Journal of the Operations Research
Society of Japan
Vol. 43, No. 1, March 2000

OPTIMIZATION IN CONTROL SYSTEM DESIGN

Hisaya Fujioka Yuji Wakasa Yutaka Yamamoto
Kyoto University

(Received January 19, 1999; Revised May 17, 1999)

Abstract This paper reviews i) the background for the modern H^∞ control theory, and then ii) how it led to the modern optimization theory such as those using LMIs (linear matrix inequalities) and BMIs (bilinear matrix inequalities). Starting from the simplest sensitivity minimization problem, we give solutions via the Nevalinna-Pick interpolation and the Nehari theorem. The latter leads to a Riccati equation on which most of the H^∞ solutions are based. This in turn leads to the modern approach using mathematical programming such as LMIs and BMIs. Two types of global optimization algorithms to solve BMIs are introduced.

1. Introduction

The theory of feedback control dates back to the time of the industrial revolution, particularly that of the celebrated discovery of the principle of feedback in the introduction of a centrifugal governor to steam engines by J. Watt.

Before the invention of Watt, it was always troublesome to keep the rotational speed of steam engines constant against various load changes. When the speed becomes slower, the valve to the cylinder should be opened so that the engine picks up more power; on the other hand, if the load becomes lighter, the valve should be again adjusted to reduce the rotational speed. Such a change may be done manually, but will be considerably more difficult if the load change becomes very frequent and fast.

The obvious desire then is to do this job (i.e., *control*) automatically by some machine. The centrifugal governor introduced by Watt precisely did this job. It measures the current rotational speed, and *feed it back* to the valve opening, thereby controlling the overall speed to a constant. The important fact we should note here is that control systems such as steam engines are exposed to fluctuations/disturbances that are often unknown or unpredictable, such as load changes, and the control systems should take care of such undesirable uncertainty in the total system. Furthermore, in operating conditions, the *real* plants (systems) such as steam engines are exposed to many different kinds of fluctuations from the *ideal* situation in which the total system was designed. First of all, there are some parasitic effects introduced by friction, back-lash, etc. There are also disturbances coming from external conditions, winds, external temperature change, just to name a few. And, most of all, the model we work with are very far from complete. There are unmodeled dynamics everywhere. In a word, if we expect a control system to work in a *real* situation, it should be effective in reducing such *uncertainties*. The primary role of feedback lies precisely in this: it reduces *sensitivity* against such undesirable fluctuations.

To obtain a more elaborate response (such as tracking to various different set-points), one may have to increase the order of the controller. This is bound to introduce more delays in signal transmission in the feedback loop. This makes it more difficult to increase the *gain* of the controller, because it tends to induce *instability* of the overall system.

This stability problem captured the interest of Sir Maxwell, and he wrote perhaps the first paper on the theory of control: "On Governors" [28]. He could derive stability conditions for systems with order less than or equal to 3, but no more. The general case was later resolved independently by Routh and Hurwitz.

Great progress has been made since then. The notion of frequency response was found and greatly contributed to the steady-state analysis of the system characteristics. Various analysis and design methods were developed in 20's and 30's. Then came the modern theory based on state space methods since late 50's. The new state space framework (vs. frequency domain methods) provide more algebraic and computationally appealing tools, often based on optimal control problems. Such modern developments later culminated into a unification of frequency-domain and state-space methods through the modern H^∞ control theory, which is a more advanced min-max version of optimal control developed in 60's.

This paper reviews the background for the modern H^∞ control theory, and then goes over how it led to the modern optimization theory such as those using LMIs (linear matrix inequalities) and BMIs (bilinear matrix inequalities). Starting from the simplest sensitivity minimization problem, we give solutions via the Nevalinna-Pick interpolation and the Nehari theorem. The latter leads a Riccati equation on which most of the H^∞ solutions are based. This in turn leads to the modern approach using mathematical programming such as LMIs and BMIs. Two types of global optimization algorithms to solve BMIs are introduced.

The paper is organized as follows: Section 2 introduces some basic concepts and design problems in control engineering. Then, Section 3 reviews the modern H^∞ control theory. Section 4 introduces an LMI approach to the H^∞ control problem. Two types of global optimization algorithms to solve BMIs are provided in Section 5.

2. Preliminaries

As usual $H^p(\mathbb{C}_+)$ $(p \geq 1)$ denotes the Hardy p space on the right-half complex plane $\mathbb{C}_+ = \{s | \operatorname{Re} s > 0\}$. It consists of all functions f analytic on $\mathbb{C}_+$ such that

$$\sup_{\sigma>0} \int_{-\infty}^{\infty} |f(\sigma + j\omega)|^p d\omega < \infty. \tag{2.1}$$

When $p = \infty$, this condition is replaced by

$$\sup_{\sigma>0} |f(\sigma + j\omega)| < \infty. \tag{2.2}$$

It is a Banach space with norm

$$\|f\|_p := \left\{ \sup_{\sigma>0} \int_{-\infty}^{\infty} |f(\sigma + j\omega)|^p d\omega \right\}^{1/p}. \tag{2.3}$$

When $p = \infty$, this definition is accordingly modified in accordance with (2.2).

The cases $p = 2, \infty$ are particularly important. The space H^2 is the Laplace transform of functions in $L^2[0, \infty)$. H^∞ acts on H^2 via multiplication. Via inverse Laplace transformation, this corresponds to convolution operation, and elements in H^∞ gives a *stable* operator (see below). While H^2 is a Hilbert space, H^∞ is not. See [44] for details.

2.1. Linear time-invariant systems

Usually the systems we consider are *linear* and *time-invariant* whose relationship between input u and output y is described by

$$y(t) = \int_0^t g(t - \tau) u(\tau) d\tau. \tag{2.4}$$

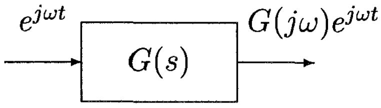

Figure 1: Steady-State Response

It is often convenient to represent this correspondence via Laplace transform as

$$\hat{y}(s) = G(s)\hat{u}(s). \tag{2.5}$$

where $G(s) = \hat{g}(s)$ and $\hat{}$ denotes the Laplace transform. $G(s)$ is called the *transfer function* of the system. We now make the following standing assumption:

Assumption 1 We assume that $G(s)$ is a proper (i.e., the degree of the numerator is less than or equal to that of the denominator) rational function.

Under this assumption, the following facts hold (e.g., [24]):

Facts:

- $G(s)$ admits a "realization" in the state space model:

$$\frac{d}{dt}x(t) = Ax(t) + Bu(t) \tag{2.6}$$
$$y(t) = Cx(t) + Du(t) \tag{2.7}$$

  Under this notation, $G(s) = D + C(sI - A)^{-1}B$. This can be ensured by taking the Laplace transforms of the both sides of (2.6) and (2.7) and comparing the results with (2.5). A realization is said to be *minimal* if it assumes the smallest dimension in the state space among all realizations. In what follows, we basically assume that the realizations are minimal. Note that when a realization (A, B, C, D) is minimal, then the spectrum of A coincides with the poles of $G(s)$.
- The system is stable $\Leftrightarrow \sigma(A) \subset \mathbb{C}_- \Leftrightarrow G(s)$ has no poles in $\bar{\mathbb{C}}_+ \Leftrightarrow G(s) \in H^\infty(\mathbb{C}_+)$, where $\mathbb{C}_-$ and $\bar{\mathbb{C}}_+$ denote the open left-half and the closed right-half complex planes, respectively.

2.2. Steady-state and frequency response

The asymptotic behavior of the system $G(s) \in H^\infty(\mathbb{C}_+)$ against sinusoidal inputs $\sin\omega t$ (or $e^{j\omega t}$) represents the system characteristic very well. The very fundamental fact is that *in the steady state, a linear time-invariant system synchronizes the input, i.e., it produces an output with the same frequency as that of the input.*

To see this, let $e^{j\omega t}$ be the input and expand $G(s)$ around $j\omega$ to get

$$G(s) = G(j\omega) + G_1(s)(s - j\omega).$$

Applying the (unilateral) Laplace transform $1/(s - j\omega)$ of $e^{j\omega t}$, we get

$$G(s)\frac{1}{s - j\omega} = G(j\omega)\frac{1}{s - j\omega} + G_1(s).$$

The second term belongs to H^2 and decays to zero in the time domain (it is stable). So the remaining steady-state term is

$$G(j\omega)\frac{1}{s - j\omega}$$

with inverse Laplace transform

$$G(j\omega)e^{j\omega t}.$$

This means that the asymptotic response against a sinusoid is the same sinusoid with

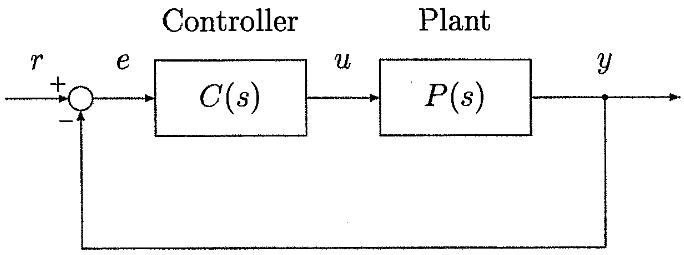

Figure 2: Unity Feedback System

- gain change by $|G(j\omega)|$
- phase shift by $\angle G(j\omega)$.

The function: $\omega \mapsto G(j\omega)$ is called the *frequency response* and is a fundamental tool for control systems. The plot of this function against each frequency ω (most often in logarithmic scale) is called the *Bode plot.*

2.3. Design problems

Now consider the unity feedback system Figure 2. The correspondence from r to e is described by the *sensitivity function* $S(s) = (1 + p(s)c(s))^{-1}$ as

$$\hat{e}(s) = S(s)\hat{r}(s).$$

Now one of the control objectives is to reduce this sensitivity (making the system low sensitive to plant, noise fluctuations). Unfortunately, it is known that it is not possible to reduce the sensitivity over all frequency range (Bode's sensitivity integral theorem). We usually make the criterion *frequency selective*, by introducing a suitable weighting function $W(s)$ and reduce $W(s)S(s)$ instead of $S(s)$. A classical solution to this is given for the H^2 norm case:

$$J = \inf_{c(s):\text{stabilizing}} \|W(s)S(s)\|_2. \tag{2.8}$$

This is essentially a minimum distance problem in Hilbert space, and its solution reduces to the projection theorem. Beautiful connections with state space algorithms (Kalman filtering), Hamilton-Jacobi theory, Riccati equations, controllability/observability notions have been known.

On the other hand, the above performance is only in the mean square sense, and is not quite suitable for elaborate design. It is certainly more desirable to solve a problem like

$$\gamma_{opt} = \inf_{c(s):\text{stabilizing}} \|W(s)S(s)\|_\infty. \tag{2.9}$$

A difficulty is, of course, that H^∞ is not a Hilbert space and the projection theorem cannot be used. This is a very rudimentary form of the more modern approach, so known as H^∞ control.

3. H^∞ control

Around 1980, George Zames introduced this optimization problem to the modern control [45]. In a sense, the frame is surely more amenable to the classical control engineering, consisting of graphical shaping of Bode plots. On the other hand, no one at the time knew how it could be solved; moreover, it was not known whether it was solvable at all. Since it does not allow a Hilbert space structure (and not quadratic structures), there was a fair amount of skepticism.

However, around 1982–3, a general solution started to take shape. Zames and his co-workers (particularly Bruce Francis) found a connection with Sarason's generalized interpolation [46, 10]. Later the connection was tied together with Nevalinna-Pick theory and Nehari's approximation theorem.

Let us restate our problem. We have to find

$$\inf_{c(s):\text{stabilizing}} \|W(s)(1+p(s)c(s))^{-1}\|_\infty.$$

For simplicity, let us assume $p(s)$ to be stable, and introduce a new parameter $q(s) = (1+p(s)c(s))^{-1}c(s)$. Then

- closed-loop stability $\Leftrightarrow q(s) \in H^\infty$
- $S(s) = 1 - p(s)q(s)$.

In this case, we need not worry about the internal stability. We just have to take stable q. Note however that the correspondence $q \leftrightarrow c$ is nonlinear.

So the problem is reduced to finding

$$\inf_{q\in H^\infty} \|W - Wpq\|_\infty. \tag{3.1}$$

Bring in the inner-outer factorization of Wp:

$$W(s)p(s) = v_o(s)v_i(s).$$

Clearly the outer factor $v_o(s)$ can be absorbed into q, so we do not need this term. Writing $m := v_i(s)$, we see that (3.1) is rewritten as

$$\inf_{q\in H^\infty} \|W - mq\|_\infty. \tag{3.2}$$

This is the *sensitivity minimization problem.*

Remark 3.1

- More general problems have been solved. Of particular importance is the so-called *mixed sensitivity problem* which intends to minimize

$$\left\| \begin{bmatrix} W_1 S \\ W_2 T \end{bmatrix} \right\|_\infty$$

  where $T = 1 - S$ is the *complementary sensitivity function.* This problem appears in sensitivity minimization while maintaining *robust stability* against model uncertainties, plant fluctuations.
- Since the problem above attempts to minimize the H^∞ norm of a two block transfer function, it is often called a *two block problem.* In contrast, the sensitivity minimization problem (3.2) is referred to as the *one block problem.*
- When $p(s)$ is not stable, the *Q parameterization* $q = (1+pc)^{-1}c$ does not work. This can be taken care of the so-called *Youla parameterization* that makes use of coprime factorization of $p(s)$ over H^∞.

Various solutions are available for our sensitivity minimization problem. In particular, a solution via two Riccati equations give a general solution to the *four block* problem, and is conveniently packaged into a CAD program (MATLAB and other commercial CAD programs). The design has successfully been applied to many advanced control problems, and has proven to be superior to more conventional LQ (H^2) designs.

We here indicate two different (but equivalent) solutions which played historical roles. One is via the Nevalinna-Pick interpolation problem and the other via Nehari's function approximation in H^∞.

3.1. Nevalinna-Pick solution

Recall that we wish to find the minimum model-matching error γ_{opt} among all γ such that

$$\|W - mq\|_\infty \leq \gamma.$$

The following solution follows the treatment given in [4]. Fix $\gamma > 0$ and consider

$$G := \frac{1}{\gamma}(W - mq).$$

If q is stable (i.e., in H^∞) so is G, but not conversely. Certain interpolation conditions must be satisfied in order that q be stable. To see this, let $\{s_1, \ldots, s_n\}$ be the unstable zeros of m, i.e., Re $s_i > 0$. Clearly the interpolation conditions

$$G(s_i) = \frac{1}{\gamma}W(s_i), \quad i = 1, \ldots, n$$

must be satisfied. Conversely, if $G \in H^\infty$ satisfies these conditions, then

$$q = m^{-1}(W - \gamma G)$$

belongs to H^∞ because all the unstable *poles* s_i arising from m^{-1} will be canceled by the numerator $W - \gamma G$. Therefore, γ_{opt} is the minimum of γ such that

$$\begin{aligned} G(s_i) &= \frac{1}{\gamma}W(s_i), \quad i = 1, \ldots, n \\ \|G\|_\infty &\leq 1 \end{aligned}$$

for some H^∞ function G. This is precisely the Nevalinna-Pick problem with interpolation data

$$\begin{array}{ccc} s_1 & \cdots & s_n \\ \updownarrow & \cdots & \updownarrow \\ W(s_1)/\gamma & \cdots & W(s_n)/\gamma. \end{array}$$

Therefore, γ satisfies the conditions above if and only if the Pick matrix

$$\left(\frac{1 - W(s_i)\overline{W(s_j)}/\gamma^2}{s_i + \overline{s_j}} \right) \geq 0.$$

The minimum γ can be found as the minimum of all such γ.

Remark 3.2 The optimal q is given by

$$q := \frac{W - \gamma_{opt} G}{m}$$

Note that $q \in H^\infty$.

3.2. Nehari's theorem

Another solution was given by Nehari's theorem. Let us review this solution following [9]. It goes as follows: Multiply $m\tilde{}(s) = m^{-1}(s)$ in (3.2) to get

$$\inf_{q\in H^\infty} \|m\tilde{}W - q\|_\infty, \tag{3.3}$$

where the norm is taken in $L^\infty(-j\infty, j\infty)$. Decompose $m\tilde{}W$ as

$$m\tilde{}W = W_1 + W_2$$

where W_1 is stable and W_2 is strictly proper and anti-stable (i.e., all poles in the open right-half complex plane). Nehari's theorem [31] tells us that

$$\gamma_{opt} = \|\Gamma_{W_2}\|$$

where Γ_{W_2} is the Hankel operator associated with W_2:

$$\Gamma_{W_2} : H^2 \to H^2_- : x \mapsto \Pi_- M_{W_2} x \tag{3.4}$$

and $\Pi_- : L^2(-j\infty, j\infty) \to H^2$ is the canonical projection and $M_f : L^2(-j\infty, j\infty) \to L^2(-j\infty, j\infty)$ is the multiplication operator with symbol $f \in L^\infty(-j\infty, j\infty)$.

To execute the computation, we should bring in the minimal realization (A, B, C) of W_2 as $W_2(s) = C(sI - A)^{-1}B$. The inverse bilateral Laplace transform of $W_2(s)$ is given by

$$\begin{aligned} f(t) &= -Ce^{At}B, \quad t < 0 \\ f(t) &= 0, \quad t \geq 0. \end{aligned}$$

Then in the time-domain the Hankel operator is represented as the correspondence from $u \in L^2[0, \infty)$ to $y \in L^2(-\infty, 0]$:

$$y(t) = \int_0^\infty f(t-\tau)u(\tau)d\tau = -Ce^{At}\int_0^\infty e^{-A\tau}Bu(\tau)d\tau, \quad t < 0.$$

Define

$$\begin{aligned} \Psi_c : L^2[0,\infty) \to \mathbb{C}^n : &\quad \Psi_c u := -\int_0^\infty e^{-A\tau}Bu(\tau)d\tau \\ \Psi_o : \mathbb{C}^n \to L^2(-\infty, 0] : &\quad (\Psi_o x)(t) := Ce^{At}x, t < 0. \end{aligned}$$

Clearly $\Gamma_f = \Psi_o\Psi_c$.

The norm of Γ_f is generally given as the maximal singular value, so one has to solve the eigenvalue equation

$$\Psi_c^*\Psi_o^*\Psi_o\Psi_c u = \lambda u. \tag{3.5}$$

A nonzero eigenvalue λ here is equal to that of

$$\Psi_c\Psi_c^*\Psi_o^*\Psi_o x = \lambda x. \tag{3.6}$$

It is routine to see that

$$\begin{aligned} \Psi_c\Psi_c^* &= \int_0^\infty e^{-At}BB^Te^{A^Tt}dt \\ \Psi_o^*\Psi_o &= \int_0^\infty e^{-A^Tt}C^TCe^{At}dt. \end{aligned}$$

In contrast to (3.5), (3.6) is a finite-dimensional eigenvalue problem, and the pertinent matrices

$$L_c := \Psi_c \Psi_c^*, \quad L_o := \Psi_o^* \Psi_o$$

are *unique* solutions of the Lyapunov equations:

$$\begin{aligned} AL_c + L_c A^T &= BB^T && (3.7)\\ A^T L_o + L_o A &= C^T C. && (3.8) \end{aligned}$$

This completes our treatment of the Nehari-type solution.

3.3. Relationship with Riccati equations

Multiply $X = L_c^{-1}$ from both left and right in (3.7). Then we obtain

$$XA + A^T X = XBB^T X.$$

This is a special case of the so-called Riccati equation

$$XA + A^T X + XR^{-1}X + Q = 0, \quad Q \geq 0$$

(except the difference of the sign in the quadratic term).

Q is often a free design parameter, and can be replaced by another Q. We can convert this equation to the matrix Linear Matrix Inequality

$$\begin{bmatrix} -XA - A^T X - Q & X \\ X & R \end{bmatrix} > 0.$$

because by the so-called Schur complement [3],

$$R > 0 \text{ and } A - BR^{-1}B^T > 0$$

is equivalent to

$$\begin{bmatrix} A & B \\ B^T & R \end{bmatrix} > 0.$$

The latter is the so-called *Linear Matrix Inequality* (LMI) and can be solved effectively via various convex optimization techniques. This will be discussed in more detail in the next section.

Preceding to such developments was Boyd's contribution that a variety of control problems can be reduced to convex optimization problems, via nonlinear variable changes. This may give the impression that such problems have relatively limited use, since convex problems are known to be rare in some contexts. However, the crux here is the variable transformation. For example, the H^∞ norm bound problem

$$\|(1 - c(s)p(s))^{-1}\|_\infty < \gamma$$

is not convex with respect to $c(s)$, but *is* convex with respect to the closed-loop transfer function $S = (1 - c(s)p(s))^{-1}$. Hence once a solution S is obtained, it can be converted back to that of $c(s)$ by solving $S(s) = (1 - c(s)p(s))^{-1}$ for $c(s)$. For some details, the reader is referred to [2].

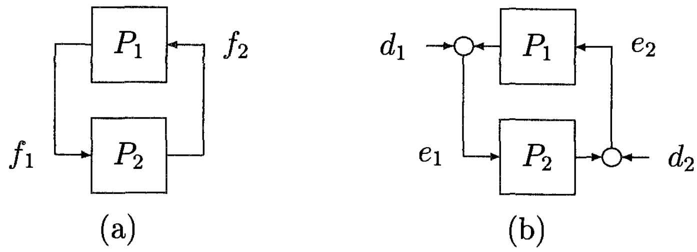

Figure 3: Internal Stability

4. LMI Approach to H^∞ Control

Let us now delve more into numerical optimization. In the past ten years, it has been shown that we can solve a wide variety of control problems by using solutions to LMIs [3, 20, 35]. LMIs are solvable effectively based on the interior point method [32] with available softwares (e.g., [14]), and hence we can obtain solutions to many control problems. Among them, the following problems are hard to solve analytically, but solvable by solving corresponding LMIs numerically: multi-objective control [27, 38], robust stability analysis [29, 47], gain scheduling control [34], state-feedback scaled H^∞ control [36].

In this section, we will show how to cast the H^∞ control problem into LMI framework [15, 22].

4.1. Internal stability

To formulate the problem, we need to define the notion of *internal stability* of feedback systems.

Consider an FDLTI (finite dimensional linear time-invariant) system P:

$$P : \begin{bmatrix} \dot{x}(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}.$$

We say that P is internally stable if the solution $x(t)$ of

$$\dot{x}(t) = Ax(t)$$

approaches to 0 when t goes to ∞ for any bounded initial state $x(0)$. Generalizing this notion of stability, the internal stability of feedback systems is defined as follows:

Definition 1 (Internal Stability) The feedback system in Figure 3 (a) composed of two systems P_1 and P_2 is *internally stable* if the map $d \mapsto e$ in Figure 3 (b) is L^2-bounded, i.e., bounded when the input and output spaces are endowed with the L^2 norm, where

$$d(t) := \begin{bmatrix} d_1(t) \\ d_2(t) \end{bmatrix}, \quad e(t) := \begin{bmatrix} e_1(t) \\ e_2(t) \end{bmatrix}.$$

Remark 1 Suppose that both P_1 and P_2 are FDLTI systems:

$$P_1 : \begin{bmatrix} \dot{x}_1(t) \\ f_1(t) \end{bmatrix} = \begin{bmatrix} A_1 & B_1 \\ C_1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ f_2(t) \end{bmatrix}, \quad P_2 : \begin{bmatrix} \dot{x}_2(t) \\ f_2(t) \end{bmatrix} = \begin{bmatrix} A_2 & B_2 \\ C_2 & 0 \end{bmatrix} \begin{bmatrix} x_2(t) \\ f_1(t) \end{bmatrix}.$$

Then the dynamics of the closed-loop system is given by

$$\dot{x}_{c\ell}(t) = \begin{bmatrix} A_1 & B_1C_2 \\ B_2C_1 & A_2 \end{bmatrix} x_{c\ell}(t); \quad x_{c\ell}(t) := \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}.$$

We see that the map $d \mapsto e$ in Figure 3 (b) is L^2-bounded if and only if the state $x_{c\ell}(t)$ approaches to 0 when t goes to ∞ for any bounded initial state $x_{c\ell}(0)$, assuming the stabilizabilities and the detectabilities of P_1 and P_2.

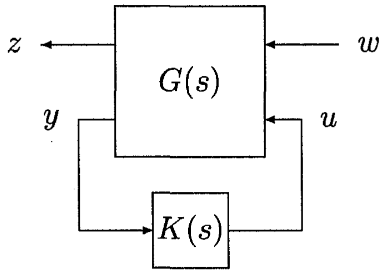

Figure 4: H^∞ Control Problem

4.2. Problem formulation

The general setup of the H^∞ control problem is depicted in Figure 4, where $G(s)$ is the transfer function of the *generalized plant*, and $K(s)$ denotes the transfer function of the controller to be designed. The problem is formulated as follows [5]:

Problem 1 (H^∞ Control Problem) For given $G(s)$, find $K(s)$ satisfying the following two specifications:

(i) The closed-loop system is internally stable.

(ii) $\|T(s)\|_\infty < 1$, where $T(s)$ is the transfer function from w to z:

$$T(s) := G_{11}(s) + G_{12}(s)K(s)\left(I - G_{22}(s)K(s)\right)^{-1} G_{21}(s). \tag{4.1}$$

Here $G(s)$ is conformably partitioned to satisfy

$$\begin{bmatrix} z(s) \\ y(s) \end{bmatrix} = \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix} \begin{bmatrix} w(s) \\ u(s) \end{bmatrix}.$$

The following are typical examples of the H^∞ control problem:

Example 1 (Sensitivity Minimization Problem) Consider the feedback system depicted in Figure 2. The sensitivity minimization problem stated in the previous section is to minimize γ such that the H^∞ control problem has a solution for

$$G(s) = \begin{bmatrix} \gamma^{-1}W(s) & -\gamma^{-1}W(s)P(s) \\ I & -P(s) \end{bmatrix}.$$

Example 2 (Robust Stabilization Problem) Consider the feedback system depicted in Figure 2. Let $\mathcal{P}$ denote a set of transfer functions of the same size. We say that $K(s)$ *robustly stabilizes* the closed-loop system if the closed-loop system is internally stable for all $P(s) \in \mathcal{P}$. The robust stabilization problem is to find $K(s)$ such that the closed-loop system is robustly stabilized for given $\mathcal{P}$.

Consider the case where $\mathcal{P}$ is defined by

$$\mathcal{P} = \{\, P(s) : P(s) = P_0(s) + \Delta(s)W(s); \quad \|\Delta(s)\|_\infty \le 1\}. \tag{4.2}$$

The robust stabilization problem for $\mathcal{P}$ in (4.2) is equivalent to the H^∞ control problem for

$$G(s) = \begin{bmatrix} 0 & W(s) \\ -I & -P(s) \end{bmatrix}$$

under mild conditions.

4.3. Solution in terms of LMIs

Suppose that a state-space realization of $G(s)$ of order n is given by

$$G: \begin{bmatrix} \dot{x}(t) \\ z(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} A & B_1 & B_2 \\ C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ w(t) \\ u(t) \end{bmatrix}; \quad x(t) \in \Re^n. \tag{4.3}$$

and a state-space realization of $K(s)$ is also given by

$$K: \begin{bmatrix} \dot{x}_K(t) \\ u(t) \end{bmatrix} = M_K \begin{bmatrix} x_K(t) \\ y(t) \end{bmatrix}; \quad M_K := \begin{bmatrix} A_K & B_K \\ C_K & D_K \end{bmatrix}; \quad x_K(t) \in \Re^{n_K} \tag{4.4}$$

where the order of $K(s)$ is n_K. We assume that (A, B_2) and (A, C_2) are stabilizable and detectable, respectively.

The following is a version of the bounded real lemma (e.g., [3, 20]), which plays a key role in the LMI approach to the H^∞ control problem:

Lemma 1 (Bounded Real Lemma) Let $T(s) := C_{c\ell}(sI_{n_{c\ell}} - A_{c\ell})^{-1}B_{c\ell} + D_{c\ell}$. The following statements are equivalent:

(i) $\sigma(A) \subset \mathbb{C}_-$ and $\|T(s)\|_\infty < 1$.

(ii) There exists $X_{c\ell} = X_{c\ell}^T \in \Re^{n_{c\ell} \times n_{c\ell}}$, $X_{c\ell} > 0$ satisfying

$$\begin{bmatrix} A_{c\ell}X_{c\ell} + X_{c\ell}A_{c\ell}^T & X_{c\ell}C_{c\ell}^T & B_{c\ell} \\ C_{c\ell}X_{c\ell} & -I & D_{c\ell} \\ B_{c\ell}^T & D_{c\ell}^T & -I \end{bmatrix} < 0. \tag{4.5}$$

Remark 2 For a fixed controller $K(s)$, the state-space data of $T(s)$ are constant and hence (4.5) is an LMI in $X_{c\ell}$. Thus we can check the stability and the norm-bound condition of $T(s)$ effectively by solving (4.5).

Taking $x_{c\ell}$ as

$$x_{c\ell}(t) = \begin{bmatrix} x(t) \\ x_K(t) \end{bmatrix}, \quad n_{c\ell} = n + n_K,$$

$A_{c\ell}$, $B_{c\ell}$, $C_{c\ell}$, and $D_{c\ell}$ are given by

$$\begin{bmatrix} A_{c\ell} & B_{c\ell} \\ C_{c\ell} & D_{c\ell} \end{bmatrix} = \left[\begin{array}{cc|c} A & 0 & B_1 \\ 0 & 0 & 0 \\ \hline C_1 & 0 & 0 \end{array}\right] + \left[\begin{array}{cc} 0 & B_2 \\ I & 0 \\ \hline 0 & D_{12} \end{array}\right] M_K \left[\begin{array}{cc|c} 0 & I & 0 \\ C_2 & 0 & D_{21} \end{array}\right]. \tag{4.6}$$

Lemma 1 implies that the following are equivalent:

(i) There exists a solution to the H^∞ control problem $K(s)$ of order n_K.

(ii) There exists a pair of matrices $(X_{c\ell}, M_K)$ that satisfies (4.5).

The state-space data of $T(s)$ affinely depend on M_K as shown in (4.6), hence (4.5) is not an LMI in $(M_K, X_{c\ell})$. We can however reduce (4.5) into a condition in terms of LMIs preserving the solvability [15, 22].

Let us introduce the following notation: For a given matrix $M \in \Re^{n \times m}$, $M^\perp$ is a matrix satisfying $M^\perp \in \Re^{(n-r) \times n}$, $M^\perp M = 0$, and $M^\perp (M^\perp)^T > 0$, where $r := \operatorname{rank} M$. Then we have the following theorem:

Theorem 1 [15, 22] For given $G(s)$ in (4.3), the following are equivalent:

(i) There exists a solution to the H^∞ control problem $K(s)$.

(ii) There exists a pair of matrices $(X, Y) \in \Re^n \times \Re^n$, $X = X^T > 0$, $Y = Y^T > 0$ satisfying

$$\begin{bmatrix} B_2 \\ D_{12} \end{bmatrix}^{\perp} \begin{bmatrix} AX + XA^T + B_1 B_1^T & XC_1^T \\ C_1 X & -I \end{bmatrix} \left(\begin{bmatrix} B_2 \\ D_{12} \end{bmatrix}^{\perp} \right)^T < 0, \tag{4.7}$$

$$\begin{bmatrix} C_2^T \\ D_{21}^T \end{bmatrix}^{\perp} \begin{bmatrix} A^T Y + YA + C_1^T C_1 & YB_1 \\ B_1^T Y & -I \end{bmatrix} \left(\begin{bmatrix} C_2^T \\ D_{21}^T \end{bmatrix}^{\perp} \right)^T < 0, \tag{4.8}$$

and

$$\begin{bmatrix} X & I \\ I & Y \end{bmatrix} > 0. \tag{4.9}$$

Proof: We assume that $n_K = n$ without loss of generality, since there exists a solution $K(s)$ of order n if there exists a solution [5]. In this case, we can take $X_{c\ell}$ as

$$X_{c\ell} = \begin{bmatrix} X & Z \\ Z & Z \end{bmatrix} \tag{4.10}$$

without loss of generality [27, 38], where $X \in \Re^{n \times n}$, $X = X^T > 0$, $Z \in \Re^{n \times n}$, $Z = Z^T > 0$, and

$$X - Z > 0.$$

We will show that condition (ii) in Lemma 1 is equivalent to condition (ii) in Theorem 1. Substituting (4.6) and (4.10) into (4.5), we have

$$\tilde{B} M_K \tilde{C} + (\tilde{B} M_K \tilde{C})^T + \tilde{Q} < 0 \tag{4.11}$$

where

$$\tilde{B} := \begin{bmatrix} 0 & B_2 \\ I & 0 \\ 0 & D_{12} \\ 0 & 0 \end{bmatrix}, \quad \tilde{C} := \begin{bmatrix} Z & Z & 0 & 0 \\ C_2 X & C_2 Z & 0 & D_{21} \end{bmatrix}, \quad \tilde{Q} := \begin{bmatrix} AX + XA^T & AZ & XC_1^T & B_1 \\ ZA^T & 0 & ZC_1^T & 0 \\ C_1 X & C_1 Z & -I & 0 \\ B_1^T & 0 & 0 & -I \end{bmatrix}.$$

Inequality (4.11) has a solution M_K *if and only if* the following conditions hold [15, 22]:

$$\tilde{B}^{\perp} \tilde{Q} (\tilde{B}^{\perp})^T < 0. \tag{4.12}$$

$$(\tilde{C}^T)^{\perp} \tilde{Q} ((\tilde{C}^T)^{\perp})^T < 0. \tag{4.13}$$

Noting that

$$\tilde{B}^{\perp} = \begin{bmatrix} \begin{bmatrix} B_2 \\ D_{12} \end{bmatrix}^{\perp} & 0 \\ 0 & I \end{bmatrix} \left[\begin{array}{cccc} I & 0 & 0 & 0 \\ 0 & 0 & I & 0 \\ \hline 0 & 0 & 0 & I \end{array} \right],$$

we see that inequality (4.12) becomes

$$\begin{bmatrix} \begin{bmatrix} B_2 \\ D_{12} \end{bmatrix}^{\perp} & 0 \\ 0 & I \end{bmatrix} \left[\begin{array}{cc|c} AX + XA^T & XC_1^T & B_1 \\ C_1 X & -I & 0 \\ \hline B_1^T & 0 & -I \end{array} \right] \begin{bmatrix} \begin{bmatrix} B_2 \\ D_{12} \end{bmatrix}^{\perp} & 0 \\ 0 & I \end{bmatrix}^T < 0.$$

Invoking the Schur complement formula [3], (4.7) follows.

Similarly, we have

$$(\tilde{C}^T)^\perp = \begin{bmatrix} \begin{bmatrix} C_2^T \\ D_{21}^T \end{bmatrix}^\perp & 0 \\ 0 & I \end{bmatrix} \left[\begin{array}{cc|c|c} I & 0 & 0 & 0 \\ 0 & 0 & 0 & I \\ \hline 0 & 0 & I & 0 \end{array}\right] \begin{bmatrix} \begin{bmatrix} X & Z \\ Z & Z \end{bmatrix}^{-1} & \begin{matrix} 0 & 0 \end{matrix} \\ \begin{matrix} 0 \\ 0 \end{matrix} & \begin{matrix} I & 0 \\ 0 & I \end{matrix} \end{bmatrix}$$

and

$$\begin{bmatrix} X & Z \\ Z & Z \end{bmatrix}^{-1} = \begin{bmatrix} Y & -Y \\ -Y & Y + Z^{-1} \end{bmatrix}$$

where $Y \in \Re^{n \times n}$ is defined by

$$Y := (X - Z)^{-1} > 0.$$

Noting that

$$\begin{aligned} \tilde{Q} &= \begin{bmatrix} X & Z \\ Z & Z \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A^T & 0 & C_1^T & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ &\quad + \left(\begin{bmatrix} X & Z \\ Z & Z \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A^T & 0 & C_1^T & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right)^T + \begin{bmatrix} 0 & 0 & 0 & B_1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -I & 0 \\ B_1^T & 0 & 0 & -I \end{bmatrix}, \end{aligned}$$

we also see inequality (4.13) turns to

$$\begin{bmatrix} \begin{bmatrix} C_2^T \\ D_{21}^T \end{bmatrix}^\perp & 0 \\ 0 & I \end{bmatrix} \left[\begin{array}{cc|c} A^TY + YA & YB_1 & C_1^T \\ B_1^TY & -I & 0 \\ \hline C_1 & 0 & -I \end{array}\right] \begin{bmatrix} \begin{bmatrix} C_2^T \\ D_{21}^T \end{bmatrix}^\perp & 0 \\ 0 & I \end{bmatrix}^T < 0.$$

Invoking the Schur complement formula again, we obtain (4.8).

Finally, we can treat X and Y as independent variables by adding a condition

$$X - Y^{-1} > 0$$

which is equivalent to $Z > 0$. Again by the Schur complement formula, (4.9) follows. This completes the proof. ■

Parameterizations of all M_K's satisfying (4.5) in terms of $X_{c\ell}$ are found in [27, 38].

5. Bilinear Matrix Inequality

In this section, we focus our attention on the BMI (bilinear matrix inequality) which is considered as one of the most flexible frameworks for control system design [12, 13, 37].

In the past five years, several researchers in the field of control engineering have tried to develop algorithms to solve BMIs [1, 11, 17, 18, 25, 26, 40, 42], and we currently have several types of local and global optimization algorithms to solve BMIs. However, the performance of global optimization algorithms are not satisfactory to be applied control system design of practical size. Hence, further improvement is still required.

In what follows, two types of global optimization algorithms are introduced after stating motivations and the formulation of the problem.

5.1. Motivations

To motivate the study of BMIs, consider the H^∞ control problem again.

We have seen analytical and numerical solutions to the H^∞ control problem. Here we consider the H^∞ control problem with an additional constraint on the order of the solution, requiring a simpler structure of the controller: For given $G(s)$ of order n and $n_\ell < n$, find $K(s)$ of order n_ℓ satisfying the specifications (i) and (ii) in Problem 1. This problem of low-order H^∞ controller synthesis is so hard that no analytical solution is known to date.

Based on (4.5), we can obtain a low-order H^∞ controller if we can find a pair $(X_{c\ell}, M_K)$ for the case of $n_K = n_\ell$ numerically. We call (4.5) as a BMI in $(X_{c\ell}, M_K)$, noting that $A_{c\ell}$, $B_{c\ell}$, $C_{c\ell}$, and $D_{c\ell}$ depend affinely on M_K as shown in (4.6). In the previous section, we have seen that we can cast (4.5) into conditions in terms of LMIs for the case $n_K = n$. However, we cannot apply techniques in the proof of Theorem 1 to the case of $n_K = n_\ell$, and it is an open question whether we can cast (4.5) into LMI conditions for the case of $n_K = n_\ell$. The study of BMIs intends to develop a numerical algorithm to solve (4.5) directly.

The motivation for the study of BMIs is not restricted to the low-order controller synthesis [39]. We can characterize a large number of problems that are considered to be hard in the field of control engineering by BMIs. The following is a part of the list of such problems: robust performance synthesis [47], multi-objective control synthesis [38], distributed control synthesis [19], simultaneous optimization of control and structure [33], just to name a few.

Remark 3 Although (4.5) is in fact a BMI in $(X_{c\ell}, M_K)$, we can reformulate it as other types of optimization problems [21]: a rank minimization problem [30], a matrix product eigenvalue problem [43], etc. Several local and global optimization algorithms are available for them.

5.2. BMI problem formulation

Now we formulate the BMI problem concretely.

Let $x \in \Re^n$, $y \in \Re^m$, and $z \in \Re^\ell$ be variable vectors. Let also $\mathcal{A}$: $\Re^n \times \Re^m \times \Re^\ell \to \mathcal{S}_p$ and $\mathcal{C}$: $\Re^\ell \to \mathcal{S}_q$ be affine maps, where p and q denote positive integers and $\mathcal{S}_p$ denotes a set of $p \times p$ real symmetric matrices. Let $\mathcal{B}$: $\Re^n \times \Re^m \to \mathcal{S}_q$ denote a bilinear map defined by

$$\mathcal{B}(x, y) := \sum_{i=1}^{n} B_{i0}x_i + \sum_{j=1}^{m} B_{0j}y_j + \sum_{i=1}^{n}\sum_{j=1}^{m} B_{ij}x_iy_j. \tag{5.1}$$

where coefficient matrices are all symmetric:

$$B_{i0} = B_{i0}^T, \quad B_{0j} = B_{0j}^T, \quad B_{ij} = B_{ij}^T.$$

We also assume that

$$\sum_{k=1}^{m} \|B_{ik}\| \neq 0, \quad \sum_{k=1}^{n} \|B_{kj}\| \neq 0$$

for all $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m\}$. Finally let $X \subset \Re^n$ and $Y \subset \Re^m$ denote bounded hyper-rectangles:

$$X := [L_{x_1}, U_{x_1}] \times \cdots \times [L_{x_n}, U_{x_n}], \quad Y := [L_{y_1}, U_{y_1}] \times \cdots \times [L_{y_m}, U_{y_m}], \tag{5.2}$$

respectively, where

$$|L_{x_i}| < \infty, \quad |U_{x_i}| < \infty, \quad |L_{y_j}| < \infty, \quad |U_{y_j}| < \infty.$$

The original objective is to find a solution $(x, y, z) \in X \times Y \times \Re^\ell$ to a BMI:

$$\begin{bmatrix} \mathcal{A}(x, y, z) & 0 \\ 0 & \mathcal{B}(x, y) + \mathcal{C}(z) \end{bmatrix} \geq 0. \tag{5.3}$$

Instead of solving (5.3) directly, we will solve the following optimization problem:
Problem 2 (OP) Compute $J(X \times Y)$ defined by

$$J(X \times Y) := \min_{(x, y, z)} \lambda$$

subject to

$$\mathcal{F}(x, y, z, \lambda) := \begin{bmatrix} \mathcal{A}(x, y, z) & 0 & 0 & 0 \\ 0 & \lambda I + \mathcal{B}(x, y) + \mathcal{C}(z) & 0 & 0 \\ 0 & 0 & \mathcal{D}_x(x; X) & 0 \\ 0 & 0 & 0 & \mathcal{D}_y(y; Y) \end{bmatrix} \geq 0 \tag{5.4}$$

and find the minimizer (x^*, y^*, z^*), where $\mathcal{D}_x$ and $\mathcal{D}_y$ are defined by

$$\mathcal{D}_x(x; X) := \mathrm{diag}\,(x_1 - L_{x_1}, -x_1 + U_{x_1}, \ldots, x_n - L_{x_n}, -x_n + U_{x_n}).$$

$$\mathcal{D}_y(y; Y) := \mathrm{diag}\,(y_1 - L_{y_1}, -y_1 + U_{y_1}, \ldots, y_m - L_{y_m}, -y_m + U_{y_m}).$$

Note that BMI (5.3) has a solution $(x, y, z) \in X \times Y \times \Re^\ell$ if and only if the optimal value of (OP) is $\lambda \leq 0$.
Remark 4 Since $\mathcal{B}$ is a bilinear map, (OP) is a non-convex optimization problem. In fact, the problem of checking the solvability of a BMI is $\mathcal{NP}$-hard [41].

In the following subsections, two types of global algorithms for (OP) are provided.

5.3. Branch and bound algorithm

This subsection provides a version of the branch and bound algorithm for (OP).

5.3.1. Semidefinite relaxation

In order to solve (OP) with a type of the branch and bound algorithm, we need to derive lower and upper bounds of the objective function.

To derive a lower bound of $J(X \times Y)$, it would be natural to introduce a new variable $W \in \Re^{n \times m}$ and linearize (5.4) by

$$\begin{bmatrix} \mathcal{A}(x, y, z) & 0 & 0 & 0 \\ 0 & \lambda I + \mathcal{R}(x, y, W) + \mathcal{C}(z) & 0 & 0 \\ 0 & 0 & \mathcal{D}_x(x; X) & 0 \\ 0 & 0 & 0 & \mathcal{D}_y(y; Y) \end{bmatrix} \geq 0 \tag{5.5}$$

where $\mathcal{R}$: $\Re^n \times \Re^m \times \Re^{n \times m} \to \mathcal{S}_q$ is a linear map defined by

$$\mathcal{R}(x, y, W) := \sum_{i=1}^{n} B_{i0} x_i + \sum_{j=1}^{m} B_{0j} y_j + \sum_{i=1}^{n} \sum_{j=1}^{m} B_{ij} W_{ij}. \tag{5.6}$$

Noting that

$$\begin{aligned} & \{(x, y, z, \lambda) : \lambda I + \mathcal{B}(x, y) + \mathcal{C}(z) \geq 0\} \\ = \; & \left\{(x, y, z, \lambda) : \lambda I + \mathcal{R}(x, y, x^T y) + \mathcal{C}(z) \geq 0\right\} \\ \subseteq \; & \left\{(x, y, z, \lambda) : {}^\exists W \text{ such that } \lambda I + \mathcal{R}(x, y, W) + \mathcal{C}(z) \geq 0\right\}, \end{aligned}$$

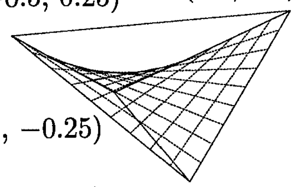

Figure 5: Relaxation of Bilinear Constraint

we see that the minimization of λ with a constraint (5.5) is a semidefinite relaxation of (OP).

The following lemma is useful to derive a tight semidefinite relaxation:

Lemma 2 Given $X \times Y$ in (5.2). One has

$$\left\{(x,\, y,\, W) :\ (x,\, y) \in X \times Y, \quad W = xy^T\right\} \subset \{(x,\, y,\, W) :\ \mathcal{T}(x,\, y,\, W;\, X,\, Y) \geq 0\} \quad (5.7)$$

where $\mathcal{T}$ is defined by

$$\begin{aligned}\mathcal{T}(x,\, y,\, W;\, X,\, Y) \ := \ \operatorname{diag}\Big(\cdots,\ & -L_{y_j}x_i - L_{x_i}y_j + W_{ij} + L_{x_i}L_{y_j},\\ & U_{y_j}x_i + L_{x_i}y_j - W_{ij} - L_{x_i}U_{y_j},\\ & L_{y_j}x_i + U_{x_i}y_j - W_{ij} - U_{x_i}L_{y_j},\\ & -U_{y_j}x_i - U_{x_i}y_j + W_{ij} + U_{x_i}U_{y_j},\ \cdots\Big).\end{aligned}$$

Proof: The convex hull of the set

$$\left\{(x_i,\, y_j,\, W_{ij}) :\ W_{ij} = x_i y_j,\ x_i \in [L_{x_i},\, U_{x_i}],\ y_j \in [L_{y_j},\, U_{y_j}]\right\}$$

is given by the interior of a tetrahedron determined by the following four vertices:

$$(L_{x_i},\, L_{y_j},\, L_{x_i}L_{y_j}),\ (L_{x_i},\, U_{y_j},\, L_{x_i}U_{y_j}),\ (U_{x_i},\, L_{y_j},\, U_{x_i}L_{y_j}),\ (U_{x_i},\, U_{y_j},\, U_{x_i}U_{y_j}).$$

Then (5.7) follows. ∎

Figure 5 is a demonstrative example of Lemma 2, where a bilinear constraint $w = xy$ with $x \in [-0.5,\, 0.5]$, $y \in [-0.5,\, 0.5]$ is relaxed by a linear constraint related to the tetrahedron.

Combination of the linearization (5.5) and Lemma 2 provides an improved semidefinite relaxation of (OP):

Theorem 2 Define J_L by

$$J_L(X \times Y) := \min_{(x,y,z,W)} \lambda$$

such that

$$\begin{bmatrix} \mathcal{A}(x,\, y,\, z) & 0 & 0 \\ 0 & \lambda I + \mathcal{R}(x,\, y,\, W) + \mathcal{C}(z) & 0 \\ 0 & 0 & \mathcal{T}(x,\, y,\, W;\, X,\, Y) \end{bmatrix} \geq 0. \quad (5.8)$$

Then one has

$$J_L(X \times Y) \leq J(X \times Y).$$

Let us define $J_U(X \times Y)$ to be the minimum eigenvalue of the following pencil:

$$\begin{bmatrix} \mathcal{A}(x_L, y_L, z_L) & 0 & 0 & 0 \\ 0 & \mathcal{B}(x_L, y_L) + \mathcal{C}(z_L) & 0 & 0 \\ 0 & 0 & \mathcal{D}_x(x_L; X) & 0 \\ 0 & 0 & 0 & \mathcal{D}_y(y_L; Y) \end{bmatrix} + \lambda \begin{bmatrix} 0_p & 0 & 0 & 0 \\ 0 & I_q & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

where (x_L, y_L, z_L) are part of the lower bound minimizer, i.e.,

$$(x_L, y_L, z_L, W_L) := \arg \min_{(x, y, z, W)} J_L(X \times Y).$$

Then we have the following lemma:

Lemma 3 $J_U(X \times Y)$ is an upper bound for $J(X \times Y)$, i.e.,

$$J(X \times Y) \leq J_U(X \times Y).$$

Furthermore, the gap between $J_U(X \times Y)$ and $J_L(X \times Y)$ is bounded above by the following inequality:

$$J_U(X \times Y) - J_L(X \times Y) \leq \frac{1}{4} \sum_{i=1}^{n} \sum_{j=1}^{m} \rho(B_{ij})(U_{x_i} - L_{x_i})(U_{y_j} - L_{y_j})$$

where $\rho(\cdot)$ denotes the spectral radius.

The proof is essentially the same as that for Theorem 1 in [11], hence it is omitted.

5.3.2. Branch and bound algorithm and performance analysis

Based on the proposed bounds of $J(X \times Y)$, the following algorithm solves (OP) with a given tolerance $\varepsilon > 0$:

Algorithm 1 Let $\varepsilon > 0$ be a given tolerance.

I. Initialization:

$k \leftarrow 0$, $S \leftarrow \{X \times Y\}$, $\lambda_L \leftarrow J_L(X \times Y)$, $\lambda_U \leftarrow J_U(X \times Y)$.

II. repeat {

R1. *branch*:

$S_L \leftarrow \{\mathcal{Q} : \mathcal{Q} \in S, \lambda_L = J_L(\mathcal{Q})\}$.

$\bar{\mathcal{Q}} \leftarrow \arg \max_{\mathcal{Q} \in S_L} \mathrm{Size}(\mathcal{Q})$.

$S \leftarrow S \setminus \{\bar{\mathcal{Q}}\}$.

Split $\bar{\mathcal{Q}}$ along its longest edge equally into $\bar{\mathcal{Q}}_1$ and $\bar{\mathcal{Q}}_2$.

$S \leftarrow S \cup \{\bar{\mathcal{Q}}_1, \bar{\mathcal{Q}}_2\}$.

R2. *update upper bound*:

$\lambda_U \leftarrow \min_{\mathcal{Q} \in S} J_U(\tilde{\mathcal{Q}})$.

R3. *bound*:

$S \leftarrow S \setminus \{\mathcal{Q} : J_L(\mathcal{Q}) > \lambda_U\}$.

R4. *update lower bound*:

$\lambda_L \leftarrow \min_{\mathcal{Q} \in S} J_L(\mathcal{Q})$.

$k \leftarrow k + 1$.

} until $\lambda_U - \lambda_L < \varepsilon$

where $\mathrm{Size}(\cdot)$ denotes the length of the longest edge defined by

$$\mathrm{Size}(X \times Y) := \max \left\{ \max_i \{U_{x_i} - L_{x_i}\}, \max_j \{U_{y_j} - L_{y_j}\} \right\}.$$

Due to Lemma 3, we can prove that Algorithm 1 terminates in finite time:
Theorem 3 Let $\varepsilon > 0$ be a given tolerance. Algorithm 1 terminates, i.e., one obtains a suboptimal value λ_ε satisfying

$$\lambda_\varepsilon - \varepsilon \leq J(X \times Y) \leq \lambda_\varepsilon$$

and a sub-optimizer $(x_\varepsilon,\ y_\varepsilon,\ z_\varepsilon)$ such that

$$\begin{bmatrix} \mathcal{A}(x_\varepsilon, y_\varepsilon, z_\varepsilon) & 0 & 0 & 0 \\ 0 & \lambda_\varepsilon I + \mathcal{B}(x_\varepsilon, y_\varepsilon) + \mathcal{C}(z_\varepsilon) & 0 & 0 \\ 0 & 0 & \mathcal{D}_x(x_\varepsilon;\ X) & 0 \\ 0 & 0 & 0 & \mathcal{D}_y(y_\varepsilon;\ Y) \end{bmatrix} \geq 0$$

until the step $k = \kappa$, where κ is bounded above by

$$\log_2 \kappa \leq \sum_{i=1}^{n} \kappa_{x_i} + \sum_{j=1}^{m} \kappa_{y_j}$$

and κ_{x_i}'s and κ_{y_j}'s are minimum positive integers satisfying

$$\kappa_{x_i} > \frac{1}{2} \log_2 \left(\sum_{i=1}^{n} \sum_{j=1}^{m} \rho(B_{ij}) \right) - \frac{1}{2} \log_2 \varepsilon + \log_2 (U_{x_i} - L_{x_i}),$$

$$\kappa_{y_j} > \frac{1}{2} \log_2 \left(\sum_{i=1}^{n} \sum_{j=1}^{m} \rho(B_{ij}) \right) - \frac{1}{2} \log_2 \varepsilon + \log_2 (U_{y_j} - L_{y_j}).$$

Proof: Lemma 3 guarantees that the gap related to a branch $\mathcal{Q}$ is less than ε if

$$\frac{1}{4} (\mathrm{Size}(\mathcal{Q}))^2 \sum_{i=1}^{n} \sum_{j=1}^{m} \rho(B_{ij}) < \varepsilon. \tag{5.9}$$

Hence Algorithm 1 terminates when all the branches satisfy (5.9).

If we split the edge of $X \times Y$ related to x_1 with equal spacing in $2^{\kappa_{x_1}}$ times, the lengths of all edges related to x_1 of resultant branches are bounded by

$$\frac{U_{x_1} - L_{x_1}}{2^{\kappa_{x_1}}} < 2\sqrt{\varepsilon} \left(\sum_{i=1}^{n} \sum_{j=1}^{m} \rho(B_{ij}) \right)^{-\frac{1}{2}}.$$

Similarly all the resultant branches satisfy (5.9) if we split all the edges related to x_i's and y_j's $2^{\kappa_{x_i}}$ and $2^{\kappa_{y_j}}$ times respectively.

By the selection scheme of $\bar{\mathcal{Q}}$, the number of loops k required for the optimization with tolerance $\varepsilon > 0$ is given by the product of $2^{\kappa_{x_i}}$'s and $2^{\kappa_{y_j}}$'s. This completes the proof. ■

5.4. Primal-relaxed dual algorithm

Basic ideas of primal-relaxed dual approaches date back to the generalized Benders decomposition method [16], which can solve only some special problems. Recently, these approaches have received significant attention in the area of global optimization since Floudas and Visweswaran [7] proposed a global optimization algorithm for a larger class of problems, i.e., mathematical programming problems whose objective and constraints are both biconvex. In this subsection, we present a global optimization algorithm for the BMI problem based on the primal-relaxed dual method [42, 1].

5.4.1. Primal and relaxed dual problems

Define the following problem as the primal problem:

$$\text{(P)} \quad \min_{x,\, z,\, \lambda} \ \lambda \quad \text{subject to} \ \mathcal{F}(x,\, y^k,\, z,\, \lambda) \geq 0 \tag{5.10}$$

where k will denote the k-th iteration in the subsequent algorithm and $y^k \in Y$. Since this problem (P) is simply (OP) solved for fixed values of $y = y^k$, it represents an upper bound on the optimal value of (OP). (P) is a semidefinite programming problem and therefore can be solved efficiently.

We now introduce a Lagrangian associated with the problem (OP):

$$L(x,\, y,\, z,\, \lambda,\, \mathcal{G}) := \lambda - \mathrm{Tr}\{\mathcal{F}(x,\, y,\, z,\, \lambda)\, \mathcal{G}\} \tag{5.11}$$

where $\mathcal{G} \in \mathcal{S}_{p+q+2n+2m}$ is the Lagrange multiplier corresponding to the LMI constraint of the primal problem (P). Since the problem (P) satisfies Slater's constraint qualification, the following dual relation holds for any $y \in Y$ [23]:

$$\begin{aligned} \min_{x,\, z,\, \lambda,\, \mathcal{F}(x,y,z,\lambda)\geq 0} \lambda &= \min_{x,\, z,\, \lambda} \max_{\mathcal{G}\geq 0} L(x,\, y,\, z,\, \lambda,\, \mathcal{G}) \\ &= \max_{\mathcal{G}\geq 0} \min_{x,\, z,\, \lambda} L(x,\, y,\, z,\, \lambda,\, \mathcal{G}). \end{aligned} \tag{5.12}$$

Hence the problem (OP) is equivalent to the following problem:

$$\min_{y\in Y,\, \mu} \ \mu \quad \text{subject to} \ \mu \geq \min_{x,\, z,\, \lambda} L(x,\, y,\, z,\, \lambda,\, \mathcal{G}), \quad {}^{\forall}\mathcal{G} \geq 0. \tag{5.13}$$

This problem is difficult to solve since it contains an infinite number of constraints. By using the Lagrange multipliers, the relaxed dual problem with a finite number of constraints is obtained:

$$\text{(RD)} \quad \min_{y\in Y,\, \mu} \ \mu \quad \text{subject to} \ \mu \geq \min_{x,\, z,\, \lambda} L(x,\, y,\, z,\, \lambda,\, \mathcal{G}^k), \quad {}^{\forall}k = 1,\, \ldots,\, K \tag{5.14}$$

where $\mathcal{G}^k \geq 0$ is the optimal Lagrange multipliers corresponding to the primal problem (P) for $y = y^k$. The problem (RD) contains fewer constraints than (5.13), and hence provides a valid lower bound for the original problem (OP).

5.4.2. Primal-relaxed dual algorithm for BMI

In the following, we present a global optimization algorithm for the BMI problem based on the primal-relaxed dual method. Roughly speaking, the algorithm consists of the following two procedures:

(i) Solve the primal problem (P) for $y = y^k$ and update the upper bound.

(ii) Solve the subproblems of the relaxed dual problem (RD) and update the lower bound.

These two procedures are to be repeated until the difference between the upper and lower bounds becomes less than the prescribed tolerance ϵ (> 0).

We state the BMI primal-relaxed dual algorithm [42] after some definitions.

Definition 2 At the K-th iteration,

(1) define x^{E_j} $(j = 1, \ldots, 2^n)$ to be the vertices of X, and let $\mathcal{E}$ be the set of all E_j. Also, define $\mathcal{I}^K$ to be the set of i's for which $\nabla_{x_i} L(x, y, z, \lambda, \mathcal{G}^K)$ is a function of y, and let $\mathcal{J}(k, K)$ be the set of j's such that

$$\left.\begin{array}{ll} \nabla_{x_i} L(x,\, y^K,\, z,\, \lambda,\, \mathcal{G}^k) \;\geq\; 0, & \text{if } x_i^{E_j} = L_{x_i} \\ \nabla_{x_i} L(x,\, y^K,\, z,\, \lambda,\, \mathcal{G}^k) \;\leq\; 0, & \text{if } x_i^{E_j} = U_{x_i} \end{array}\right\} \;\forall i \in \mathcal{I}^k$$

(2) define $\mu^{\text{stor}}(K, E_l)$ and $y^{\text{stor}}(K, E_l)$ to be the optimal solutions of the following subproblem (SUBP) associated with (RD)

$$\mu^{\text{stor}}(K, E_l)$$
$$= \left\{\begin{array}{l} \min\limits_{y \in Y,\, \mu} \;\; \mu \\ \text{subject to} \\ \left.\begin{array}{l} \mu \geq L(x^{E_j},\, y,\, z,\, \lambda,\, \mathcal{G}^k) \\ \quad \left.\begin{array}{l} \nabla_{x_i} L(x,\, y,\, z,\, \lambda,\, \mathcal{G}^k) \geq 0, \text{ if } x_i^{E_j} = L_{x_i} \\ \nabla_{x_i} L(x,\, y,\, z,\, \lambda,\, \mathcal{G}^k) \leq 0, \text{ if } x_i^{E_j} = U_{x_i} \end{array}\right\} \forall i \in \mathcal{I}^k \\ \qquad \text{for } j = \mathcal{J}(k, K), \quad k = 1, 2, \cdots, K-1 \end{array}\right\} \\ \left.\begin{array}{l} \mu \geq L(x^{E_l},\, y,\, z,\, \lambda,\, \mathcal{G}^K) \\ \quad \left.\begin{array}{l} \nabla_{x_i} L(x,\, y,\, z,\, \lambda,\, \mathcal{G}^K) \geq 0, \text{ if } x_i^{E_l} = L_{x_i} \\ \nabla_{x_i} L(x,\, y,\, z,\, \lambda,\, \mathcal{G}^K) \leq 0, \text{ if } x_i^{E_l} = U_{x_i} \end{array}\right\} \forall i \in \mathcal{I}^K \end{array}\right\} \end{array}\right.$$

Algorithm 2 [BMI Primal-Relaxed Dual Algorithm]

Step 0. Initialization of Parameters.

Let P^{UBD} and M^{LBD} be a very large *positive* number and a very large *negative* number, respectively. Select a convergence tolerance parameter ϵ (> 0). Set $K = 1$ and select an initial fixed value $y^1 \in Y$.

Step 1. Primal Problem.

Store the value of y^K. Solve the primal problem (P) for $y = y^K$. Store the optimal Lagrange multiplier $\mathcal{G}^K$. Update the upper bound so that

$$P^{\text{UBD}} = \min\,(P^{\text{UBD}}, P^K)$$

where P^K is the solution of the K-th primal problem.

Step 2. Selection of Lagrangians from the Previous Iterations.

For $k = 1, 2, \cdots, K-1$, select the Lagrangian corresponding to $j = \mathcal{J}(k, K)$.

Step 3. Relaxed Dual Problem.

For all $E_l \in \mathcal{E}$, solve the subproblem (SUBP) and store the solutions μ^{stor}, y^{stor}.

Step 4. Selection of a New Lower Bound and y^{K+1}.

From the stored set μ^{stor}, select the minimum μ^K_{min}, and set $M^{\text{LBD}} = \mu^K_{\text{min}}$. Also, select the corresponding stored value of y^{stor} as y^{K+1}. Delete μ^K_{min} and y^{K+1} from μ^{stor} and y^{stor}, respectively.

Step 5. Check for Convergence.

Check if $P^{\text{UBD}} - M^{\text{LBD}} \leq \epsilon$. If yes, stop, else set $K = K + 1$, and return to Step 1.

Remark 5 In Step 3 of the algorithm, the relaxed dual problem (RD) is decomposed into the subproblems (SUBP) that are formulated as linear programming problems. By solving these subproblems, the lower bound on the original problem is obtained. It should be noted that the convergence and global optimality of the algorithm is proved in Floudas and Visweswaran [7].

6. Conclusion

We have given a brief and partial review of the relationship between control and optimization. Emphasis is placed upon the modern H^∞ control and how it can be placed in a unified scope of semidefinite programming. For the benefit of the reader, we have given some motivating discussions of the simplest sensitivity minimization problem. The solutions via the Nevalinna-Pick interpolation and the Nehari theorem have been discussed, which in turn motivated the solutions via LMIs.

The LMI approach to the H^∞ control problem leads to the study of BMIs by putting an additional constraint on the order of the controller, which constitutes the forefront of the modern control issues. Finally, two types of global optimization algorithms for solving BMIs have been provided.

References

[1] E. Beran, L. Vandenberghe and S. Boyd: A global BMI algorithm based on the generalized Benders decomposition. *Proceedings of the 1997 European Control Conference, CD-ROM Version* (1997).

[2] S. Boyd and C. Barratt: *Linear Controller Design: Limits of Performance* (Prentice Hall, 1991).

[3] S. Boyd et al.: *Linear Matrix Inequalities in Systems and Control Theory* (SIAM Publications, 1994).

[4] J. C. Doyle, B. A. Francis and A. R. Tannenbaum: *Feedback Control Theory* (McMillan, 1992).

[5] J. C. Doyle et al.: State-space solutions to standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control problems. *IEEE Transactions on Automatic Control*, **AC-34**-8 (1989), 831–847.

[6] J. C. Doyle, A. Packard and K. Zhou: Review of LFTs, LMIs, and μ. *Proceedings of the 29th IEEE Conference on Decision and Control*, (1991) 1227–1232.

[7] C. A. Floudas and V. Visweswaran: Primal-relaxed dual global optimization approach. *Journal of Optimization Theory and Applications*, **78**-2 (1993) 187–225.

[8] C. Foias, H. Özbay and A. Tannenbaum: *Robust Control of Infinite Dimensional Systems: Lecture Notes in Control and Information Sciences, Vol. 209* (Springer, 1991).

[9] B. A. Francis: *A Course in H_∞ Control Theory: Lecture Notes in Control and Information Sciences, Vol. 88* (Springer, 1987).

[10] B. A. Francis, J. W. Helton and G. Zames: H_∞-optimal feedback controllers for linear multivariable systems. *IEEE Transactions on Automatic Control*, **AC-29** (1984) 888–900.

[11] H. Fujioka and K. Hoshijima: Bounds for the BMI eigenvalue problem. *Transactions of the Society of Instrument and Control Engineers*, **33**-7 (1997) 616–621.

[12] H. Fujioka and T. Iwasaki: BMI: A challenge to control synthesis based on non-convex matrix Inequalities (in Japanese). *Journal of the Society of Instrument and Control Engineers*, **36**-11 (1997) 762–767.

[13] H. Fujioka and Y. Wakasa: BMI: A non-convex optimization problem for control systems design (in Japanese). *Proceedings of the 10th RAMP Symposium*, (1998) 81–94.

[14] P. Gahinet et al.: *LMI Control Toolbox* (The Math Works, Inc., 1995).

[15] P. Gahinet and P. Apkarian: A linear matrix inequality approach to $\mathcal{H}_\infty$ control. *International Journal of Robust and Nonlinear Control*, **4** (1994) 421–448.

[16] A. M. Geoffrion: Generalized Benders decomposition. *Journal of Optimization Theory and Applications*, **10**-4 (1972) 237–260.

[17] K. C. Goh et al.: Biaffine matrix inequality properties and computational methods. *Proceedings of the 1994 American Control Conference*, (1994) 850–855.

[18] K. C. Goh et al.: A global optimization approach for the BMI problem. *Proceedings of the 33rd IEEE Conference on Decision and Control*, (1994) 2009–2014.

[19] M. Ikeda, G. Zhai and Y. Fujisaki: Decentralized $\mathcal{H}_\infty$ controller design for large-scale systems. *Proceedings of the 35th IEEE Conference on Decision and Control*, (1996) 1–6.

[20] T. Iwasaki: *LMI and Control* (in Japanese) (Shokodo, 1997).

[21] T. Iwasaki: Open problems and partial solutions in control engineering. *Proceedings of the 10th RAMP Symposium*, (1998) 53–65.

[22] T. Iwasaki and R. E. Skelton: All controllers for the general H_∞ control problem: LMI existence conditions and state space formulas. *Automatica*, **30**-8 (1994) 1307–1317.

[23] J. Jahn: *Introduction to the Theory of Nonlinear Optimization* (Springer-Verlag, 1994).

[24] T. Kailath: *Linear Systems* (Prentice-Hall, 1980).

[25] M. Kawanishi, T. Sugie and H. Kanki: BMI global optimization based on branch and bound method taking account of the property of local minima. *Proceedings of the 36th IEEE Conference on Decision and Control*, (1997) 781–786.

[26] S. Liu and G. P. Papavassilopoulos: Numerical experience with parallel algorithms for solving the BMI problem. *Proceedings of the 13th World Congress of IFAC*, **D** (1996) 387–391.

[27] I. Masubuchi, A. Ohara and N. Suda: LMI-based output feedback controller design. *Proceedings of the 1995 American Control Conference*, (1995) 3473–3477.

[28] J. C. Maxwell: On governors. *Proceedings of the Royal Soc. London*, **16** (1868) 270–283.

[29] A. Megretski and A. Rantzer: System analysis via integral quadratic constraints. *IEEE Transactions on Automatic Control*, **AC-42**-6 (1997) 819–830.

[30] M. Mesbahi and G. P. Papavassilopoulos: On the rank minimization problem over a positive semidefinite linear matrix inequality. *IEEE Transactions on Automatic Control*, **AC-42**-2 (1997) 239–243.

[31] Z. Nehari: On bounded bilinear forms. *Ann. of Math.*, **65** (1957) 153–162.

[32] Y. Nesterov and A. Nemirovski: *Interior Point Polynomial Methods in Convex Programming: Theory and Applications* (SIAM Publications, 1994).

[33] G. Obinata: Simultaneous optimal design of structural and control systems. *Journal of the Society of Instrument and Control Engineers*, **36**-4 (1997) 254–261.

[34] A. Packard: Gain scheduling via linear fractional transformations. *Systems and Control Letters*, **22** (1994) 72–92.

[35] A. Packard et al.: A collection of robust control problems leading to LMI's. *Proceedings of the 30th IEEE Conference on Decision and Control*, (1991) 1245–1250.

[36] A. Packard et al.: Optimal, constant I/O similarity scaling for full-information and state-feedback control problems. *Systems and Control Letters*, **19** (1992) 271–280.

[37] M. G. Safonov, K. C. Goh and J. H. Ly: Control system synthesis via bilinear matrix inequalities. *Proceedings of the 1994 American Control Conference*, (1994) 45–49.

[38] C. Scherer, P. Gahinet and M. Chilali: Multiobjective output-feedback control via LMI optimization. *IEEE Transactions on Automatic Control*, **AC-42**-7 (1997) 896–911.

[39] V. L. Syrmos et al.: Static output feedback – a survey. *Automatica*, **33**-2 (1997) 125–137.

[40] S. Takano, T. Watanabe and K. Yasuda: Branch and bound technique for global solution of BMI. *Transactions of the Society of Instrument and Control Engineers*, **33**-7 (1997) 701–708.

[41] O. Toker and H. Özbay: On the $\mathcal{NP}$-hardness of solving bilinear matrix inequalities and simultaneous stabilization with static output feedback. *Proceedings of the 1995 American Control Conference*, (1995) 2525–2526.

[42] Y. Wakasa, M. Sasaki, T. Tanino and T. Katayama: A global approach to the BMI problem using the primal-relaxed dual method (in Japanese). *Transactions of the Society of Instrument and Control Engineers*, **35**-6 (1999) 748–753.

[43] Y. Yamada and S. Hara: The matrix product eigenvalue problem. *Proceedings of the 36th IEEE Conference on Decision and Control*, (1997) 4926–4931.

[44] Y. Yamamoto: *Mathematics for Systems and Control* (in Japanese) (Asakura, 1998).

[45] G. Zames: Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms and approximate inverses. *IEEE Transactions on Automatic Control*, **AC-26** (1981) 301–320.

[46] G. Zames and B. A. Francis: Feedback, minimax sensitivity and optimal robustness. *IEEE Transactions on Automatic Control*, **AC-28** (1983) 585–601.

[47] K. Zhou, J. C. Doyle and K. Glover: *Robust and Optimal Control* (Prentice Hall, 1996).

Yuji Wakasa
Graduate School of Informatics
Kyoto University
Kyoto 606-8501, JAPAN
E-mail: `wakasa@acs.i.kyoto-u.ac.jp`