On Optimal Solutions to Two-Block H^∞ Problems *

Babak Hassibi and Thomas Kailath

Information Systems Laboratory
Stanford University

Abstract

In this paper we obtain a new formula for the minimum achievable disturbance attenuation in two-block H^∞ problems. This new formula has the same structure as the optimal H^∞ norm formula for noncausal problems, except that doubly-infinite (so-called Laurent) operators must be replaced by semi-infinite (so-called Toeplitz) operators. The benefit of the new formula is that it allows us to find explicit expressions for the optimal H^∞ norm in several important cases: the equalization problem (or its dual, the tracking problem), and the problem of filtering signals in additive noise. Furthermore, it leads us to the concepts of “worst-case non-estimability”, corresponding to when causal filters cannot reduce the H^∞ norms from their a priori values, and “worst-case complete estimability”, corresponding to when causal filters offer the same H^∞ performance as noncausal ones. We also obtain an explicit characterization of worst-case non-estimability and study the consequences to the problem of equalization with finite delay.

1. Introduction

The formula for the minimum achievable disturbance attenuation in two-block H^∞ problems (denoted, hereafter, by γ_c) was obtained by Verma and Jonckheere [1] and Feintuch and Francis [2] in the mid 1980's. In these works, γ_c was described as the spectral radius of a mixed Toeplitz-plus-Hankel operator. Despite the elegance of this result, little physical insight into the properties of the two-block problem have been obtained in this framework. Moreover, the explicit computation of the spectral radius of the mixed Toeplitz-plus-Hankel operator has been superseded by state-space Riccati-based approaches that compute this quantity only implicitly (see e.g., [3]).

In this paper we attempt to show that much insight into the two-block H^∞ problem can be obtained by studying the minimum achievable disturbance attenuation. The main result in this attempt is a new formula for γ_c which is, in our view, simpler than the mixed Toeplitz-plus-Hankel spectral radius formula, and which retains the same structure as the formula for the minimum achievable disturbance attenuation in noncausal two-block H^∞ problems. To demonstrate this fact, we give a complete analysis of γ_c for the important problems of equalization

and of filtering signals in additive noise.

While such a complete analysis for general two block problems is not currently available (primarily due to the fact that simple frequency-domain characterizations of γ_c currently do not exist), we do introduce two concepts that we believe are of considerable importance. The first is *worst-case complete estimability* which corresponds to when causal estimators have the same H^∞ performance as non-causal ones, and roughly speaking represents an “easy” estimation problem. The second is *worst-case non-estimability* which corresponds to when causal estimators cannot reduce γ_c from their a priori bounds, and roughly speaking represents a “difficult” estimation problem. These concepts are important since in estimation (and control) one would like to set up problems that are close to worst-case complete estimable and that avoid worst-case non-estimability. Although we have not been able to give a simple characterization of worst-case complete estimability, we do give one for worst-case non-estimability and demonstrate its merits by studying the problem of equalization with finite delay.

1.1. Notation

In this paper we shall often deal with time-invariant operators that map $l^{2,m}$ to $l^{2,p}$, i.e., the space of square-summable sequences of m -vectors to the space of square-summable sequences of p -vectors, according to the rule

$$y_i = \sum_{j=-\infty}^{\infty} T_{i-j} u_j, \quad (1)$$

where $u = \{u_i\} \in l^{2,m}$ and $y = \{y_i\} \in l^{2,p}$. Note that the above equation can also be written as

$$\begin{bmatrix} \vdots \\ y_{-1} \\ y_0 \\ y_1 \\ \vdots \end{bmatrix} = \underbrace{\begin{bmatrix} \ddots & \ddots & \vdots & \vdots \\ \ddots & T_0 & T_{-1} & T_{-2} & \dots \\ \dots & T_1 & T_0 & T_{-1} & \dots \\ \dots & T_2 & T_1 & T_0 & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}}_{\triangleq \mathcal{T}} \begin{bmatrix} \vdots \\ u_{-1} \\ u_0 \\ u_1 \\ \vdots \end{bmatrix}.$$

Another characterization of the operator $\mathcal{T}$ is through its z -transform,

$$T(z) = \sum_{i=-\infty}^{\infty} T_i z^{-i}. \quad (2)$$

*This work was supported in part by DARPA through the Department of Air Force under contract F49620-95-1-0525-P00001 and by the Joint Service Electronics Program at Stanford under contract DAAH04-94-G-0058-P00003.

We shall further assume that the above (doubly) infinite-series expansion converges absolutely on an annulus containing the unit circle, $|z| = 1$. When $T(z)$ is nonrational this is slightly stronger than the requirement that $\mathcal{T}$ maps $l^{2,m}$ to $l^{2,p}$, since it guarantees that $\mathcal{T}$ maps $l^{\infty,m}$ to $l^{\infty,p}$, i.e., that it maps bounded-amplitude sequences to bounded-amplitude sequences. In anycase, this implies that the Fourier transform, $T(e^{j\omega})$, is well-defined and bounded for all $\omega \in [0, 2\pi]$.

We shall also find it useful to partition the input and output, u and y , into their past, $u_- \triangleq \{u_i, i < 0\}$ and $y_- \triangleq \{y_i, i < 0\}$, and present and future, $u_+ \triangleq \{u_i, i \geq 0\}$ and $y_+ \triangleq \{y_i, i \geq 0\}$, components. We shall also denote the corresponding spaces of semi-infinite sequences by $l_-^{2,m}$, $l_-^{2,p}$ and $l_+^{2,m}$, $l_+^{2,p}$, respectively. With this partitioning of the input and output spaces, the operator $\mathcal{T}$ can be partitioned as follows:

$$\mathcal{T} = \left[\begin{array}{c|c} \mathcal{T}_- & \mathcal{T}_A \\ \hline \mathcal{T}_H & \mathcal{T}_+ \end{array} \right], \quad (3)$$

where

$$\mathcal{T}_- \triangleq \begin{bmatrix} \ddots & \ddots & \vdots & \vdots \\ \ddots & T_0 & T_{-1} & T_{-2} \\ \dots & T_1 & T_0 & T_{-1} \\ \dots & T_2 & T_1 & T_0 \end{bmatrix} \quad \mathcal{T}_A \triangleq \begin{bmatrix} \vdots & \vdots & \ddots \\ T_{-2} & T_{-3} & \dots \\ T_{-1} & T_{-2} & \dots \end{bmatrix}$$

$$\mathcal{T}_H \triangleq \begin{bmatrix} \dots & T_2 & T_1 \\ \dots & T_3 & T_2 \\ \vdots & \vdots & \vdots \end{bmatrix} \quad \mathcal{T}_+ \triangleq \begin{bmatrix} T_0 & T_{-1} & T_{-2} & \dots \\ T_1 & T_0 & T_{-1} & \dots \\ T_2 & T_1 & T_0 & \dots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix}$$

The operator $\mathcal{T}_-$ maps $l_-^{2,m}$ to $l_-^{2,p}$, i.e., past inputs to past outputs, and is called a *Toeplitz* operator, whereas the operator $\mathcal{T}_H$ maps $l_-^{2,m}$ to $l_+^{2,p}$, i.e., past inputs to present and future outputs, and is called a *Hankel* operator. Similar remarks apply to $\mathcal{T}_+$ and $\mathcal{T}_A$. The doubly-infinite operator $\mathcal{T}$, however, is referred to as a *Laurent* operator.

We will also be interested in the induced 2-norm, or so-called H^∞ norm, of such operators which is defined as

$$\|\mathcal{T}\|_\infty \triangleq \sup_{u \neq 0 \in l^{2,m}} \frac{\|\mathcal{T}u\|_2}{\|u\|_2}, \quad \|\mathcal{T}_-\|_\infty \triangleq \sup_{u_- \neq 0 \in l_-^{2,m}} \frac{\|\mathcal{T}_-u_-\|_2}{\|u_-\|_2}$$

where we have used $\|a\|^2 = \sum_i a_i^* a_i$. Moreover, it turns out that, since Laurent and Toeplitz operators are time-invariant, there is a very simple frequency-domain characterization of their H^∞ norms:

$$\|\mathcal{T}\|_\infty = \|\mathcal{T}_-\|_\infty = \sup_{\omega \in [0, 2\pi]} \bar{\sigma}[T(e^{j\omega})],$$

where $\bar{\sigma}(\cdot)$ denotes the maximum singular value of its argument. [Note that $\|\mathcal{T}\|_\infty = \|\mathcal{T}_-\|_\infty$.]

Finally, we should mention that the operators $\mathcal{T}$ and $\mathcal{T}_-$ will be called causal if $T_i = 0$, for all $i < 0$, or, in other words, if $\mathcal{T}_A = 0$.

2. The Two-Block Problem

Consider the following "two-block" operator,

$$\mathcal{T}_K = \begin{bmatrix} \mathcal{L} - \mathcal{K}\mathcal{H} & -\mathcal{K} \end{bmatrix}, \quad (4)$$

where $\mathcal{L}$ and $\mathcal{H}$ are causal Laurent operators. The operator $\mathcal{K}$ is also Laurent, though not necessarily causal.

Many estimation problems lead to the operator $\mathcal{T}_K$. To this end, consider Fig. 1 where $\mathcal{H}$ and $\mathcal{L}$ are known causal linear time-invariant operators (or simply, causal LTI systems), $\{u_i\}$ and $\{v_i\}$ are unknown sequences, $\{y_i\}$ is the known observations sequence, and $\{s_i\}$ is the unobservable desired sequence we wish to estimate. The goal in estimation is to appropriately design the (so-called) estimator $\mathcal{K}$ that provides the estimates $\{\hat{s}_i\}$ based on the observations $\{y_i\}$. When $\mathcal{K}$ is a non-causal operator the estimation problem is referred to as a *smoothing* problem since the estimator has access to future observations, whereas when $\mathcal{K}$ is a causal operator the estimation problem is referred to as a *filtering* problem since the estimator does not have access to future observations.

Figure 1: A general estimation problem.

In this setting, $\mathcal{T}_K$ is the transfer operator that maps the unknown disturbances $\{u_j\}$ and $\{v_j\}$ to the estimation errors $\tilde{s} \triangleq \{\tilde{s}_i\}$. Thus, the problem of estimation becomes one of making $\mathcal{T}_K$ "small" in some sense. In the two block H^∞ problem, it is proposed to minimize $\mathcal{T}_K$ in the H^∞ norm sense.

Problem 1 (Two-Block H^∞ Problem) Consider the causal Laurent operators, $\mathcal{L}$ and $\mathcal{H}$.

(a) Find γ_s , where

$$\gamma_s \triangleq \inf_K \left\| \begin{bmatrix} \mathcal{L} - \mathcal{K}\mathcal{H} & -\mathcal{K} \end{bmatrix} \right\|_\infty. \quad (5)$$

(b) Find γ_c , where

$$\gamma_c \triangleq \inf_{\text{causal } K} \left\| \begin{bmatrix} \mathcal{L} - \mathcal{K}\mathcal{H} & -\mathcal{K} \end{bmatrix} \right\|_\infty. \quad (6)$$

Note that in Problem 1-(a) there is no restriction on the Laurent operator $\mathcal{K}$, whereas in Problem 1-(b) the Laurent operator is restricted to being causal. It is thus clear that

$$\gamma_s \leq \gamma_c. \quad (7)$$

Note, moreover, that γ_s^2 and γ_c^2 can be regarded as the min-max energy gains from the disturbances to the estimation errors, which is the reason why γ_s and γ_c are also referred to as the minimum disturbance attenuations.

Finding an expression for the minimum disturbance attenuation, when $\mathcal{K}$ is not restricted to be causal, is quite straightforward and only requires a "completion of squares" argument.

Theorem 1 (Non-Causal Optimal H^∞ Norm)
Consider the setting of Problem 1-(b). Then we have

$$\begin{aligned}\gamma_s^2 &= \left\| \mathcal{L} (I + \mathcal{H}^* \mathcal{H})^{-1} \mathcal{L}^* \right\|_\infty \\ &= \sup_{\omega \in [0, 2\pi]} \bar{\sigma} \left[L(e^{j\omega}) (I + H^*(e^{j\omega}) H(e^{j\omega}))^{-1} L^*(e^{j\omega}) \right].\end{aligned}\quad (8)$$

3. Optimal H^∞ Norm in the Causal Case

The main result of this paper has to do with a new expression for γ_c , which is given below.

Theorem 2 (Causal Optimal H^∞ Norm) Consider the setting of Problem 1-(b). Then we have

$$\gamma_c^2 = \left\| \mathcal{L}_- (I + \mathcal{H}_-^* \mathcal{H}_-)^{-1} \mathcal{L}_-^* \right\|_\infty. \quad (9)$$

Note that the expressions for the minimum disturbance attenuation in the non-causal and causal cases (of Theorems 1 and 2) have the exact same structure! The only difference between the two is that the doubly-infinite Laurent operators $\mathcal{L}$ and $\mathcal{H}$ must be replaced by the semi-infinite Toeplitz operators $\mathcal{L}_-$ and $\mathcal{H}_-$.

The reason why we were able to give a simple frequency-domain expression for γ_s in Theorem 1 is that products and inverses of Laurent operators are themselves Laurent operators. Thus $\mathcal{L} (I + \mathcal{H}^* \mathcal{H})^{-1} \mathcal{L}^*$ is a Laurent operator and the second equation in (8) readily follows. However, products and inverses of Toeplitz operators are not necessarily Toeplitz (the reader may want to check, for example, that $\mathcal{H}_-^* \mathcal{H}_-$ is not Toeplitz). Thus $\mathcal{L}_- (I + \mathcal{H}_-^* \mathcal{H}_-)^{-1} \mathcal{L}_-^*$ is not necessarily a Toeplitz operator, and so simple frequency-domain formulas for γ_c in Eq. (9) cannot generally be given.

However, some further characterization of $\mathcal{L}_- (I + \mathcal{H}_-^* \mathcal{H}_-)^{-1} \mathcal{L}_-^*$ can indeed be given. To this end, let us define

$$\mathcal{E} \triangleq \mathcal{L} (I + \mathcal{H}^* \mathcal{H})^{-1} \mathcal{L}^* = \left[\begin{array}{c|c} \mathcal{E}_- & \mathcal{E}_H^* \\ \hline \mathcal{E}_H & \mathcal{E}_+ \end{array} \right], \quad (10)$$

and note that

$$\mathcal{E} = \mathcal{L} \mathcal{L}^* - \mathcal{L} \mathcal{H}^* (I + \mathcal{H} \mathcal{H}^*)^{-1} \mathcal{H} \mathcal{L}^*. \quad (11)$$

Now, if we further define the spectral factorization

$$\Delta \Delta^* = I + \mathcal{H} \mathcal{H}^*, \quad (12)$$

with Δ causal and causally invertible, and the operator

$$\mathcal{P} \triangleq \Delta^{-1} \mathcal{H} \mathcal{L}^* = \left[\begin{array}{c|c} \mathcal{P}_- & \mathcal{P}_A \\ \hline \mathcal{P}_H & \mathcal{P}_+ \end{array} \right], \quad (13)$$

then using (11) we can identify the Toeplitz part of $\mathcal{E}$ as

$$\mathcal{E}_- = (\mathcal{L} \mathcal{L}^*)_ - - (\mathcal{P}^* \mathcal{P})_- = \mathcal{L}_- \mathcal{L}_-^* - \mathcal{P}_-^* \mathcal{P}_- - \mathcal{P}_H^* \mathcal{P}_H. \quad (14)$$

But it is easy to see that

$$\mathcal{P}_- = \Delta_-^{-1} \mathcal{H}_- \mathcal{L}_-^*, \quad (15)$$

which implies that $\mathcal{E}_-$ can be written as

$$\mathcal{E}_- = \mathcal{L}_- \mathcal{L}_-^* - \mathcal{L}_- \mathcal{H}_-^* \Delta_-^{-1} \mathcal{H}_- \mathcal{L}_-^* - \mathcal{P}_H^* \mathcal{P}_H. \quad (16)$$

But from (12) it follows that $\Delta_- \Delta_-^* = I + \mathcal{H}_- \mathcal{H}_-^*$, so that

$$\begin{aligned}\mathcal{E}_- &= \mathcal{L}_- \mathcal{L}_-^* - \mathcal{L}_- \mathcal{H}_-^* (I + \mathcal{H}_- \mathcal{H}_-^*)^{-1} \mathcal{H}_- \mathcal{L}_-^* - \mathcal{P}_H^* \mathcal{P}_H \\ &= \mathcal{L}_- (I + \mathcal{H}_-^* \mathcal{H}_-)^{-1} \mathcal{L}_-^* - \mathcal{P}_H^* \mathcal{P}_H.\end{aligned}$$

We thus finally have

$$\mathcal{L}_- (I + \mathcal{H}_-^* \mathcal{H}_-)^{-1} \mathcal{L}_-^* = \mathcal{E}_- + \mathcal{P}_H^* \mathcal{P}_H, \quad (17)$$

which, in view of Theorem 2, recovers a wellknown result of Verma and Jonckheere [1] and Feintuch and Francis [2] that states that the minimum disturbance attenuation is given by the spectral radius of the so-called mixed Toeplitz-plus-Hankel operator $\mathcal{E}_- + \mathcal{P}_H^* \mathcal{P}_H$.

Eq. (17) is significant since it shows that the operator $\mathcal{L}_- (I + \mathcal{H}_-^* \mathcal{H}_-)^{-1} \mathcal{L}_-^*$ is generally not Toeplitz, and that it differs from the Toeplitz operator $\mathcal{E}_-$ by the amount of $\mathcal{P}_H^* \mathcal{P}_H$. Note also that $\gamma_s^2 = \|\mathcal{E}_-\|_\infty$. Therefore (17) shows, as expected, that $\gamma_c \geq \gamma_s$, and that the increase in the minimum disturbance attenuation depends on the Hankel operator $\mathcal{P}_H$.

Although the mixed Toeplitz-plus-Hankel spectral radius result of [1, 2] is intriguing, it has not proven to be a very useful tool in analyzing the behaviour of γ_c with respect to $\mathcal{L}$ and $\mathcal{H}$. In the remainder of this paper, however, we shall see that our new formula (9) allows us to draw various qualitative and quantitative conclusions about γ_c . The main reasons for this are that in the new formula the dependence of γ_c on $\mathcal{L}$ and $\mathcal{H}$ is much simpler, and that the formula retains the same structure as in the non-causal case of Eq. (8).

4. Some Special Cases

In this section we shall use the result of Theorem 2 to compute γ_c for the two important special cases of equalization (or its dual problem tracking) and filtering signals from additive noise.

4.1. The Equalization Problem

In the equalization problem we have $\mathcal{L} = I$. Referring back to Fig. 1, this means that we would like to estimate the unknown input signal $\{u_i\}$ from noisy measurements of the output of the linear time-invariant system $\mathcal{H}$. In this sense, the causal equalization problem can be regarded as the problem of causally inverting a linear system in the presence of additive noise. [For further details see [4].]

Theorem 3 (H^∞ Equalization) Consider the causal Laurent operator $\mathcal{H}$, where the impulse response $\{H_i\}$ is a sequence of elements of $\mathbb{C}^{m \times m}$, and suppose we want to find γ_c , where

$$\gamma_c = \inf_{\text{causal } \mathcal{K}} \left\| \begin{bmatrix} I - \mathcal{K}\mathcal{H} & -\mathcal{K} \end{bmatrix} \right\|_\infty.$$

(i) If $\mathcal{H}$ is minimum phase, i.e., if $\mathcal{H}^{-1}$ is causal, then

$$\gamma_c^2 = \gamma_s^2 = \sup_{\omega \in [0, 2\pi]} \bar{\sigma} \left[(I + H^*(e^{j\omega})H(e^{j\omega}))^{-1} \right]. \quad (18)$$

(ii) If $\mathcal{H}$ is non-minimum phase, i.e., if $\mathcal{H}^{-1}$ is non-causal, then

$$\gamma_c^2 = 1. \quad (19)$$

In other words, if the system $\mathcal{H}$ has a causal inverse then causal equalizers offer the same performance as non-causal equalizers that have access to future observations. On the other hand, if $\mathcal{H}$ does not have a causal inverse then causal equalization is not possible from the H^∞ point of view, since $\gamma = 1$ is the disturbance attenuation obtained by performing no equalization at all! Indeed $\mathcal{K} = 0$ yields

$$\mathcal{T}_\mathcal{K} = \begin{bmatrix} I & 0 \end{bmatrix}, \quad (20)$$

so that $\|\mathcal{T}_\mathcal{K}\|_\infty = 1$.

Proof of Theorem 3: To prove part (i) assume that $\mathcal{H}$ is minimum phase. We will show that

$$\gamma > \gamma_s \Rightarrow \gamma > \gamma_c$$

which establishes $\gamma_s \geq \gamma_c$. But since we already know that $\gamma_s \leq \gamma_c$ this readily implies that $\gamma_s = \gamma_c$.

Now since $\mathcal{H}$ is minimum phase, $H(z)$ has no unit circle zeros and hence $\gamma_s < 1$. Choose now a γ such that $1 > \gamma > \gamma_s$. We thus have the following series of arguments: $1 > \gamma > \gamma_s$

$$\begin{aligned} &\Rightarrow \gamma^2 I > (I + \mathcal{H}^* \mathcal{H})^{-1} \\ &\Rightarrow (\gamma^{-2} - 1)I < \mathcal{H}^* \mathcal{H} \\ &\Rightarrow \frac{1}{\gamma^{-2} - 1} I > \mathcal{H}^{-1} \mathcal{H}^{-*} \\ &\Rightarrow \frac{1}{\gamma^{-2} - 1} I > \left[\begin{array}{c|c} \mathcal{H}_-^{-1} & 0 \\ \times & \times \end{array} \right] \left[\begin{array}{c|c} \mathcal{H}_-^{-*} & \times \\ 0 & \times \end{array} \right] \\ &\Rightarrow \frac{1}{\gamma^{-2} - 1} \left[\begin{array}{c|c} I & 0 \\ 0 & I \end{array} \right] > \left[\begin{array}{c|c} \mathcal{H}_-^{-1} \mathcal{H}_-^{-*} & \times \\ \times & \times \end{array} \right] \\ &\Rightarrow \frac{1}{\gamma^{-2} - 1} I > \mathcal{H}_-^{-1} \mathcal{H}_-^{-*}, \end{aligned}$$

where $\times$ denotes irrelevant entries. We can now unwind this last expression to get $\gamma^2 I > (I + \mathcal{H}_-^* \mathcal{H}_-)^{-1}$. But this implies that $\gamma > \gamma_c$, and we are done. [Note that the fourth step in the above argument is the key step, since we used the fact that $\mathcal{H}^{-1}$ is causal, and hence lower triangular.]

To prove part (ii) assume now that $\mathcal{H}^{-1}$ is non-causal. This implies that $H(z)$ has a zero outside the unit circle. Let p ($|p| > 1$) be such a zero of $H(z)$. Then

$$\underbrace{\begin{bmatrix} \ddots & & & \\ \ddots & H_0 & & \\ \cdots & H_1 & H_0 & \\ \cdots & H_2 & H_1 & H_0 \end{bmatrix}}_{\mathcal{H}_-} \underbrace{\begin{bmatrix} \vdots \\ p^{-2} \\ p^{-1} \\ 1 \end{bmatrix}}_X = H(p) \begin{bmatrix} \vdots \\ p^{-2} \\ p^{-1} \\ 1 \end{bmatrix} = 0. \quad (21)$$

Note that $X \in l_2^-$ since $|p| > 1$. The above equation shows that X is an eigenvector of $\mathcal{H}_-$ with eigenvalue $H(p)$. Now

$$(I + \mathcal{H}_-^* \mathcal{H}_-)^{-1} X = (I - \mathcal{H}_-^* (I + \mathcal{H}_- \mathcal{H}_-^*)^{-1} \mathcal{H}_-) X = X,$$

so that $(I + \mathcal{H}_-^* \mathcal{H}_-)^{-1}$ has an eigenvalue of unity. Since $(I + \mathcal{H}_-^* \mathcal{H}_-)^{-1} \leq I$ this implies the desired result that

$$\|(I + \mathcal{H}_-^* \mathcal{H}_-)^{-1}\|_\infty = 1. \quad \blacksquare$$

4.2. Filtering Signals from Additive Noise

In this problem we have $\mathcal{L} = \mathcal{H}$. Referring back to Fig. 1, this means that we would like to estimate the unknown signal $\{s_i\}$ from the "signal-plus-noise" observations $\{y_i = s_i + v_i\}$. The result is given below. [For further details see [5].]

Theorem 4 (H^∞ Filtering of Signals in Noise)

Consider the causal Laurent operator $\mathcal{H}$, and suppose we want to find γ_c , where

$$\gamma_c = \inf_{\text{causal } \mathcal{K}} \left\| \begin{bmatrix} \mathcal{H} - \mathcal{K}\mathcal{H} & -\mathcal{K} \end{bmatrix} \right\|_\infty.$$

Then we have, $\gamma_c = \gamma_s$, i.e.,

$$\gamma_c^2 = \sup_{\omega \in [0, 2\pi]} \bar{\sigma} \left[H(e^{j\omega}) (I + H^*(e^{j\omega})H(e^{j\omega}))^{-1} H^*(e^{j\omega}) \right].$$

Proof: We will show that

$$\gamma > \gamma_s \Rightarrow \gamma > \gamma_c$$

which establishes $\gamma_s \geq \gamma_c$. But since we already know that $\gamma_s \leq \gamma_c$ this readily implies that $\gamma_s = \gamma_c$.

First note that since $H(z)$ is analytic on $|z| = 1$, $H(e^{j\omega})$ is bounded for all $\omega \in [0, 2\pi]$, and hence $\gamma_s < 1$. Choose now a γ such that $1 > \gamma > \gamma_s$. We thus have the following series of arguments: $1 > \gamma > \gamma_s$

$$\begin{aligned} &\Rightarrow \gamma^2 I > \mathcal{H} (I + \mathcal{H}^* \mathcal{H})^{-1} \mathcal{H}^* = I - (I + \mathcal{H} \mathcal{H}^*)^{-1} \\ &\Rightarrow (I + \mathcal{H} \mathcal{H}^*)^{-1} > (1 - \gamma^2) I \\ &\Rightarrow I + \mathcal{H} \mathcal{H}^* < \frac{1}{1 - \gamma^2} I \\ &\Rightarrow \left[\begin{array}{c|c} I + \mathcal{H}_- \mathcal{H}_-^* & \times \\ \times & \times \end{array} \right] < \frac{1}{1 - \gamma^2} \left[\begin{array}{c|c} I & 0 \\ 0 & I \end{array} \right] \\ &\Rightarrow I + \mathcal{H}_- \mathcal{H}_-^* < \frac{1}{1 - \gamma^2} I, \end{aligned}$$

where $\times$ denotes irrelevant entries. We can now unwind this last expression to get $\gamma^2 I > \mathcal{H}_-(I + \mathcal{H}_-^* \mathcal{H}_-)^{-1} \mathcal{H}_-^*$. But this implies that $\gamma > \gamma_c$, and we are done. ■

5. Worst-Case Non-Estimability and Worst-Case Complete Estimability

As mentioned earlier, explicit frequency-domain expressions for γ_c are currently not available. However, in general one can always claim that

$$\gamma_s \leq \gamma_c \leq \gamma_n \triangleq \|\mathcal{L}\|_\infty, \quad (22)$$

where the upper bound follows from the fact that γ_n corresponds to performing no estimation, i.e., $\mathcal{K} = 0$, since for this choice of $\mathcal{K}$ we have $\mathcal{T}_{\mathcal{K}} = \begin{bmatrix} \mathcal{L} & 0 \end{bmatrix}$.

Therefore a natural question to ask is what are the conditions on $\mathcal{L}$ and $\mathcal{H}$ for γ_c to achieve either of the above upper and lower bounds? This is also an important question since:

- $\gamma_c = \gamma_s$ corresponds to an *easy* estimation problem, since here causal estimators have the same performance as noncausal ones.
- $\gamma_c = \gamma_n$ corresponds to a *difficult* estimation problem, since here causal estimators cannot offer any improvement over not estimating at all ($\mathcal{K} = 0$).

Due to the importance of the above concepts, we shall call the pair $\{\mathcal{L}, \mathcal{H}\}$ *worst-case complete estimable* if $\gamma_c = \gamma_s$ and *worst-case non-estimable* if $\gamma_c = \gamma_n$.

From our analysis of the previous section we already know the answer to whether $\{\mathcal{L}, \mathcal{H}\}$ is worst-case complete estimable or worst-case non-estimable for the special cases of equalization and filtering signal from additive noise. Unfortunately, we do not currently have a simple characterization of when the arbitrary pair $\{\mathcal{L}, \mathcal{H}\}$ is worst-case complete estimable. However, for worst-case non-estimability we do have such a characterization, as described next.

5.1. Worst-Case Non-Estimability

Theorem 5 (Worst-Case Non-Estimability)

Consider the setting of Problem 1, and suppose that $\gamma_s < \gamma_n = \|\mathcal{L}\|_\infty$. Then we have the following results.

- (i) $\gamma_c = \gamma_n$, iff., there exists an $x \in l_-^2$ such that
 - (a) $\|\mathcal{L}_- x\|_2 = \|\mathcal{L}\|_\infty \|x\|_2$.
 - (b) $\mathcal{H}_- x = 0$.
- (ii) If the corresponding z -transforms $L(z)$ and $H(z)$ are scalar and rational, then $\gamma_c = \gamma_n$, iff.,
 - (a) $L(z)$ is all-pass, i.e., $|L(e^{j\omega})| = \text{constant}$, for all $\omega \in [0, 2\pi]$.
 - (b) The number of non-minimum phase zeros of $H(z)$ (counting multiplicities) is greater than the McMillan degree of $L(z)$.

The above theorem gives a very simple characterization of worst-case non-estimability. This is quite useful since it will allow us to recognize the occurrence of worst-case non-estimability in various applications, and to design estimation (and control) scenarios to avoid it. We shall presently demonstrate this by considering the problem of equalizing a scalar rational LTI system, $H(z)$.

5.2. Equalization with Delay

First note that Theorem 5 confirms the result of Theorem 3, part (ii), on equalization. Indeed in the equalization problem $L(z) = 1$ is clearly all-pass and has McMillan degree zero. Thus if $H(z)$ has any non-minimum phase zero, the equalization problem is worst-case non-estimable and $\gamma_c = \gamma_n = 1$.

One may now speculate whether it is possible to causally equalize $H(z)$ by allowing a finite amount of delay. In this case, one would like to estimate u_{i-d} , for some $d > 0$, using the observations $\{y_j, j \leq i\}$. Mathematically, this corresponds to choosing $L(z) = z^{-d}$ in the two-block problem of Sec. 2. The natural question to ask is what is the delay necessary to guarantee $\gamma_c < \gamma_n = 1$?

Lemma 1 (Equalization with Delay) Consider the Laurent operator with scalar rational transfer function $H(z)$, and define

$$\gamma_c = \inf_{\text{causal } K(\cdot)} \left\| \begin{bmatrix} z^{-d} - K(z)H(z) & -K(z) \end{bmatrix} \right\|_\infty.$$

Then if the number of non-minimum phase zeros of $H(z)$ is given by n , we have:

- (i) $\gamma_c < 1$ if $d \geq n$.
- (ii) $\gamma_c = 1$ if $d < n$.

Proof: The lemma follows from Theorem 5, part (ii), since $L(z) = z^{-d}$ is all-pass and has McMillan degree d . ■

Therefore the minimum amount of delay is given by the number of non-minimum phase zeros of $H(z)$.

References

- [1] M. Verma and E.A. Jonckheere. l^∞ compensation with mixed sensitivity as a broadband matching problem. *Systems and Control Letters*, 4:125–129, 1984.
- [2] A. Feintuch and B.A. Francis. Uniformly optimal control of linear systems. *Automatica*, 21:563–574, 1985.
- [3] J.C. Doyle, K. Glover, P.P. Khargonekar, and B.A. Francis. State-space solutions to standard H^2 and H^∞ control problems. *IEEE Trans. on Automatic Control*, 34:831–847, 1989.
- [4] B. Hassibi, A.E. Erdogan, and T. Kailath. Equalization with an H^∞ criterion. *Submitted to IEEE Transactions on Information Theory*, 1997.
- [5] B. Hassibi and T. Kailath. H^∞ filtering of signals in additive noise. *Submitted to IEEE Transactions on Signal Processing*, 1998.