ norm. A first characterization of the solutions to the H_∞ model reduction problem has been derived in [34], by converting it into a Hankel norm approximation problem through an appropriate embedding scheme. Numerical issues concerning this embedding step seem to be still open. An alternative approach for *suboptimal* H_∞ model reduction has been proposed by Grigoriadis [22] in terms of an LMI system and a coupling non-convex rank constraint set. Further, the main result in [22] is reproduced for continuous-time systems with complex coefficients. The discrete-time counterpart is similar, and can be found in the same paper.

H_∞ model reduction via LMI Given the stable and proper rational transfer matrix of order n

$$G(s) = C(sI - A)^{-1}B + D, \tag{3.19}$$

find a stable and proper rational matrix function $G_r(s) = C_r(sI - A_r)^{-1}B_r + D_r$ of order r such that :

1. $r < n$.
2. $\|G - G_r\|_\infty$ is minimized or, for pre-specified $\gamma > 0$, $\|G - G_r\|_\infty < \gamma$ (γ-suboptimal H_∞ model reduction).

The next result [22] provides necessary and sufficient conditions for the solution of the γ-suboptimal H_∞ lower order approximation problem.

Theorem 3.9. *There exists an r-th order transfer function G_r to solve the γ-suboptimal H_∞ model reduction problem if and only if there exist matrices $X \succ 0$ and $Y \succ 0$ such that the following conditions are satisfied:*

$$AX+XA^*+BB^* \prec 0, \;\; YA+A^*Y+C^*C \prec 0, \;\; \begin{bmatrix} X & \gamma I \\ \gamma I & Y \end{bmatrix} \succeq 0, \;\; \text{rank} \begin{bmatrix} X & \gamma I \\ \gamma I & Y \end{bmatrix} \leq n+r. \tag{3.20}$$

All γ-suboptimal r-th order models that correspond to a feasible matrix pair (X, Y) are given by

$$\begin{bmatrix} D_r & C_r \\ B_r & A_r \end{bmatrix} = G_1 + G_2 L G_3,$$

where $L \in \mathbb{C}^{(p+r)\times(m+r)}$ is strictly contractive and G_i, $i = 1,2,3$ depend on X, Y and γ.

The above LMI (3.20) are Lyapunov type inequalities combined with a rank condition. As discussed in Section 4, this makes the problem intrinsically difficult to handle. However, computing an optimal zeroth order approximation $G_0(s) = D_0$ does not require this rank condition, as pointed out in [22]. For this particular approximation problem, standard methods can thus be used.

Passive model reduction Passivity is a natural property for many physical systems, such as RLC networks, or certain mechanical systems. Passive systems are represented by *positive real* transfer matrix functions, see Definition 3.1. The above mentioned model reduction methods do not guarantee automatically the positive realness of the reduced order transfer matrix function, if the higher order transfer matrix function is *positive real*. A first solution to this problem has been obtained by Opdenacker and Jonckheere [45], who used balanced stochastic truncation techniques (see, for instance, [11, 21, 27]) to obtain a stable, positive real reduced order system. Later on, H_∞ norm error bounds on the approximation error have been derived in [9], by using similar methods. More recently, Wang and Huang [56], combined Hankel norm approximation techniques with the LMI approach in [22], and provided only sufficient conditions for the existence of a reduced order positive real transfer function with prescribed H_∞ norm error bound. They formulated the *optimal* H_∞ norm lower order approximation problem as a non-convex optimization problem in the following manner:

$$\inf_{P\succ 0, X\succ 0, Y\succ 0} \gamma \text{ subject to } \begin{bmatrix} A^*P + PA - L^*L & -PBW^{-1} - 2L^* \\ -W^{-T}B^*P - 2L & (\gamma^2+4)(D^*+D)^{-1} - 2I \end{bmatrix} \preceq 0 \qquad (3.21)$$

and to the LMI system (3.20). Here $\Xi(s) = L(sI-A)^{-1}B + W$ is a spectral factor of $G_*(s) + G(s)$. Let us point out that LMI (3.21) expresses a condition written exclusively in the terms of the initial data, $G(s)$, which severely restricts the class of systems to deal with.

Alternative reduced order passive modeling techniques for large RLC networks are based on Padé approximations and proved to be accurate and numerically stable. Unfortunately, the Padé reduced order model cannot be guaranteed to be passive for general RLC circuits. On the other hand, the Arnoldi based dimensional reduction technique [44] produces, under certain conditions, passive reduced-order for general RLC circuits. Using a similar approach, a general projection technique for computing reduced-order models, based on block Krylov subspaces has been proposed recently by Freund [15]. This type of technique ensures the passivity of the reduced-order model only for a particular class of transfer functions and matches only half as many moments as the Padé based techniques.

4 Conic problems and interior-point methods

In the previous sections we have seen that many important problems in control theory can be written in some special convex settings. Nowadays, linear matrix inequalities are considered to be easily solvable by standard software. However, the blind use of any LMI toolbox is potentially hazardous. Important issues like strict feasibility, nondegenerate operators, rank conditions, numerical complexity and stability are often disregarded by practitioners. Our aim is mainly to bring the reader's attention to these mathematical issues, which can prevent the numerical schemes from solving the problems at hand.

4.1 Elements of Convex Analysis

In order to treat our convex problems by the modern numerical schemes, we need to put them in a standard form. Let us start from the description of corresponding objects. We do that on a quite abstract level since the problems we are interested in have various nature (real/complex variables, real/hermitian/complex matrices, etc.).

Denote by $\mathcal{E}$ a finite dimensional linear vector space. And let $\mathcal{E}^*$ be the dual space; that is the finite-dimensional space of real-valued linear functions on $\mathcal{E}$. Clearly, $(\mathcal{E}^*)^* = \mathcal{E}$. For any $x \in \mathcal{E}$ and $s \in \mathcal{E}^*$ we denote

$$\langle s, x \rangle$$

the value of the function s at x. Sometimes, when we need to emphasize the space of the variables, the notation $\langle s, x \rangle_{\mathcal{E}}$ is used.

Let $\mathcal{E}_r$ be another finite dimensional linear vector space. Consider a linear operator

$$\mathcal{A} : \mathcal{E} \to \mathcal{E}_r^*.$$

Then we can define an *adjoint* linear operator $\mathcal{A}^*$ as follows:

$$\langle \mathcal{A}(x), u \rangle_{\mathcal{E}_r} = \langle \mathcal{A}^*(u), x \rangle_{\mathcal{E}} \quad \forall x \in \mathcal{E},\ u \in \mathcal{E}_r.$$

Thus, $\mathcal{A}^* : \mathcal{E}_r \to \mathcal{E}^*$. If $\mathcal{A} : \mathcal{E} \to \mathcal{E}^*$, then $\mathcal{A}^* : \mathcal{E} \to \mathcal{E}^*$.

Let $\mathcal{M}$ be a linear operator from $\mathcal{E}$ to $\mathcal{E}^*$:

$$\mathcal{M}(x) \in \mathcal{E}^*, \quad \forall x \in \mathcal{E}.$$

The operator $\mathcal{M}$ is called symmetric if $\mathcal{M}^* = \mathcal{M}$. It is called positive definite if

$$\langle \mathcal{M}(x), x \rangle > 0, \quad \forall x \in \mathcal{E} \setminus \{0\}.$$

If the above inequality is not strict, the operator $\mathcal{M}$ is called positive semidefinite. The set of all symmetric operators from $\mathcal{E}$ to $\mathcal{E}^*$ is denoted by $S(\mathcal{E})$; we use $S_+(\mathcal{E})$ to denote positive semidefinite operators and $S_+^0(\mathcal{E})$ to denote positive definite operators.

Let us fix some $\mathcal{M} \in S_+^0(\mathcal{E})$. Then we can define the Euclidean norm on $\mathcal{E}$:

$$\|x\|_{\mathcal{M}}^2 = \langle \mathcal{M}(x), x \rangle, \quad x \in \mathcal{E}.$$

Since $\mathcal{E}$ is a finite-dimensional space, all Euclidean norms on $\mathcal{E}$ are topologically equivalent.

In what follows we often work with convex cones in $\mathcal{E}$. A cone $\mathcal{K} \subseteq \mathcal{E}$ is called convex if

$$\begin{array}{rcl} x \in \mathcal{K}, \tau \geq 0 & \Rightarrow & \tau x \in \mathcal{K}, \\ x, y \in \mathcal{K} & \Rightarrow & x + y \in \mathcal{K}. \end{array}$$

For any convex cone $\mathcal{K} \subseteq \mathcal{E}$ we can define the dual cone $\mathcal{K}^* \subseteq \mathcal{E}^*$:

$$\mathcal{K}^* = \{s \in \mathcal{E}^* : \ \langle s, x \rangle \geq 0 \ \forall x \in \mathcal{K}\}.$$

If the cone $\mathcal{K} \subseteq \mathcal{E}$ with $\operatorname{int} \mathcal{K} \neq \emptyset$ is closed, convex and proper (contains no straight line), then the dual cone $\mathcal{K}^*$ also has all these properties. We call such cones the *regular* cones. Note that $(\mathcal{K}^*)^* = \mathcal{K}$. It is important that $s \in \operatorname{int} \mathcal{K}^*$ if and only if $\langle s, x \rangle > 0$ for all $x \in \mathcal{K}$.

Consider a real-valued function $f(x)$, $\operatorname{dom} f \subseteq \mathcal{E}$, $\operatorname{int}(\operatorname{dom} f) \neq \emptyset$. For differentiable functions we can define the gradient $f'(x)$ at some $x \in \operatorname{int}(\operatorname{dom} f)$:

$$f(x + h) = f(x) + \langle f'(x), h \rangle + o(\|h\|_{\mathcal{M}}), \quad x + h \in \operatorname{int}(\operatorname{dom} f).$$

Clearly, $f'(x) \in \mathcal{E}^*$. Note that the gradient does not depend on the choice of $\mathcal{M} \in S^0_+(\mathcal{E})$. Similarly, for twice differentiable functions we can define the Hessian $f''(x)$:

$$f(x+h) = f(x) + \langle f'(x), h\rangle + \frac{1}{2}\langle f''(x)h, h\rangle + o(\|h\|^2_{\mathcal{M}}), \quad x + h \in \mathrm{int}(\mathrm{dom}\, f).$$

Thus, $f''(x) \in S(\mathcal{E})$. If $f(x)$ is convex then $f''(x) \in S_+(\mathcal{E})$. If $f(x)$ is strongly convex then $f''(x) \in S^0_+(\mathcal{E})$.

Let $f(x)$, $x \in \mathrm{dom}\, f \subseteq \mathcal{E}$, be a closed convex function (its epigraph is a closed convex set). Then we can define the *conjugate* (or *dual*) function $f_*(s)$, $s \in \mathrm{dom}\, f_* \subseteq \mathcal{E}^*$:

$$f_*(s) = \sup_x \{-\langle s, x\rangle - f(x),\ x \in \mathrm{dom}\, f\}.$$

Note that $f_*(s)$ is also a closed convex function. Moreover, $(f^*)^* = f$.

Finally, let us introduce some special barrier functions, which are used in the interior-point schemes. Let $Q \subset \mathcal{E}$ be a convex set with nonempty interior. A closed convex three times differentiable function $f(x)$ is called *self-concordant* if $\mathrm{dom}\, f = \mathrm{int}\, Q$ and for any $x \in \mathrm{int}\, Q$ and $h \in \mathcal{E}$ we have

$$D^3 f(x)[h,h,h] \le 2[\langle f''(x)h, h\rangle]^{3/2},$$

where $D^3 f(x)[h,h,h]$ is the third-order differential of $f(x)$ at the direction h (notation: $f \in \mathcal{SC}(Q)$). Such a function is called a *self-concordant barrier* (denoted as $f \in \mathcal{SCB}_\nu(Q)$) if in addition there exists a constant $\nu \ge 1$ such that

$$\langle f'(x), h\rangle^2 \le \nu \langle f''(x)h, h\rangle, \quad \forall x \in \mathrm{int}\, Q,\ h \in \mathcal{E}.$$

The constant ν is called the parameter of the barrier. The simplest example of such a function is the logarithmic barrier for $\mathbb{R}_+ = \{x \in \mathbb{R} :\ x \ge 0\}$:

$$f(x) = -\ln x, \quad \nu = 1.$$

The most important operations, which preserve the above properties, are as follows:

$$f(x) \in \mathcal{SCB}_\nu(Q) \Rightarrow f(x) + \langle c, x\rangle \in \mathcal{SCB}_\nu(Q), \tag{4.1}$$

$$f(x) \in \mathcal{SCB}_\nu(Q) \Rightarrow f(x) \in \mathcal{SCB}_\nu(\{x \in Q :\ \mathcal{A}(x) = b\}), \tag{4.2}$$

$$f(x) \in \mathcal{SCB}_\nu(Q) \Rightarrow f(\mathcal{A}(y) + b) \in \mathcal{SCB}_\nu(\{y :\ \mathcal{A}(y) + b \in Q\}), \tag{4.3}$$

$$f_i(x) \in \mathcal{SCB}_{\nu_i}(Q_i),\ i = 1,2, \Rightarrow f_1(x) + f_2(x) \in \mathcal{SCB}_{\nu_1+\nu_2}(Q_1 \bigcap Q_2). \tag{4.4}$$

For regular cones we need to impose one more property of the barriers. Let $\mathcal{K} \subset \mathcal{E}$ be a regular cone. The function $f(x)$ is called a *ν-normal barrier* for $\mathcal{K}$ (denoted as $f \in \mathcal{NB}_\nu(\mathcal{K})$) if $f \in \mathcal{SCB}_\nu(\mathcal{K})$ and for any $x \in \mathrm{int}\, \mathcal{K}$ and $\tau > 0$ we have

$$f(\tau x) = f(x) - \nu \ln \tau. \tag{4.5}$$

For example,

$$f(X) = -\ln \det X \in \mathcal{NB}_n(S_+(\mathbb{R}^n)).$$

It is important that if $f(x)$ is a ν-normal barrier for $\mathcal{K}$ then $f_*(s)$ is a ν-normal barrier for $\mathcal{K}^*$.

4.2 Standard conic problems

In the previous sections we have seen two different types of convex problems: the *feasibility* problems and the *optimization* problems. Let us describe these problems in a general form. In what follows we always assume that we work with regular cones.

4.2.1 Feasibility problems

In feasibility problems we need to find a point from a convex set, which existence justifies some property of our object. We will see that any of these conic problems can be written either in the "primal" or in the "dual" form. Since the conic duality is symmetric, we have a choice to call any of these form the "primal" one. In order to avoid a misleading terminology, let us follow the following convention. The problems, in which the feasible set is formed as an intersection of a convex cone with an affine subspace, given by the equality constraints, will be always called the *primal* problems. The convex cone in this setting will be seen as the *primal* cone. The problems in which the feasible set is formed by an intersection of an image of a linear operator with some convex cone will be called the *dual* problems. The convex cone in this setting will be seen as the *dual* cone.

Let $c \in \mathcal{E}^*$ and $\mathcal{A}$ be a linear operator, $\mathcal{A} : \mathcal{E} \to \mathcal{E}_r^*$. Then $\mathcal{A}^* : \mathcal{E}_r \to \mathcal{E}^*$. The general conic feasibility problem has the following form:

$$\text{Find } y \in \mathcal{E}_r : \quad c - \mathcal{A}^*(y) \in \mathcal{K}^*. \tag{4.6}$$

In accordance to our convention, this is the dual problem. Note that the feasible set of this problem is a convex set (which may be empty). This problem can be treated by interior-point schemes if it is *strictly feasible*. That is there exists $\hat{y} \in \mathcal{E}_r$ such that $c - \mathcal{A}^*(\hat{y}) \in \operatorname{int} \mathcal{K}^*$.

In order to solve the problem (4.6) by interior-point schemes, we need to specify which point from the feasible set we are looking for. In our case it is convenient to specify this point as a minimizer of the following strictly convex function:

$$f_*(s) + \langle s, x_0 \rangle, \quad s = c - \mathcal{A}^*(y) \in \mathcal{K}^*, \; y \in \mathcal{E}_r,$$

where x_0 is an arbitrary (*scaling*) point from $\operatorname{int} \mathcal{K}$. Note that this function is strictly convex and its unique minimum exists for any feasible set in (4.6) with non-empty interior.

Thus, we come to the following form of our dual problem

$$\min_{s,y} [\; f_*(s) + \langle s, x_0 \rangle : \; s = c - \mathcal{A}^*(y) \;]. \tag{4.7}$$

The advantage of this formulation is that it can be rewritten in an equivalent primal form. For a completeness of presentation, let us sketch this transformation.

$$\begin{aligned}
&\min_{s,y} [\; f_*(s) + \langle s, x_0 \rangle : \; s = c - \mathcal{A}^*(y) \;] \\
&= \min_{s,y} \max_{x} [\; f_*(s) + \langle c - \mathcal{A}^*(y), x_0 \rangle + \langle s - c + \mathcal{A}^*(y), x \rangle \;] \\
&= \max_{x} \min_{s,y} [\, f_*(s) + \langle c, x_0 \rangle + \langle \mathcal{A}(x) - \mathcal{A}(x_0), y \rangle + \langle s - c, x \rangle] \\
&= \max_{x} [\; -f(x) - \langle c, x \rangle + \langle c, x_0 \rangle : \; \mathcal{A}(x) = \mathcal{A}(x_0) \;].
\end{aligned}$$

Thus, up to a constant term, our primal problem is

$$\min_{x} [\; f(x) + \langle c, x \rangle : \; \mathcal{A}(x) = \mathcal{A}(x_0) \;]. \tag{4.8}$$

Note that the objective function of this problem is self-concordant and we have a strictly feasible point $x = x_0$, which can be used as a starting point for interior-point schemes.

In fact, there is a particular case of the problem (4.8), which is much easier for numerical methods. That is the case $c = 0$, which corresponds to homogeneous feasibility problem (4.6). Then the objective function in (4.8) is a self-concordant barrier and (4.8) becomes the problem of finding the *analytic center* of the corresponding convex set. In Section 4.3 we will see how we can solve this problem by interior point schemes. And now, let us show that any non-homogeneous feasibility problem can be rewritten in homogeneous form.

Indeed, the initial non-homogeneous feasibility problem (4.6) is equivalent to the following:

$$\text{Find } y \in \mathcal{E}_r,\ \gamma \in \mathbb{R}: \quad \gamma c - \mathcal{A}^*(y) \in \mathcal{K}^*,\ \gamma > 0. \tag{4.9}$$

In order to put this problem in a standard form, let us define

$$\hat{\mathcal{E}}_r = \mathcal{E}_r \times \mathbb{R}, \quad \hat{\mathcal{E}} = \mathcal{E} \times \mathbb{R}, \quad \hat{\mathcal{K}}^* = \mathcal{K}^* \times \mathbb{R}_+,$$

$$\hat{\mathcal{A}}^* : \hat{\mathcal{E}}_r \to \hat{\mathcal{E}}^*, \quad \hat{\mathcal{A}}^* \left(\begin{bmatrix} y \\ \gamma \end{bmatrix} \right) = \begin{bmatrix} \mathcal{A}^* y - \gamma c \\ -\gamma \end{bmatrix}.$$

Then the problem (4.9) can be rewritten as

$$\text{Find } u \in \hat{\mathcal{E}}_r,\ -\hat{\mathcal{A}}^*(u) \in \hat{\mathcal{K}}^*. \tag{4.10}$$

Choosing some scaling points $x_0 \in \operatorname{int} \mathcal{K}$ and $\tau_0 > 0$, and using the regular barrier $f(x) - \ln \tau$ for $\hat{\mathcal{K}} = \mathcal{K} \times \mathbb{R}_+$, we come to the following primal problem:

$$\min_{x,\tau} \left[f(x) - \ln \tau : \ \hat{\mathcal{A}} \left(\begin{bmatrix} x \\ \tau \end{bmatrix} \right) = \hat{\mathcal{A}} \left(\begin{bmatrix} x_0 \\ \tau_0 \end{bmatrix} \right) \right].$$

Eliminating τ by the last linear equation in the system, we get the problem

$$\min_x [\ f_0(x) \equiv f(x) - \ln(\tau_0 - \langle c, x - x_0 \rangle) : \ \mathcal{A}(x) = \mathcal{A}(x_0)\ ]. \tag{4.11}$$

Note that the objective function in this problem is a self-concordant barrier:

$$f_0(x) \in \mathcal{SCB}_{\nu+1}(Q \equiv \{x \in \mathcal{K} : \ \langle c, x - x_0 \rangle \le \tau_0\}),$$

and the point $x = x_0$ is strictly feasible. At the same time, this problem captures both homogeneous and non-homogeneous feasibility problems. The only difference is that in the first case Q is a cone.

4.2.2 Convex optimization problems

Let $c \in \mathcal{E}^*$, $b \in \mathcal{E}_r^*$ and $\mathcal{A}$ be a linear operator, $\mathcal{A} : \mathcal{E} \to \mathcal{E}_r^*$. The general conic convex optimization problem has the following form:

$$\begin{array}{rl} & \min\limits_x \quad \langle c, x \rangle_{\mathcal{E}}, \\ \text{s.t.} & \mathcal{A}(x) = b, \\ & x \in \mathcal{K}. \end{array} \tag{4.12}$$

In accordance to our convention, that is the *primal* problem. This problem can be rewritten in the following dual form:

$$\begin{array}{rl} & \max\limits_{s,y} \quad \langle b, y \rangle_{\mathcal{E}_r}, \\ \text{s.t.} & s + \mathcal{A}^*(y) = c, \\ & s \in \mathcal{K}^*,\ y \in \mathcal{E}_r. \end{array} \tag{4.13}$$

In order to guarantee the equivalence of these two problems, we need the following assumption.

Assumption 4.1. The pair of the problems (4.12) and (4.13) is strictly feasible: there exist $\hat{x}$, $\hat{s}$ and $\hat{y}$ such that

$$\hat{x} \in \operatorname{int} \mathcal{K}, \quad \mathcal{A}(\hat{x}) = b, \quad \hat{s} \in \operatorname{int} \mathcal{K}^*, \quad \hat{s} + \mathcal{A}^*(\hat{y}) = c.$$

Sometimes, it is reasonable to treat the problems (4.12), (4.13) simultaneously, as a *primal-dual problem:*

$$\begin{array}{rl} & \min\limits_{x,s,y} \quad \langle c, x \rangle_{\mathcal{E}} - \langle b, y \rangle_{\mathcal{E}_r}, \\ \text{s.t.} & \mathcal{A}(x) = b, \ s + \mathcal{A}^*(y) = c, \\ & x \in \mathcal{K}, \ s \in \mathcal{K}^*, \ y \in \mathcal{E}_r. \end{array} \tag{4.14}$$

From Assumption 4.1 we get the optimal value of this problem equal to zero. Therefore the problem of finding ϵ-solution to (4.14) can be written as a feasibility problem:

$$\text{Find } (x, s, y) : \quad \left\{ \begin{array}{l} x \in \mathcal{K}, \ s \in \mathcal{K}^*, \ y \in \mathcal{E}_r, \\ \mathcal{A}(x) = b, \ s + \mathcal{A}^*(y) = c, \\ \langle c, x \rangle_{\mathcal{E}} - \langle b, y \rangle_{\mathcal{E}_r} \leq \epsilon. \end{array} \right. \tag{4.15}$$

4.3 Interior-point methods

As we have seen, the general convex feasibility problem can be written in the form

$$\min_x [f(x) : \ \mathcal{A}(x) = \mathcal{A}(x_0) \ ], \tag{4.16}$$

where Q is a convex set and $f \in \mathcal{SCB}_\mu(Q)$. This problem is solvable if and only if our initial feasibility problem has an interior solution.

The simplest algorithm for solving the problem (4.16) is the *Newton method.* For $s \in \mathcal{E}^*$ and $x \in \operatorname{int} Q$ define

$$\begin{array}{rcl} p_x(s) & = & \arg \max\limits_{h \in \mathcal{E}} [\ \langle s, h \rangle - \frac{1}{2} \langle f''(x)h, h \rangle : \ \mathcal{A}(h) = 0 \ ], \\ r_x(s) & = & \langle s, p_x(s) \rangle^{1/2}. \end{array} \tag{4.17}$$

Note that this problem is equivalent to the following system of linear equations :

$$\left\{ \begin{array}{c} f''(x) p_x(s) + \mathcal{A}^*(y_x(s)) = s \\ \mathcal{A}(p_x(s)) = 0 \end{array} \right. \tag{4.18}$$

Then the scheme of the *Damped Newton method* is as follows:

$$x_{k+1} = x_k - \frac{p_{x_k}(f'(x_k))}{1 + r_{x_k}(f'(x_k))}, \quad k = 0, 1, \ldots. \tag{4.19}$$

The method stops when $r_{x_k}(f'(x_k)) \leq \beta$ with some $\beta \in (0, 1)$. If (4.16) corresponds to (4.11), then the point

$$\bar{y} = (\tau_0 - \langle c, x_k - x_0 \rangle) \cdot y_{x_k}(f'(x_k))$$

is a solution to the feasibility problem (4.6). Denote by x^* the optimal solution to (4.16). Then the number of iterations of the Newton method is bounded by

$$\frac{f(x_0) - f(x^*)}{\beta - \ln(1 + \beta)}. \tag{4.20}$$

Another approach for solving the problem (4.16) is the path-following scheme. Let us define the *central path* of the problem (4.16) :

$$x(t) = \arg\min_x [f(x) - t\langle f'(x_0), x\rangle : \ \mathcal{A}(x) = \mathcal{A}(x_0)], \quad t \in [0,1].$$

Note that $x(1) = 0$ and $x(0) = x^*$. The path-following scheme updates the approximations to the central path as $t \to 0$:

$$\left.\begin{array}{rcl} s_{k+1} & = & f'(x_k) - \left(1 - \frac{0.2}{\sqrt{\mu}}\right)^{k+1} f'(x_0) \\ x_{k+1} & = & x_k - p_{x_k}(s_{k+1}) \end{array}\right\} \quad k = 0, 1, \ldots. \tag{4.21}$$

The stopping criterion for this scheme is again $r_{x_k}(f'(x_k)) \leq \beta$. The upper bound for the number of iterations of this scheme looks as follows:

$$\mathcal{O}\left(\sqrt{\mu} \ln \frac{\mu}{r_{x^*}(f'(x_0))}\right). \tag{4.22}$$

4.4 Applications

Let us now apply the above formalism to the problems described in the previous sections. For each of them, we need to specify

- two finite dimensional vector spaces and the corresponding inner products;
- the cone $\mathcal{K}$ and/or its dual $\mathcal{K}^*$;
- the linear operator $\mathcal{A}$ and/or its dual $\mathcal{A}^*$;
- the associated barrier functions.

From these data, one can compute all the necessary elements presented at the beginning of this section.

In the sequel, the set of $n \times n$ hermitian matrices and the set of $n \times n$ hermitian positive semidefinite matrices are denoted by $\mathcal{S}^n \doteq S(\mathbb{C}^n)$ and $\mathcal{S}^n_+ \doteq S_+(\mathbb{C}^n)$, respectively. In the context of linear matrix inequalities, the cone of positive semidefinite matrices is our main object of interest. In order to be concise, we now briefly mention all the related mathematical objects, which are common to these problems. The standard inner product on $\mathcal{S}^n$ is defined by

$$\langle X, Y\rangle_{\mathbb{C}} = \mathrm{Re}\,\mathrm{Trace}(XY^*). \tag{4.23}$$

where $\mathrm{Trace}\,A$ denotes the trace of A. Since the cone $\mathcal{S}^n_+$ is regular, we can associated a ν- normal barrier function to it. It is the logarithmic barrier function

$$f(X) = -\ln\det X, \quad X \in \mathrm{int}\,\mathcal{S}^n_+. \tag{4.24}$$

with parameter $\nu = n$. Moreover, the cone $\mathcal{S}^n_+$ is self-dual, i.e. $(\mathcal{S}^n_+)^* = \mathcal{S}^n_+$. The conjugate function is $f_*(S) = f(S) - n$.

4.4.1 Kalman-Yakubovich-Popov Lemma

The KYP Lemma is a characterization of positivity equivalent to checking feasibility of a given LMI

$$H(\tilde{X}) = H + \begin{bmatrix} E \\ F \end{bmatrix} \tilde{X} \begin{bmatrix} A^* & C^* \end{bmatrix} + \begin{bmatrix} A \\ C \end{bmatrix} \tilde{X} \begin{bmatrix} E^* & F^* \end{bmatrix} \succeq 0. \tag{4.25}$$

where $H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}$ is hermitian.

This is exactly a conic feasibility problem. To see this, we need to identify our objects. Our two finite dimensional vector spaces are $\mathcal{E} = \mathcal{S}^{n+p}$ and $\mathcal{E}_r = \mathcal{S}^n$. The standard inner products (4.23) are used on both spaces. Since we are dealing with LMI, it is not surprising that the primal and dual cones are both $\mathcal{S}_+^{n+p}$. They are associated with the standard logarithmic barriers. The dual linear operator $\mathcal{A}^* : \mathcal{E}_r \to \mathcal{E}^*$ is given by

$$\mathcal{A}^* : \mathcal{E}_r \to \mathcal{E}^* : \tilde{X} \to -\begin{bmatrix} E \\ F \end{bmatrix} \tilde{X} \begin{bmatrix} A^* & C^* \end{bmatrix} - \begin{bmatrix} A \\ C \end{bmatrix} \tilde{X} \begin{bmatrix} E^* & F^* \end{bmatrix}. \tag{4.26}$$

Identifying y with $\tilde{X}$ and c with H in (4.6) completes the proof of our statement.

Let us now write the primal problem which we actually solve, see (4.11). Let us fix $X_0 \in \operatorname{int} \mathcal{S}_+^{n+p}$ and $\tau_0 > 0$. The linear operator $\mathcal{A}$ is

$$\mathcal{A} : \mathcal{E} \to \mathcal{E}_r^* : X \to -\begin{bmatrix} E^* & F^* \end{bmatrix} X \begin{bmatrix} A \\ C \end{bmatrix} - \begin{bmatrix} A^* & C^* \end{bmatrix} X \begin{bmatrix} E \\ F \end{bmatrix}. \tag{4.27}$$

Since the barrier function associated to $\mathcal{E}$ is $F(X) = -\ln\det X$, we get the analytic center problem

$$\begin{array}{ll} \min & -\ln\det X - \ln(\tau_0 - \langle H, X - X_0\rangle_{\mathbb{C}}) \\ \text{s. t.} & \begin{bmatrix} E^* & F^* \end{bmatrix} X \begin{bmatrix} A \\ C \end{bmatrix} + \begin{bmatrix} A^* & C^* \end{bmatrix} X \begin{bmatrix} E \\ F \end{bmatrix} \\ & = \begin{bmatrix} E^* & F^* \end{bmatrix} X_0 \begin{bmatrix} A \\ C \end{bmatrix} + \begin{bmatrix} A^* & C^* \end{bmatrix} X_0 \begin{bmatrix} E \\ F \end{bmatrix}, \\ & X \in \mathcal{S}_+^{n+p}. \end{array} \tag{4.28}$$

This problem can be solved using the methods presented in the previous subsection. In order to apply a Damped Newton method or to follow the central path, we must be able to compute the gradient and the Hessian of the self-concordant function

$$f(X) = -\ln\det X - \ln(\tau_0 - \langle H, X - X_0\rangle_{\mathbb{C}}). \tag{4.29}$$

One can easily check that they are given by the expressions

$$f'(X) = -X^{-1} + \frac{H}{\tau_0 - \langle H, X - X_0\rangle_{\mathbb{C}}} \tag{4.30}$$

$$f''(X)\Delta = X^{-1}\Delta X^{-1} + \frac{H\Delta H}{(\tau_0 - \langle H, X - X_0\rangle_{\mathbb{C}})^2}. \tag{4.31}$$

The linear system (4.18), which is at the core of both optimization schemes, can thus be rewritten as

$$\begin{aligned} X^{-1}P_X(S)X^{-1} + \frac{HP_X(S)H}{(\tau_0 - \langle H, X - X_0\rangle_{\mathbb{C}})^2} \\ - \begin{bmatrix} E \\ F \end{bmatrix} Y_X(S) \begin{bmatrix} A^* & C^* \end{bmatrix} - \begin{bmatrix} A \\ C \end{bmatrix} Y_X(S) \begin{bmatrix} E^* & F^* \end{bmatrix} = S \end{aligned} \tag{4.32}$$

$$-\begin{bmatrix} E^* & F^* \end{bmatrix} P_X(S) \begin{bmatrix} A \\ C \end{bmatrix} - \begin{bmatrix} A^* & C^* \end{bmatrix} P_X(S) \begin{bmatrix} E \\ F \end{bmatrix} = 0 \tag{4.33}$$

where $P_X(S) \in S^{n+p}$ and $Y_X(S) \in S^n$. According to our notation, $P_X(S)$ is the common Newton-like direction associated to both schemes. Since we optimize on the matrix space $\mathcal{S}_+^{n+p}$, it is not surprising that this direction is a matrix. In order to obtain a unique solution to (4.32)–(4.33), we must assume that the linear operator $\mathcal{A}^*$ is nondegenerate . The following theorem gives a necessary and sufficient condition for our linear operator $\mathcal{A}^*$, see (4.26).

Theorem 4.2. *Let $A, B \in \mathbb{C}^{m\times n}, m \geq n$. Then the symmetric continuous Sylvester equation*

$$AXB^* + BXA^* = 0, \quad X = X^* \tag{4.34}$$

has only the trivial solution $X = 0$ if and only if the Kronecker canonical form of $\lambda B - A$ is such that (i) $\lambda B - A$ is a full column normal rank, (ii) there is no nilpotent block, (iii) the generalized eigenvalues of the Jordan block are not symmetric w.r.t the imaginary axis.

A sketch of the proof is included in the appendix. A similar result is easily derived for the symmetric discrete Stein equation $AXA^* - BXB^* = 0, X = X^*$.

4.4.2 Bounded real lemma and L_∞ norm

The bounded real lemma checks that the L_∞ norm of a transfer function $G(\lambda)$ is bounded by γ by checking feasibility of the LMI (4.25) with

$$H = \begin{bmatrix} 0 & 0 \\ 0 & \gamma^2 I \end{bmatrix} - \begin{bmatrix} B \\ D \end{bmatrix} \begin{bmatrix} B^* & D^* \end{bmatrix}, \tag{4.35}$$

see Section 2.3. The definitions of the above paragraph can thus be applied as well.

However, if we actually want to compute the L_∞ norm of a transfer function $G(\lambda)$, we can solve an optimization problem. First we need to introduce a non-negative variable $\delta = \gamma^2$ to convexify our problem. Its optimal value δ_{opt} is the square of the L_∞ norm of $G(\lambda)$ and is obtained through the optimization problem

$$\begin{array}{ll} \min & \delta \\ \text{s.t.} & -\begin{bmatrix} B \\ D \end{bmatrix} \begin{bmatrix} B^* & D^* \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \delta I \end{bmatrix} + \begin{bmatrix} E \\ F \end{bmatrix} X \begin{bmatrix} A^* & C^* \end{bmatrix} + \begin{bmatrix} A \\ C \end{bmatrix} X \begin{bmatrix} E^* & F^* \end{bmatrix} \succeq 0, \\ & \delta \geq 0, X = X^*. \end{array} \tag{4.36}$$

The finite dimensional vector spaces are $\mathcal{E} = \mathcal{S}^{n+p} \times \mathbb{R}$ and $\mathcal{E}_r = \mathcal{S}^n \times \mathbb{R}$. The inner product on $\mathcal{E}$ ($\mathcal{E}_r$) is obviously defined as

$$\langle (X, \alpha), (Y, \beta) \rangle = \langle X, Y \rangle_{\mathbb{C}} + \alpha\beta, \tag{4.37}$$

where $\langle X, Y \rangle_{\mathbb{C}}$ is the standard inner product on $\mathcal{S}^{n+p}$ ($\mathcal{S}^n$). The dual linear operator $\mathcal{A}^*$ is defined by

$$\mathcal{A}^* : \mathcal{E}_r \to \mathcal{E} : (X, \delta) \to (-\begin{bmatrix} 0 & 0 \\ 0 & +\delta I \end{bmatrix} - \begin{bmatrix} E \\ F \end{bmatrix} X \begin{bmatrix} A^* & C^* \end{bmatrix} - \begin{bmatrix} A \\ C \end{bmatrix} X \begin{bmatrix} E^* & F^* \end{bmatrix}, -\delta). \tag{4.38}$$

As can be seen immediately from the constraints, the dual cone $\mathcal{K}^*$ is $\mathcal{S}_+^{n+p} \times \mathbb{R}_+$. This cone is regular and self-dual. The corresponding $n+1$-normal barriers are

$$f((X, \delta)) = -\ln\det X - \ln\delta, \quad (X, \delta) \in \operatorname{int} K \tag{4.39}$$

$$f_*((S, \mu)) = -\ln\det S - \ln\mu - (n + p + 1), \quad (S, \mu) \in \operatorname{int} K^* \tag{4.40}$$

Identifying y with X and c with $(-\begin{bmatrix} B \\ D \end{bmatrix} \begin{bmatrix} B^* & D^* \end{bmatrix}, 0)^*$ in (4.13), it follows that (4.36) is a conic optimization problem. It can be solved up to an accuracy ϵ via the feasibility problem (4.15) and one of the optimization schemes presented before. The key ingredients are thus the gradient and the Hessian of an appropriate self-concordant barrier. Since the methodology has already been presented, the details are left out here.

4.4.3 Model Reduction

In the context of model reduction problems, our problems can usually be recast as conic feasibility or optimality problems, provided that the rank constraint is dropped. These constraints are non-convex and can greatly increase the complexity of solving our optimization problems. For instance, several well-known NP-hard problems can be formulated using rank constraints. We refer to [3, 40] for complexity results in systems and control.

Alternating projection algorithms can be used to enforce these "hard" constraints on the solution [12]. Note that these algorithms can only guarantee a local convergence to a feasible solution, which might be suboptimal. In general, optimization problems with rank constraints can be efficiently solved if and only if the optimal solutions to the corresponding relaxed problems can easily be converted to satisfy the rank constraints. For instance, conic optimization problems formulated on any cone of non-negative scalar polynomials naturally induce low-rank optimal matrices.

4.4.4 Non-negative matrix polynomials

First let us focus on the cone of matrix polynomials non-negative on the unit circle, see Section 3.2. The scalar product to be used for pseudo-polynomials $P(z) = \sum_{k=-n}^{n} P_k z^k$ and $Q(z) = \sum_{k=-n}^{n} Q_k z^k$ is defined as follows :

$$\langle P, Q \rangle_{\mathbb{D}} \doteq \sum_{k=-n}^{n} \langle P_k, Q_k \rangle_{\mathbb{C}}.$$

The primal optimization problem reads :

$$\min_{P \in \mathcal{K}} \{ \langle C, P \rangle_{\mathbb{D}} : \langle A_r, P \rangle_{\mathbb{D}} = b_r, r = 1, \ldots, q \}, \tag{4.41}$$

where $\mathcal{K}$ is the cone of matrix coefficients $P = [P_{-n}, \ \ldots, \ P_n \ ]$ of non-negative pseudo-polynomial matrices $P(z)$ on the unit circle. Remember that the coefficients of such matrices satisfy $P_{-i} = P_i^*$ and that $P \in \mathcal{K}$ necessarily implies

$$P_i = \sum_{k-l=i} Y_{k,l}, \quad i = -n, \ldots, n \tag{4.42}$$

where Y is a non-negative block matrix with blocks $Y_{k,l}, k, l = 0, \ldots n$ of dimension $p \times p$.

The dual cone $\mathcal{K}^*$ is made of the matrix coefficients $Q = [Q_{-n}, \ , \ldots, \ Q_n \ ]$ of the parahermitian pseudo-polynomials satisfying the constraint

$$\langle Q, P \rangle_{\mathbb{D}} \geq 0, \quad \forall P \in \mathcal{K}.$$

If $\mathcal{T}(Q)$ denotes the block Toeplitz matrix

$$\mathcal{T}(Q) \doteq \begin{bmatrix} Q_0 & Q_1 & \dots & Q_n \\ Q_1^* & Q_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & Q_1 \\ Q_n^* & \dots & Q_1^* & Q_0 \end{bmatrix}, \tag{4.43}$$

the dual cone $\mathcal{K}^*$ is characterized by $\mathcal{T}(Q) \succeq 0$ [17].

Therefore the dual optimization problem of (4.41) becomes

$$\max_{u_1,\dots,u_r} \{\sum_{r=1}^{q} b_r u_r : \mathcal{T}(C - \sum_{r=1}^{q} u_r A_r) \succeq 0\} \tag{4.44}$$

for which the appropriate barrier function is $f(u) = -\ln\det \mathcal{T}(C - \sum_{r=1}^{q} A_r u_r)$.

From a numerical point of view, this dual formulation has a considerable advantage over the primal form (4.41) since it involves an optimization scheme in a space of variables of dimension q rather than $(2n+1)p^2$. Any optimization problem of this type can be solved efficiently with the help of interior point methods [42]. Their numerical implementation requires the calculation of the first and second derivatives of the barrier function. They can be expressed as follows :

$$\begin{aligned} \frac{\partial f(u)}{\partial u_r} &= \langle \mathcal{T}(S)^{-1}, \mathcal{T}(A_r)\rangle_{\mathbb{C}}, \\ \frac{\partial^2 f(u)}{\partial u_r \partial u_s} &= \langle \mathcal{T}(S)^{-1}\mathcal{T}(A_r)\mathcal{T}(S)^{-1}, \mathcal{T}(A_s)\rangle_{\mathbb{C}}, \end{aligned} \tag{4.45}$$

where $S = C - \sum_{r=1}^{q} A_r u_r$.

All these inner products can be efficiently computed using the displacement structure of block Toeplitz matrices [17, 30]. The cost of one Newton step is found to be approximately equal to $\mathcal{O}\left(q r_t p^2 n \ln^2 n + q^2 p^2 n\right)$, where r_t is the displacement rank of $\mathcal{T}$ ($r_t \leq 2m$). Since interior-points methods require $\mathcal{O}\left(\sqrt{np}\ln\frac{1}{\epsilon}\right)$ Newton steps to solve the optimization problem (4.44) up to an accuracy ϵ [42], this is a remarkable result for solving an optimization problem in a $(2n+1)p^2$-dimensional vector space, subject to q linear constraints and p semi-infinite inequality constraints (see (4.41)).

This example illustrates the fact that tailor-made algorithms are often much faster than general algorithms : they can rely on both the problem formulations and the underlying problem structures. Note that the optimal polynomial $P_{\text{opt}}(z)$ can often be obtained from a low-rank block matrix Y_{opt}, via spectral factorization of $P_{\text{opt}}(z)$.

Similar results have been developed to deal with other curves of the complex plane, see [17]. However, it should be pointed out that the resulting LMI can be inherently difficult to solve. For instance, if we use the standard polynomial basis $1, x, \dots, x^{2n}$ to describe the set of matrix polynomials non-negative on the real line, the previous formula hold provided that the block Toeplitz structure $\mathcal{T}(Q)$ is replaced by a block Hankel structure. In particular, the constraint of the dual optimization problem involves a positive semidefinite block Hankel matrix. Since the set of positive definite Hankel matrices is composed by exponentially ill-conditioned matrices, numerical errors definitely corrupt the computation of the gradient and the Hessian. By using the Chebyshev basis as functional basis, the block Hankel structure is replaced by a block Hankel-plus-Toeplitz structure which is intrinsically better than the Hankel one. One could also reformulate the problem directly on the unit circle and get a block Toeplitz structure.

This last example illustrates that formulating a problem as an LMI is only half the way to solving the problem. It should be checked that (i) the problem is well-posed and (ii) the algorithm is sufficiently stable or accurate for this particular problem.

5 Alternative methods

5.1 Level sets

In order to simplify our notation, we restrict ourselves here to a parahermitian transfer function $\Phi(s)$ with respect to the $j\omega$ axis although everything easily extends to the unit circle as well. Since $\Phi(s)$ is parahermitian it is also hermitian for every point $s = j\omega$:

$$[\Phi(j\omega)]^* = \Phi(j\omega), \quad \omega \in \mathbb{R}. \tag{5.1}$$

Two important quantities of such transfer functions are the maximum and minimum eigenvalues on $\omega \in \mathbb{R}$:

$$\xi_{min} \doteq \min_{\omega \in \mathbb{R}} \lambda_{min} \Phi(j\omega), \quad \xi_{max} \doteq \max_{\omega \in \mathbb{R}} \lambda_{max} \Phi(j\omega). \tag{5.2}$$

It is easy to see e.g. that $\Phi(j\omega)$ is non-negative on $\omega \in \mathbb{R}$ if and only if $\xi_{min} \geq 0$ and that $\Phi(j\omega)$ is bounded on $\omega \in \mathbb{R}$ if and only if $\xi_{max} \prec \infty$. This can be used to solve important problems related to transfer functions :

1. The L_∞ norm of an arbitrary transfer function $G(s)$ is defined as

   $$\gamma_{max} \doteq \max_{\omega \in \mathbb{R}} \sigma_{max} G(j\omega)$$

   which is clearly equal to $\xi_{max}^{1/2}$ of $\Phi(s) \doteq G(s)G_*(s)$. Notice that if $G(s)$ is stable, the L_∞ norm is also the H_∞ norm of the transfer function.

2. Positive realness of $G(s)$ implies

   $$0 \leq \min_{\omega \in \mathbb{R}} \lambda_{min}[G(j\omega) + G_*(j\omega)]$$

   which amounts to checking $0 \leq \xi_{min}$ for $\Phi(s) = G(s) + G_*(s)$.

We describe now a level set algorithm to find the extremal values ξ_{min} and ξ_{max} of the eigenvalues of $\Phi(j\omega)$, $\omega \in \mathbb{R}$. Since both problems are dual to each other, we focus on ξ_{max} only. It is shown in [33] that a hermitian matrix $\Phi(j\omega)$ of a real variable ω has real analytic eigenvalues as a function of ω, if $\Phi(j\omega)$ is itself analytic in ω. Since we consider here rational functions of $s = j\omega$ – where the "frequency" ω is real – this is certainly the case. In Figure 2, we show these functions $\lambda_i(\omega)$ for a 2×2 matrix $\Phi(j\omega)$. We also indicate a level ξ_0 for which we want to check if there is any eigenvalue $\lambda_i(\omega) = \xi_0$. Clearly these are the intersections of the level ξ_0 with the eigenvalues of $\Phi(j\omega)$. Assume that these intersections occur at frequencies ω_i. Since

$$\det(\Phi(j\omega_i) - \xi_0 I) = 0$$

each frequency ω_i is an imaginary axis zero of the shifted transfer function $\Phi(s) - \xi_0 I$. These can be computed as the eigenvalues of the corresponding zero pencil that are located on the $j\omega$ axis. Notice that if there are no imaginary axis eigenvalues, then the level ξ_0 does not intersect the eigenvalue plots and hence

$$\xi_0 < \xi_{min} \quad \text{or} \quad \xi_0 > \xi_{max}.$$

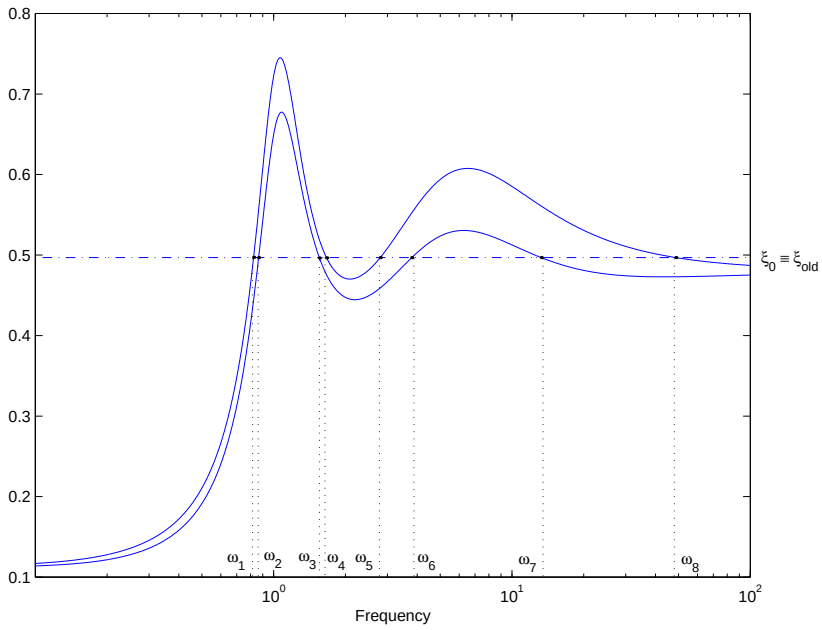

Figure 2: Level set iterations

In order to find a value ξ_0 for which there are eigenvalue crossings one can e.g. choose $\xi_0 = \lambda_{max}\Phi(j\omega_0)$ for an arbitrary value ω_0. Using these ingredients, one then derives a bisection algorithm to find ξ_{max} : each interval will contain an upper bound ξ_{up} and a lower bound ξ_{lo} for ξ_{max} and the bisection method checks whether there are eigenvalues on the $j\omega$ axis equal to the new level $(\xi_{lo} + \xi_{up})/2$ [8]. This algorithm has obviously linear convergence.

A faster convergence can be obtained by using information on the eigenvalue functions (see [4, 18]). Start from a point ξ_{old} which intersects the eigenvalues of $\Phi(j\omega)$ as in Figure 2, and obtain from this the intervals for which $\lambda_{max}(\omega) > \xi_{old}$ (these are called the level sets for ξ_{old}). In Figure 2 these are the intervals $[\omega_1, \omega_4]$ and $[\omega_5, \omega_8]$ (in this context we need to define $\lambda_{max}(\omega)$ as the piecewise analytic function that is maximal at each frequency ω). In [18] it is shown how to use the information of the derivative of $\lambda_{max}(\omega)$ in each point in order to determine the relevant "level sets". It is also shown there how to obtain these derivatives at little extra cost from the eigenvalue problem of the underlying zero pencil. Using these level sets and the derivative of $\lambda_{max}(\omega)$ at their endpoints, one then constructs a new frequency ω_{new} which is a good estimate of an extremal frequency ω_{max} :

$$\xi_{max} = \lambda_{max}[\Phi(j\omega_{max})] = \max_{\omega} \lambda_{max}[\Phi(j\omega)].$$

It is shown in [18] that such a scheme has global linear convergence and at least cubic asymptotic convergence. Each step requires the calculation of the largest eigenvalue ξ_{new} of $\Phi(j\omega_{new})$ and the eigenvalues and eigenvectors of the zero pencil defining the zeros of $\Phi(s) - \xi_{new}I$. The complexity of each iteration is thus cubic in the dimensions of the system matrix of $\Phi(s)$.

5.2 Cutting plane methods

In order to simplify our description of cutting plane methods, we now consider the general formulation of a convex optimization problem :

$$\min\{f(x) : x \in Q \subset \mathbb{R}^n\} \tag{5.3}$$

where $f(x)$ is a convex function and Q is a closed convex set.

As the function $f(x)$ is convex, there exists at any $x \in \operatorname{dom} f$ at least one *subgradient* $g(x)$ such that

$$f(y) \geq f(x) + \langle g(x), y - x\rangle, \quad \forall y \in \operatorname{dom} f. \tag{5.4}$$

If f is not differentiable at x, then the subgradient is not unique. In that case, the set $\partial f(x)$ of all subgradients is called the *subdifferential* of f at x.

Cutting plane methods belong to the set of optimization methods based on *first-order oracles*. At each point $x \in \mathbb{R}^n$, such an oracle returns the following information :

- If $x \in Q$, it computes the function value $f(x)$ and an associated subgradient $g(x)$.
- If $x \notin Q$, it returns a separating hyperplane g : $\langle g, y - x\rangle \leq 0, \quad \forall y \in Q$.

At each iteration, these schemes update the current iterate using the information returned by this oracle at the current *query point*.

The generic cutting plane algorithm works as follows. At each iteration k, it refines a so-called *localization set* $\mathcal{L}_k$, which contains the optimal solution x^*, using the information returned by the oracle. More precisely, the following operations are performed

1. Get a new query point x_k as a "center" of the localization set $\mathcal{L}_k$;
2. Call the oracle at this point and generate an associated cutting plane : $\langle g_k, x_k - x^*\rangle \geq 0$. Depending on the location of x_k, this cut is called either an optimality cut ($x_k \in Q$) or a feasibility cut ($x_k \notin Q$).
3. Update the localization set : $\mathcal{L}_{k+1} = \mathcal{L}_k \cap \{x : \langle g_k, x_k - x\rangle \geq 0\}$.

There exist several cutting planes methods (inscribed ellipsoid method, volumetric center method, analytic center method,...), which mainly differ in the choice of the next query point [14, 13]. The goal of an optimality cut is to reduce the objective function value by selecting the appropriate part of the localization set. A feasibility cut separates the current query point from the feasible set. Both cuts are performed using exclusively the oracle output. Figure 3 illustrates the procedure. Note that the oracle is clearly problem-specific : it should be provided by the user of any cutting plane algorithm, see e.g. [20].

The main drawback is the memory requirement needed by this class of methods. Indeed, most of the schemes keep all the cutting planes from the very first iteration. Hopefully, the use of appropriate cuts and of modern computers tends to reduce this problem. The advantages are the optimal efficiency estimate and the possible avoidance of extra variables. Let us exemplify this last issue with a one-dimensional problem. As seen in Section 2, the problem of minimizing the L_∞ norm of a transfer function $G(\lambda)$ can be formulated as

$$\begin{array}{ll} \min & \gamma^2 \\ \text{s. t.} & \gamma^2 I - G(\lambda)G_*(\lambda) \succ 0 \end{array} . \tag{5.5}$$

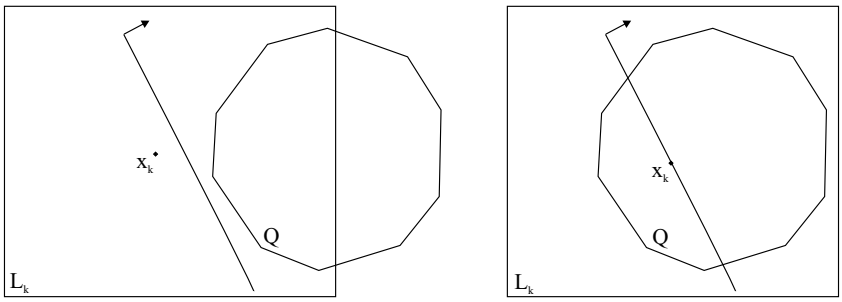

Figure 3: Cutting plane methods : feasibility cut (left) and optimality cut (right)

Using the KYP Lemma, the equivalent optimization problem for the continuous-time case is

$$\begin{array}{ll} \min & \gamma^2 \\ \text{s. t.} & -\begin{bmatrix} B \\ D \end{bmatrix} \begin{bmatrix} B^* & D^* \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \gamma^2 I \end{bmatrix} + \begin{bmatrix} E \\ F \end{bmatrix} X \begin{bmatrix} A^* & C^* \end{bmatrix} + \begin{bmatrix} A \\ C \end{bmatrix} X \begin{bmatrix} E^* & F^* \end{bmatrix} \succ 0. \\ & X = X^* \\ & \gamma^2 \geq 0 \end{array} \quad (5.6)$$

However, the latter formulation introduces an extra symmetric matrix X of size n. For large n, these $\frac{n(n+1)}{2}$ variables can be difficult to handle : they greatly increase the size of the linear systems to be solved at each iteration. Apart from an important complexity increase, it also use an excessive amount of memory. This could prevent any generic LMI solver from resolving the problem with an acceptable execution time. The former formulation does not introduce this variable X. It allows us to use a cutting plane algorithm, for which we briefly sketch an oracle. Let $\delta_k = \gamma_k^2$ be the current square L_∞ norms of $G(\lambda)$. Then at each iteration, the linear matrix inequality

$$\delta_k I - G(\lambda)G_*(\lambda) \succ 0, \quad \forall \lambda \in \Gamma \quad (5.7)$$

is checked by using appropriate pencils. Indeed, we need to find λ such that $\det(\delta_k I - G(\lambda)G_*(\lambda)) = 0$, which is a procedure similar to the level set iteration (see the previous subsection). Cuts are then easily derived from the above inequality. For instance, feasibility cuts are obtained using any eigenvector of (5.7) corresponding to a violating λ. Note that the level sets method applied to a convex function is a one-dimensional cutting plane method.

Nowadays, there is a growing interest for the use of cutting planes methods in systems and control [32, 46]. Theoretical properties (complexity estimates, ...) of these schemes are extensively studied in the optimization literature : it gives a strong background to any problem-specific algorithm based on them. Indeed, it is of paramount importance to provide these schemes with oracles adapted to the problem structures. A well devised oracle is the key ingredient for an efficient algorithm.

However, these cutting planes method could also suffer from (numerical) instabilities. For instance, let us consider a conic optimization problem on the cone of polynomials of degree $2n$ non-negative on the real line and the analytic center cutting plane method. The dual formulation allows us to use the Hankel structure, which is potentially ill-conditioned. The corresponding semidefinite constraint is thus difficult to check numerically for moderate n. The primal problem does not necessarily display a better behavior, as shown by a careful analysis of the cutting plane scheme : the next query is obtained by solving a sequence of linear systems with a potentially positive definite Hankel structure ! This class of problems must thus be solved with care. In this case, the

rule of thumb is to translate the problem from the real line to the unit circle, to solve the problem in this much more stable setting and to translate back the solution.

A Appendix

A.1 KYP Lemma for generalized state-space systems

Lemma A.1. *Let* $\left[\begin{array}{c} A-\lambda E \\ C-\lambda F \end{array}\right]$ *have only right Kronecker blocks in its canonical form, then there exists an invertible transformation* $M = \left[\begin{array}{cc} U & W \\ V & Y \end{array}\right]$ *such that*

$$M \left[\begin{array}{c} A-\lambda E \\ C-\lambda F \end{array}\right] = \left[\begin{array}{c} \tilde{A}-\lambda I \\ \tilde{C} \end{array}\right], \tag{A.1}$$

where $(\tilde{A}, \tilde{C})$ *is observable and* $\tilde{A}$ *is strictly stable.*

Proof. Since $\left[\begin{array}{c} E \\ F \end{array}\right]$ has full column rank, it is easy to construct $M \cdot \left[\begin{array}{c} E \\ F \end{array}\right] = \left[\begin{array}{c} I \\ 0 \end{array}\right]$ to obtain the required form (A.1). This does not affect the Kronecker canonical form of the new pencil and hence $(\tilde{A}, \tilde{C})$ must be observable. The strict stability of $\tilde{A}$ is then easily obtained from a pole placement argument, of which the corresponding transformation can be incorporated in M. □

If we now define $T(\lambda) = (C-\lambda F)(\lambda E - A)^{-1}$ and $\tilde{T}(\lambda) = \tilde{C}(\lambda I - \tilde{A})^{-1}$ then clearly the rows of

$$\left[\begin{array}{cc} T(\lambda) & I_p \end{array}\right] M \doteq \left[\begin{array}{cc} T(\lambda)U + V & T(\lambda)W + Y \end{array}\right] \quad \text{and} \quad \left[\begin{array}{cc} \tilde{T}(\lambda) & I_p \end{array}\right]$$

both span the left null space of (A.1). Therefore $T_1(\lambda) \doteq T(\lambda)W + Y$ must be invertible and

$$\left[\begin{array}{cc} T(\lambda) & I_p \end{array}\right] M = T_1(\lambda)^{-1} \left[\begin{array}{cc} \tilde{T}(\lambda) & I_p \end{array}\right]. \tag{A.2}$$

This then allows us to prove the following theorem :

Theorem A.2. *Given* $A, E \in \mathbb{C}^{n\times n}$, $C, F \in \mathbb{C}^{p\times n}$, $H \in \mathbb{C}^{(p+n)\times(p+n)}$ *where* $H = H^*$, $\det(\lambda E - A) \not\equiv 0$ *and* $P \left[\begin{array}{c} A-\lambda E \\ C-\lambda F \end{array}\right] Q = \mathrm{diag}\{L_{c_j}(\lambda)\}$, *the following statements are equivalent for* $T(\lambda) = (C-\lambda F)(\lambda E - A)^{-1}$ *:*

1. $$\Phi(\lambda) \doteq \left[\begin{array}{cc} T(\lambda) & I_p \end{array}\right] \left[\begin{array}{cc} H_{11} & H_{12} \\ H_{21} & H_{22} \end{array}\right] \left[\begin{array}{c} T_*(\lambda) \\ I_p \end{array}\right] \succeq 0 \quad \text{for almost all } \lambda \in \Gamma \tag{A.3}$$

2. *there exist a hermitian matrix* $X = X^* \in \mathbb{C}^{n\times n}$ *such that*

$$H(X) \succeq 0.$$

Proof. Since $T_1(\lambda) \doteq T(\lambda)W + Y$ is invertible we have

$$T_1(\lambda)^{-1}\Phi(\lambda)T_{1*}(\lambda)^{-1} = \left[\begin{array}{cc} \tilde{T}(\lambda) & I_p \end{array}\right] \tilde{H} \left[\begin{array}{c} \tilde{T}_*(\lambda) \\ I_p \end{array}\right] \succeq 0 \quad \text{for almost all } \lambda \in \Gamma \tag{A.4}$$

where $\tilde{H} \doteq MHM^*$. The expression on the right is now in standard state space form with a stable matrix $\tilde{A}$, for which it is known to be equivalent to the existence of a hermitian matrix X such that $\tilde{H}(X) \succeq 0$ [48]. Since M is invertible, this is also equivalent to $H(X) \succeq 0$ for the same matrix X. □

A.2 Symmetric Sylvester equation

Theorem A.3. *Let $A, B \in \mathbb{C}^{m\times n}, m \geq n$. Then the symmetric continuous Sylvester equation*

$$AXB^* + BXA^* = 0, \quad X = X^* \tag{A.5}$$

has only the trivial solution $X = 0$ if and only if the Kronecker canonical form of $\lambda B - A$ is such that (i) $\lambda B - A$ is a full column normal rank pencil, (ii) there is no nilpotent block, (iii) the generalized eigenvalues of the Jordan block are not symmetric w.r.t the imaginary axis.

Proof. Without loss of generality, the pencil $\lambda B - A$ is assumed to be in Kronecker canonical form, i.e.

$$\lambda B - A = \mathrm{diag}(\{L_{l_i}(\lambda)\}_{i=1}^{s}, \{L_{r_i}^T(\lambda)\}_{i=1}^{t}, I - \lambda N, \lambda I - J) \tag{A.6}$$

where $\{L_k(\lambda)\}$ is the $(k+1)\times k$ bidiagonal pencil

$$L_k(\lambda) = \begin{bmatrix} \lambda & & & \\ 1 & \ddots & & \\ & \ddots & & \\ & & \ddots & \\ & & \ddots & \lambda \\ & & & 1 \end{bmatrix}, \tag{A.7}$$

N is nilpotent and both N and J are in Jordan canonical form.

"If part". Let us partition the matrix X according to the structure of the canonical form of $\lambda B - A$:

$$A_k X_{kl} B_l^* + B_k X_{kl} A_l^* = 0. \tag{A.8}$$

If conditions (i)–(iii) are satisfied, all the equations (A.8) have only the trivial solution $X_{kl} = 0$ as can easily be checked by inspection. Therefore, $X = 0$.

"Only if part". If at least one of the conditions (i)–(iii) does not hold, there always exists a non-zero solution to at least one of the equations (A.8). To see this, consider the diagonal block X_{ii} associated with the unsatisfied condition. For instance, if (i) is not satisfied, set $X_{kk} = [\,1, 0, \cdots, 0\,]^T\,[\,1, 0, \cdots, 0\,]$ for the index k corresponding to one of the $L_{r_l}^T$ blocks. □

Acknowledgments

This paper presents research supported by NSF contract CCR-99-12415 and by the Belgian Programme on Inter-university Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. A research fellowship from the Belgian National Fund for Scientific Research is also gratefully acknowledged by the second author. The fourth author has been partially supported by NATO. The scientific responsibility rests with its authors.

References

[1] B. Alkire and L. Vandenberghe. Convex optimization problems involving finite autocorrelation sequences. *Math. Programming*, 2001. Submitted.

[2] H. Berghuis and H. Nijmeijer. A passivity approach to controller-observer design for robots. *IEEE Trans. on Robotics and Automation*, 9(6):741–754, 1993.

[3] V.D. Blondel and J.N. Tsitsiklis. A survey of computational complexity results in systems and control. *Automatica*, 36(9):1249–1274, 2000.

[4] S. Boyd, V. Balakrishnan, and P. Kabamba. A bisection method for computing the H_∞ norm of a transfer matrix and related problems. *Maths. Contr. Sig. Sys.*, 2(3):207–219, 1989.

[5] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*, volume 15 of *SIAM Studies in Applied Mathematics*. Society of Industrial and Applied Mathematics, Philadelphia, PA, 1994.

[6] B. Brogliato, I.D. Landau, and R. Lozano. Adaptive motion control of robot manipulators: a unified approach based on passivity. *Int. J. of Robust and Nonlinear Control*, 1(3):187–202, 1992.

[7] B. Brogliato and R. Lozano. Passive least-squares-type estimation algorithm for direct adaptive control. *Int. J. of Adaptive Control and Signal Processing*, 6(1):35–44, 1992.

[8] R. Byers. A bisection method for measuring the distance of a stable matrix to the unstable matrices. *SIAM J. Sci. Statist. Comput.*, 9(5):875–881, September 1988.

[9] X. Chen and J.T. Wen. Positive realness preserving model reduction with H^∞ norm bounds. *IEEE Trans. Circuits Syst. I*, 42(1):23–29, 1995.

[10] T.N. Davidson, Z.-Q. Luo, and J.F. Sturm. Linear matrix inequality formulation of spectral mask constraints. *Submitted to the IEEE Transactions on Signal Processing, September 2000.*, 2000.

[11] U.B. Desai and D. Pal. A transformation approach to stochastic model reduction. *IEEE Trans. Automat. Control*, 29(12):1097–1900, 1984.

[12] L. El Ghaoui and S.-I. Niculescu, editors. *Advances in Linear Matrix Inequality Methods in Control*. Advances in Design and Control. SIAM, 2000.

[13] S. Elhedhli, J.-L. Goffin, and J.-P. Vial. Nondifferentiable optimization: Cutting plane methods. In C. Floudas and P. Pardalos, editors, *Encyclopedia on Optimization*. Kluwer Academic Publishers, June 2001.

[14] S. Elhedhli, J.-L. Goffin, and J.-P. Vial. Nondifferentiable optimization: Introduction, applications and algorithms. In C. Floudas and P. Pardalos, editors, *Encyclopedia on Optimization*. Kluwer Academic Publishers, 2001.

[15] R. W. Freund. Passive reduced-order modeling via Krylov-subspace methods. In *Proceedings of the 14th Symposium on Mathematical Theory of Networks and Systems*, Perpignan, France, June 19–23, 2000.

[16] Y. Genin, Y. Hachez, Yu. Nesterov, and P. Van Dooren. Convex optimization over positive polynomials and filter design. In *Proceedings "UKACC International Conference on Control 2000"*, University of Cambridge, UK, September 2000.

[17] Y. Genin, Y. Hachez, Yu. Nesterov, and P. Van Dooren. Optimization problems over positive pseudo-polynomial matrices. *SIAM J. Matrix Anal. Appl.*, 2000. submitted.

[18] Y. Genin, P. Van Dooren, and V. Vermaut. Convergence of the calculation of H_∞ norms and related questions. *Proceedings MTNS-98*, pages 429–432, July 1998.

[19] K. Glover. All optimal Hankel norm approximations of linear multivariable systems and their L_∞-error bounds. *Int. J. Control*, 26:1115–1193, 1984.

[20] J. Gondzio, O. du Merle, R. Sarkissian, and J.-Ph. Vial. ACCPM – a library for convex optimization based on an analytic center cutting plane method. *European Journal of Operational Research*, 94:206–211, 1996. See also `http://ecolu-info.unige.ch/logilab/software/` for the latest implementation.

[21] M. Green. Balanced stochastic realizations. *Linear Algebra Appl.*, 98:211–247, 1988.

[22] K.M. Grigoriadis. Optimal H_∞ model reduction via linear matrix inequalities: continuous- and discrete-time cases. *Systems Control Lett.*, 26:321–333, 1995.

[23] E. J. Grimme, D. C. Sorensen, and P. Van Dooren. Model reduction of state space systems via an implicitly restarted Lanczos method. *Numerical Algorithms*, 12(1–2):1–31, 1995.

[24] W. M. Haddad and D. S. Bernstein. Robust stabilization with positive real uncertainty: beyond the small gain theorem. *Systems Control Lett.*, 17:191–208, 1991.

[25] W. M. Haddad and D. S. Bernstein. Explicit construction of quadratic Lyapunov functions for the small gain, positivity, circle, and Popov theorems and their application to robust stability. part I: continuous-time theory. *Int. J. Robust Nonlinear Control*, 3:313–339, 1993.

[26] W. M. Haddad and D. S. Bernstein. Explicit construction of quadratic Lyapunov functions for the small gain, positivity, circle, and Popov theorems and their application to robust stability. part II: discrete-time theory. *Int. J. Robust Nonlinear Control*, 4:249–265, 1994.

[27] P. Harshavardhana, E.A. Jonckheere, and L.M. Silverman. Stochastic balancing and approximation - stability and minimality. *IEEE Trans. Automat. Control*, 29(8):744–746, 1984.

[28] D.J. Hill and P.J. Moylan. Dissipative dynamical systems: basic input-output and state properties. *J. Franklin Institute*, 309(5):327–357, 1980.

[29] T. Kailath. *Linear Systems*. Prentice-Hall, Englewood Cliffs, New Jersey, 1980.

[30] T. Kailath and A. H. Sayed. Displacement structure: theory and applications. *SIAM Rev.*, 37:297–386, 1995.

[31] R.E. Kalman. Lyapunov functions for the problem of Lur'e in automatic control. *Proc. Nat. Acad. Sci. USA*, 49:201–205, 1963.

[32] C. Kao, U. Jonsson, and A. Megretski. A cutting plane algorithm for robustness analysis of periodically time-varying systems. *IEEE Transactions on Automatic Control*, 46(4):579–592, April 2001.

[33] T. Kato. *Perturbation Theory for Linear Operators*. Springer-Verlag, New York, 1965.

[34] D. Kavranoglu and M. Bettayeb. Characterization of the solution to the optimal H_∞ model reduction problem. *Systems Control Lett.*, 20:99–107, 1993.

[35] P. Lancaster and L. Rodman. *The Algebraic Riccati Equation*. Oxford University Press, 1995.

[36] M. S. Mahmoud, Y. C. Soh, and L. Xie. Observer-based positive real control of uncertain linear systems. *Automatica*, 35:749–754, 1999.

[37] D.C. McFarlane and K. Glover. *Robust Controller Design using Normalized Coprime Factor Plant Descriptions*, volume 138 of *Lecture notes in Control and Information Sciences*. Springer-Verlag, 1990.

[38] A. Megretski and A. Rantzer. System analysis via integral quadratic constraints. *IEEE Trans. Automat. Control*, 42:819–830, 1997.

[39] B. Moore. Principal components analysis in linear systems: controllability, observability and model reduction. *IEEE Trans. Automat. Control*, 26(1):17–31, 1981.

[40] A. Nemirovskii. Several NP-hard problems arising in robust stability analysis. *Maths. Contr. Sig. Sys.*, 6:99–105, 1993.

[41] Yu. Nesterov. Squared functional systems and optimization problems. In H.Frenk et al., editors, *High Performance Optimization*, chapter 17, pages 405–440. Kluwer Academic Publishers, 2000.

[42] Yu. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*, volume 13 of *SIAM Studies in Applied Mathematics*. Society of Industrial and Applied Mathematics, Philadelphia, PA, 1994.

[43] C. Oara and A. Varga. Computation of general inner-outer and spectral factorization. *IEEE Trans. Automat. Control*, 45(12):2307–2325, December 2000.

[44] A. Odabasioglu, M. Celik, and L.T. Pallege. Prima: passive reduced-order interconnect macromodeling algorithm. In *Proceedings ICCAD'97*, pages 58–65, 1997.

[45] P.C. Opdenacker and E.A. Jonckheere. A contraction mapping preserving balanced reduction scheme and its infinity norm error bounds. *IEEE Trans. Circuits Syst.*, 35:184–189, 1988.

[46] P. A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, California Institute of Technology, May 2000.

[47] L. Pernebo and L.M. Silverman. Model reduction via balanced state-space representation. *IEEE Trans. Automat. Control*, 27(2):382–387, 1982.

[48] V. M. Popov. *Hyperstability of control systems*. Springer Verlag, Berlin, 1973. (Romanian version in 1966).

[49] Vl. Rasvan. Some system theory ideas connected with the stability problem. *Journal of Cybernetics*, 8:203–215, 1978.

[50] M. G. Safonov, E. A. Jonckheere, M. Verma, and D. J. D. Limebeer. Synthesis of positive real multivariable feedback systems. *Int. J. Control*, 45:817–842, 1987.

[51] A. S. Shiriaev. Some remarks on "systems analysis via intregral quadratic constraints". *IEEE Trans. Automat. Control*, 45(8):1527–1532, 2000.

[52] Y. C. Soh and L. Xie. Positive real control of linear uncertain systems. In *Proc. 3rd European Control Conference*, pages 1032–1036, Rome, Italy, 1995.

[53] W. Sun, P. P. Khargonekar, and D. Shim. Solution to the positive real control problem for linear time-invariant systems. *IEEE Trans. Automat. Control*, 39:2034–2046, 1994.

[54] P. Van Dooren. The generalized eigenstructure problem in linear system theory. *IEEE Trans. Automat. Control*, AC-26(1):111–129, 1981.

[55] H.-S. Wang and F.-R Chang. The generalized state-space description of positive realness and bounded realness. In *IEEE MSCS*, pages 893–896, 1997.

[56] J. Wang and L. Huang. Model reduction problem for some strictly positive real systems. *Int. J. Systems Sci.*, 30(12):1319–1324, 1999.

[57] J. T. Wen. Time domain and frequency domain conditions for positive realness. *IEEE Trans. Automat. Control*, 33:988–992, 1988.

[58] J. C. Willems. Least squares stationary optimal control and the algebraic Riccati equation. *IEEE Trans. Automat. Control*, AC-16(6):621–634, December 1971.

[59] J.C. Willems. Disspative dynamical systems. part i: General theory. *Arch. Rat. Mech. Anal.*, 45(1):321–351, dec 1972.

[60] J.C. Willems. Disspative dynamical systems. part ii: Linear systems with quadratic supply rates. *Arch. Rat. Mech. Anal.*, 45(1):352–393, dec 1972.

[61] L. Xie and Y. C. Soh. Positive real control for uncertain linear time-invariant systems. *Systems Control Lett.*, 24:265–271, 1995.

[62] V.A. Yakubovich. Solution of certain matrix inequalities in the stability theory of nonlinear control systems. *Dokl. Akad. Nauk. SSSR*, 143:1304–1307, 1962.

[63] V.A. Yakubovich. Frequency conditions for the absolute stability of control systems with several nonlinear or linear stationary units. *Automation Remote Control*, pages 5–30, 1967.

[64] D. Youla. On the factorization of rational matrices. *IRE Trans. Inform. Theory*, IT7:172–189, 1961.

[65] L. Zhang, J. Lam, and S. Xu. On positive realness of descriptor systems. *IEEE Trans. Circuits Syst. I*, 2000. under revision.