A General Time-Varying Estimation and Control Problem

Avraham Feintuch [†*] and Alexander Markus[†]

Abstract. This paper solves two problems raised in [HSK]. The first is a uniform norm time-varying linear estimation problem and the second is a general time-varying linear control problem.

Key words. Time-varying linear systems, Estimation, Control, H-infinity, Nest algebras, Four-block problems.

[†] Department of Mathematics, Ben-Gurion University of the Negev, Beer-Sheva, 84105, Israel
e-mail: abie@math.bgu.ac.il, markus@math.bgu.ac.il

[*] Correspondent author. Telephone number: 972-8-6461613, fax number: 972-8-6477648.

1. Introduction

We consider here two problems described by Hassibi and Kailath ([HSK], Section 10.2 and Subsection 11.1.2; [H], [HK1], [HK2], [HK3]). The first is a general estimation problem:

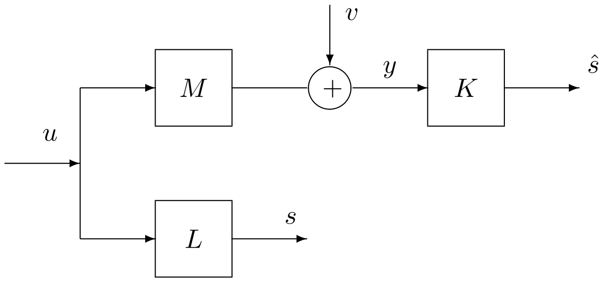

Fig. 1

M and L are known causal linear transfer operators which are in general time-varying. $u = \{u_j\}$ is an unknown input sequence, often representing a disturbance to the linear systems represented by M and L respectively. The output $s = \{s_j\}$ of L is unobservable, and the observable sequence $y = \{y_j\}$ is a measurement of the output of M, corrupted by an unknown disturbance $v = \{v_j\}$. The objective is to design a causal, linear, possibly time-varying transfer operator K that estimates s using y. The estimates will be denoted by $\hat{s} = \{\hat{s}_j\}$ and the estimation errors are $\tilde{s}_j = s_j - \hat{s}_j$. We assume that all the sequences are finite energy, namely that $u,\ v,\ s,\ y,\ \hat{s}$ are vectors in a appropriate Hilbert sequence space $\mathcal{H}$. The transfer operator that maps the unknown input pair $\begin{bmatrix} u \\ v \end{bmatrix}$ to the estimation error $\hat{s}$ is given by

$$\hat{s} = (L - KM)u - Kv = [L - KM \quad -K] \begin{bmatrix} u \\ v \end{bmatrix}.$$

We relate to $T_K = [L - KM \quad -K]$ as a linear transformation from $\mathcal{H} \oplus \mathcal{H}$ into $\mathcal{H}$.

The goal of the estimation is to make the transfer operator T_K small in some sense. Here we consider what is referred to in [HSK] as H^∞- estimation. This terminology is entirely appropiate for time-invariant systems, where the H^∞ norm is equivalent to the induced operator norm for H^∞ transfer operators. However this terminology has no meaning for time-varying operators. The analogous problem for time-varying systems is the minimization of T_K in the standard induced operator norm on the appropriate Hilbert sequence space $\mathcal{H}$.

The problem considered here is the heart of Problems 10.4.1 and 10.4.2 of [HSK]:

$$\text{Find } \gamma_{opt} = \inf\{\|T_K\| \ : \ K \text{ stable}\}.$$

Hassibi and Kailath give formulas for γ_{opt} in the following situations:

(1) The non-causal case: $L,\ M,\ K \in \mathcal{L}(\mathcal{H})$, the algebra of all bounded linear operators on $\mathcal{H}$ (Theorem 10.4.1).

(2) The time-invariant case: $L,\ M,\ K \in H^\infty$ (Theorem 10.4.2).

(3) The "finite-horizon" case: $L,\ M,\ K$ are lower triangular finite block matrices where each block is finite-dimensional (Theorem 10.4.2). This is the time-varying case for finite-input sequences.

As the authors of [HSK] point out, while similar problems have been considered by other authors (see [FF], [F]), their formulas are new and have the advantage that they depend only on the given data, the pair $\{L,\ M\}$. Our main result, using a completely different approach, generalizes the formula of Theorem 10.4.2 of [HSK] to the "infinite-horizon" time-varying case. We show how our approach leads immediately to (1), the formula of Theorem 10.4.1. While we don't derive the formula of (2) it is clear that it can be obtained using exactly the same approach.

The main idea is to use inner-outer factorizations for operator in nest algebras developed by W. Arveson [A] (see [F], Section 4.2) to reduce the problem of computing γ_{opt} to a classical two-block problem of the type first considered in [FF] and then use the special properties of the particular factorization obtained for the estimation problem.

The second problem we consider from [HSK] is the "Measurement Feedback Control Problem":

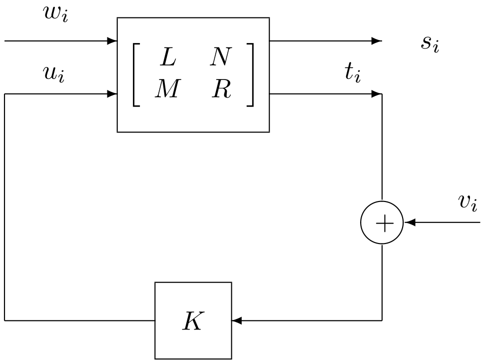

Fig. 2

Given known causal operators $L,\ M,\ N,\ R$ possibly time-varying, that map the input sequences $w = \{w_i\},\ u = \{u_i\}$ to the output sequences $s = \{s_i\},\ t = \{t_i\}$ according to the formula

$$\begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} L & N \\ M & R \end{bmatrix} \begin{bmatrix} w \\ u \end{bmatrix}.$$

The sequences w and u are exogenous signals. w can be considered process noise and v the measurement noise that corrupt the output signal t. The signal u is the control output used to influence the dynamic behavior of the plant. The relationship between y and u is given by $u = Ky$ where the controller K is a causal, linear, possibly time-varying, transfer operator. s is the regulated output which must be kept small. In addition, it is required that the control signal u be kept small.

We consider the transfer matrix T_K from $\begin{bmatrix} w \\ v \end{bmatrix}$ to $\begin{bmatrix} s \\ u \end{bmatrix}$. Since

$$s = Lw + Nu$$

$$y = Mw + Ru + v$$

we obtain that

$$T_K = \begin{bmatrix} L + NK(I - RK)^{-1} & NK(I - RK)^{-1} \\ K(I - RK)^{-1}M & K(I - RK)^{-1} \end{bmatrix}.$$

Appropriate stabilization assumptions (see [HSK], Subsection 11.2.4) give that $Q = K(I - RK)^{-1}$ is stable and that K and Q determine each other uniquely. Thus T_K can be replaced by

$$T_Q = \begin{bmatrix} L + NQM & NQ \\ QM & Q \end{bmatrix} = \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} N \\ I \end{bmatrix} Q[M \quad I].$$

The measurement feedback control problem is:

$$\text{Find } \beta_{opt} = \inf\{\|T_Q\| \ : \ Q \text{ stable}\}$$

(see Problem 11.3.1 of [HSK]).

In this case we start from scratch. No formulas are given in [HSK] for β_{opt} for any of the cases (1) – (3).

We will give a formula with a complete proof for the infinite-horizon time-varying case and obtain the non-causal case as a direct corollary. We will state without proof (since conceptually the proof is the same as the one given) the formula for the time-invariant case.

Up to this point we have related only to the formulas for γ_{opt}, β_{opt}. Other issues raised in [HSK] are:
(a) the existence of controllers for which γ_{opt} and β_{opt} are obtained;
(b) the parametrized description of all solutions to the suboptimal problems:

$$\begin{array}{l} \|T_K\| < \gamma \text{ for } \gamma > \gamma_{opt} \\ \\ \|T_Q\| < \beta \text{ for} \beta > \beta_{opt} \end{array} .$$

The issue of (a) is easily resolved using weak-compactness arguments identical to those given in [F, Chapter 7].

The issue of (b) is more complex. The solution to the four-block time-varying Nehari problem ([F], Chapter 7) plays an important role in our analysis (see Section 5). It was shown in [FFGK] (Section XIII.6) that this problem is equivalent to the "three chains completion problem" which is a time-varying version of the Sz.-Nagy – Foias commutant lifting theorem. Chapter XIV of [FFGK] gives, by means of a deep analysis, a parameteric formula for the set of all solutions to the suboptimal three chains completion problem. It would be of significant interest to adapt this formula to the special case considered here. This is beyond the scope of this paper.

2. Preliminaries

Let $\mathcal{H}$ be a separable Hilbert space. As it is well known ([F], Chapter 5) the physical notion of causality for linear systems is formulated for linear transformations in terms of leaving invariant a totally ordered set of closed subspaces of $\mathcal{H}$. In this paper we will deal with three concrete situations which we describe in terms of the chains of orthogonal projections associated with these subspaces (all of them are discrete):
(1) Finite chains: $\mathcal{P} = \{0 = P_0 < P_1 < ... < P_N = I\}$.
(2) Semi-infinite chains: $\mathcal{P} = \{0 = P_0 < P_1 < ... < P_n < P_{n+1} < ...;\ I\}$
(3) Double-infinite chains: $\mathcal{P} = \{0; ... < P_{-m} < P_{-m+1} < ... < P_0 < ...P_{n+1};\ I\}$

We assume also that in cases (2) and (3) $\lim_{n\to\infty} P_n = I$ and in case (3) $\lim_{n\to-\infty} P_n = 0$. Note that in all three cases the corresponding nest of subspaces are complete (see the general definition in [F], p. 47). While at this stage we don't impose any assumptions on $\dim(P_{n+1} - P_n)$, our main results will relate to the case where $\dim(P_{n+1} - P_n) < \infty$ for all n.

Having fixed $\mathcal{P}$, an operator A is causal if $P_i A P_i = P_i A$ for all $P_i \in \mathcal{P}$ and is stable if it causal and bounded. The family of stable operators, denoted by $\mathcal{S}$, is a weakly closed algebra containing the identity, refered to in the operator theory literature as a nest algebra. In terms of the coordinate spaces determined by $\mathcal{P}$, the stable operators are the bounded operators whose matrix representations are lower triangular.

The time-varying estimation problem is:

$$\text{Given } L,\ M \in \mathcal{S},\ \text{find } \gamma_{opt} = \inf\{\|[L - KM \quad -K]\| \ :\ K \in \mathcal{S}\}.$$

The other problems considered here will be formulated in a similar fashion.

3. Detectibility

In general the given causal linear systems L and M are not necessarily stable. This means that L and M may be unbounded operators on $\mathcal{H}$. In this case the problem of minimizing $\|T_K\|$ may not make sense. For this reason the notion of detectability was introduced in [HSK].

Definition. The causal pair $\{L,\ M\}$ is detectable if there is $K \in \mathcal{S}$ such that $L - KM \in \mathcal{S}$.

While the theory of robust control and estimation considers unstable systems the viability of this theory requires that all systems be stabilizable (see [F], Chapter 6). In particular they must have left and right fractional representations. This means that there exist $A, B,\ S,\ T \in \mathcal{S}$ such that $L = B^{-1}A,\ V = ST^{-1}$ (where $B^{-1},\ T^{-1}$ are the in general unbounded inverses of B and T respectively). We have the following:

Theorem 1. *There exists $K \in \mathcal{S}$ such that $L - KM \in \mathcal{S}$ if and only if $L \in \mathcal{S}$.*

Proof. If L is stable, just take $K = 0$. If there exists $K \in \mathcal{S}$ such that $L - KM = C \in \mathcal{S}$, then with $L = B^{-1}A,\ M = ST^{-1},\ B^{-1}A - KST^{-1} = C$ or $B^{-1}(AT - BKS)T^{-1} = C$. Thus $BKS = (A - BC)T$ or $KS = B^{-1}A - CT = L - CT$. Therefore $L = KS + CT$ is bounded. $\Box$

It is obvious that $L - KM \in \mathcal{S}$ for all $K \in \mathcal{S}$ if and only if $L,\ M \in \mathcal{S}$, and we will assume this to be the case from this point on.

4. Computation of γ_{opt}

The classical two-block problem studied in [FF] is:

$$\text{Given } A,\ B \in \mathcal{L}(\mathcal{H}) \text{ compute } \inf\{\|[A - K \quad B]\| \ :\ K \in \mathcal{S}\}.$$

The formula obtained there will apply easily to the case where $B = [B_1 \quad B_2]$ with $B_1,\ B_2 \in \mathcal{S}(\mathcal{H})$ (or, for that matter, for any number of blocks). The first stage of our analysis is the reduction of the general estimation problem to a two-block problem.

Note that

$$\|[L - KM \quad -K]\| = \|[L \quad 0] - K[M \quad I]\| = \left\| \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} K & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} M & I \\ 0 & 0 \end{bmatrix} \right\|.$$

We will use the inner-outer factorization of the operator $\begin{bmatrix} M^* & 0 \\ I & 0 \end{bmatrix}$ with respect to the nest algebra defined by the sequence $\left\{ \begin{bmatrix} P_n & 0 \\ 0 & P_n \end{bmatrix} \right\}$ as given in Example 4.2.11 of [F]. We summarize what we need for our purposes in the following lemma.

Lemma 1. *For $M \in \mathcal{S}$ there exist $B,\ W_1,\ W_2 \in \mathcal{S}$ with the following properties:*
(1) $W_1W_1^* + W_2W_2^* = I$.
(2) For each $n,\ B^*P_n\mathcal{H} = \operatorname{ran} B_n^* \cap P_n\mathcal{H}$.
(3) If R_{B^*} *is the orthogonal projection on* $\operatorname{ran} B^*$, *then for each* $n,\ R_{B^*}P_n = P_nR_{B^*}$.
(4) $[M \quad I] = B[W_1 \quad W_2]$.

The factorization $\begin{bmatrix} M^* \\ I \end{bmatrix} = \begin{bmatrix} W_1^* \\ W_2^* \end{bmatrix} B^*$, the dual of (4), is a special case of the well known inner-outer factorization for nest algebras ([F], Section 4.2). It follows from (4) that $W_2^*B^* = I$. It will be decisive for our analysis of the problems considered there that B is invertible. A sufficient condition for this to hold is that for each $n,\ \dim(P_{n+1} - P_n) < \infty$. In the finite and semi-infinite case the verification that under this assumption B is invertible is straightforward. For causality gives that for each $n,\ B^*P_n\mathcal{H} \subset P_n\mathcal{H}$. Since $W_2^*B^* = I$ implies that Ker $B^* = \{0\}$, and, since $P_n\mathcal{H}$ is finite dimensional for all $n,\ B^*P_n\mathcal{H} = P_n\mathcal{H}$ for each n. Thus ran B^* is dense in $\mathcal{H}$. Since B^* is left invertible, it is in fact invertible and thus so is B.

The doubly infinite case is more delicate. We begin with some elementary facts wich we prove for the convenience of the reader.

Lemma 2. *Let $\mathcal{N}_1$ and $\mathcal{N}_2$ be subspaces of $\mathcal{H}$ of dimension $n < \infty$. If there exists a non-zero vector $x \in \mathcal{N}_1$ such that $x \perp \mathcal{N}_2$, then there exists a non-zero vector $y \in \mathcal{N}_2$ such that $y \perp \mathcal{N}_1$.*

Proof. Let $\hat{\mathcal{N}_1} = \mathcal{N}_1 \ominus \{\alpha x \ : \ \alpha \in \mathbf{C}\}$. Then $\hat{\mathcal{N}_1}^\perp$ has co-dimension $n-1$ and there exists a non-zero $y \in \hat{\mathcal{N}_1}^\perp \cap \mathcal{N}_2$. Then $y \perp \hat{\mathcal{N}_1}$, and also $y \perp x$, so $y \perp \mathcal{N}_1$. $\square$

Lemma 3. *Let $\mathcal{N}_1, \ \mathcal{N}_2$ be subspaces of $\mathcal{N}$ with corresponding orthogonal projections $P_1, \ P_2$ respectively. If $P_1P_2 = P_2P_1$, then the subspaces $\mathcal{N}_1 \ominus (\mathcal{N}_1 \cap \mathcal{N}_2)$ and $\mathcal{N}_2$ are orthogonal.*

Proof. For $y_1 \in \mathcal{N}_1 \ominus (\mathcal{N}_1 \cap \mathcal{N}_2), \ y_2 \in \mathcal{N}_2$,

$$(y_1, \ y_2) = (P_1 y_1, \ P_2 y_2) = (y_1, \ P_1 P_2 y_2).$$

But $P_1P_2 = P_2P_1$ implies that $P_1P_2y_2 \in \mathcal{N}_1 \cap \mathcal{N}_2$, so $(y_1, \ P_1P_2y) = 0$. $\square$

Lemma 4. *Under the assumptions of Lemma 3, if $\mathcal{N}_1$ is not contained in $\mathcal{N}_2$, then there exists a non-zero vector $v \in \mathcal{N}_1$ such that $v \perp \mathcal{N}_2$*

Proof. Let $\mathcal{N}_3 = \mathcal{N}_1 \ominus (\mathcal{N}_1 \cap \mathcal{N}_2)$. Since $\mathcal{N}_1$ is not contained in $\mathcal{N}_2$, $\mathcal{N}_3 \neq \{0\}$. By Lemma 3, $\mathcal{N}_3 \perp \mathcal{N}_2$.

Theorem 2. *Let $\mathcal{P} = \{P_n\}_{-\infty}^{\infty}$ be a doubly infinite complete chain of projections such that $\dim(P_n - P_{n-1}) < \infty$ for all n. Then the operator B from Lemma 1 is invertible in $\mathcal{S}$.*

Proof. Since $BW_2 = I$ and $W_2 \in \mathcal{S}$ it is sufficient to prove that $\operatorname{ran} B^*$ is dense in $\mathcal{H}$.

By Lemma 1, $B^*P_n\mathcal{H} = \operatorname{ran} B^* \cap P_n\mathcal{H}$ for all n. This means that if $x \in \operatorname{ran} B^* \cap P_n\mathcal{H}$, there exists $z \in P_n\mathcal{H}$ such that $B^*z = x$. Suppose y is a non-zero vector in $(P_n - P_{n-1})\mathcal{H}$. Then $B^*y = P_{n-1}B^*y + (P_n - P_{n-1})B^*y$. We show $(P_n - P_{n-1})B^*y \neq 0$. If $B^*y = P_{n-1}B^*y$ then $B^*y \in \operatorname{ran} B^* \cap P_{n-1}\mathcal{H}$. Thus there exists $z \in P_{n-1}\mathcal{H}$ with $B^*z = B^*y$. Since $\operatorname{Ker} B^* = \{0\}$, $z = y$. But $z \perp y$, so $(P_n - P_{n-1})B^*y \neq 0$.

We show that for every $n \in \mathbf{Z}$, $(P_n - P_{n-1})\mathcal{H} \subset \operatorname{ran} B^*$. For suppoze this inclusion doesn't hold for some n. It follows from (3) of Lemma 1 and from Lemma 4 that there exists a non-zero $v \in (P_n - P_{n-1})\mathcal{H}$ such that $v \perp \operatorname{ran} B^*$. Let $\mathcal{M} = B^*(P_n - P_{n-1})\mathcal{H}$. Then $\dim \mathcal{M} = \dim(P_n - P_{n-1})\mathcal{H}$ and $v \perp \mathcal{M}$. By Lemma 2 there exists a non-zero vector $u \in \mathcal{M}$ such that $u \perp (P_n - P_{n-1})\mathcal{H}$. But $u \in \mathcal{M}$ implies that $u = B^*y$ for some $y \in (P_n - P_{n-1})\mathcal{H}$ and by the previous paragraph $u = P_{n-1}u + (P_n - P_{n-1})u$ with $(P_n - P_{n-1})u \neq 0$. This contradiction shows that $(P_n - P_{n-1})\mathcal{H} \subset \operatorname{ran} B^*$ for every $n \in \mathbf{Z}$. Therefore $\operatorname{ran} B^*$ is dense in $\mathcal{H}$ and the rest follows as above. $\square$

The requirement that $\dim(P_n - P_{n-1}) < \infty$ for all n is not necessary for B in the factorization $[M \quad I] = B[W_1 \quad W_2]$ to be invertible. For consider the trivial chain $\mathcal{P} = \{0, I\}$ which corresponds to the non-causal case. In this case such a factorization is given by $B = (I + MM^*)^{\frac{1}{2}}$, $W_1 = (I + MM^*)^{-\frac{1}{2}}M$, $W_2 = (I + MM^*)^{-\frac{1}{2}} = B^{-1}$. For this chain Theorem 10.4.1 of [HSK] gives a formula for γ_{opt}. We give a new proof of this theorem which highlights the basic strategy used in our more general results.

Theorem 3. *For $L, \ M \in \mathcal{L}(\mathcal{H})$,*

$$\inf\{\|[L - KM \quad -K]\| \ : \ K \in \mathcal{L}(\mathcal{H})\} = \|L(I + M^*M)^{-\frac{1}{2}}\|.$$

Proof. Let $W_1 = (I + MM^*)^{-\frac{1}{2}}M$, $W_2 = (I + MM^*)^{-\frac{1}{2}}$, $W = [W_1 \quad W_2]$. Then W and $[W^* \quad I - W^*W]$ are easily seen to be co-isometries, and for $B = (I + MM^*)^{\frac{1}{2}}$, $BW = [M \quad I]$. Then

$$\begin{aligned} \|[L - KM \quad -K]\| &= \|[L \ \ 0] - K[M \ \ I]\| = \|[L \ \ 0] - KBW\| = \|\{[L \ \ 0] - KBW\}[W^* \quad I - W^*W]\| \\ &= \|[LW_1^* - KB \quad L(I - W_1^*W_1) \quad -LW_1^*W_2]\| \end{aligned} .$$

Then

$$\gamma_{opt} = \inf\{[\|[LW_1^* - KB \;\; L(I - W_1^*W_1) \;\; -LW_1^*W_2]\| : \;\; K \in \mathcal{L}(\mathcal{H})\}.$$

This infimum is clearly obtained for $LW_1^* - KB = 0$, or $K = LW_1^*B^{-1} = LM^*(I + MM^*)^{-1}$. In this case,

$$\gamma_{opt} = \|[L\{I - M^*(I + MM^*)^{-1}M\} \;\; -LM^*(I + MM^*)^{-1}]\| = \|[L(I + M^*M)^{-1} \;\; -L(I + M^*M)^{-1}M^*]\|.$$

We simplify this formula:

$$\gamma_{opt}^2 = \left\| \begin{bmatrix} (I + M^*M)^{-1}L^* \\ -M(I + M^*M)^{-1}L^* \end{bmatrix} \right\|^2$$

$$= \sup_{\|f\|=1} \{(L(I + M^*M)^{-2}L^*f, f) + (L(I + M^*M)^{-1}M^*M(I + M^*M)^{-1}L^*f, f)\}$$

$$= \sup_{\|f\|=1} ([L(I + M^*M)^{-1}(I + M^*M)(I + M^*M)^{-1}L^*f, f) = \|L(I + M^*M)^{-1}L^*\|$$

$$= \|L(I + M^*M)^{-\frac{1}{2}}\|^2.$$

□

It will be striking that the formulas in the causal case will be analogous to the formula of Theorem 3.

Theorem 4. *Given the factorization $[M \;\; I] = B[W_1 \;\; W_2]$ from Lemma 1 with B invertible. Then*

$$\gamma_{opt} = \sup_n \|P_nLP_n(P_n + P_nM^*P_nMP_n)^{-\frac{1}{2}}\|$$

where the inverse is understood to be the inverse of the restriction of $P_n + P_nM^*P_nMP_n$ to $P_n\mathcal{H}$.

Proof. We begin as in the proof of Theorem 3. For $W = [W_1 \;\; W_2]$, $[W^* \;\; I - W^*W]$ is a co-isometry and

$$\gamma_{opt} = \inf\{\|[LW_1^* - KB \;\; L(I - W_1^*W_1) \;\; -LW_1^*W_2]\| \; : \; K \in \mathcal{S}\}.$$

Since B is invertible and $B^{-1} = W_2 \in \mathcal{S}$,

$$\{KB \; : \; K \in \mathcal{S}\} = \mathcal{S},$$

and therefore

$$\gamma_{opt} = \inf\{\|[LW_1^* - K \;\; L(I - W_1^*W_1) \;\; -LW_1^*W_2]\| \; : \; K \in \mathcal{S}\}.$$

This is the standard two-block problem of [FF] and its solution is given by $\gamma_{opt} = \sup_n\{\gamma_n\}$ where

$$\gamma_n = \|[P_nLW_1^*(I - P_n) \;\; P_nL(I - W_1^*W_1) \;\; -P_nLW_1^*W_2]\| = \left\| \begin{bmatrix} (I - P_n)W_1L^*P_n \\ (I - W_1^*W_1)L^*P_n \\ -W_2^*W_1L^*P_n \end{bmatrix} \right\|.$$

Then,

$$\gamma_n^2 = \sup_{\|f\|=1} \{\|(I - P_n)W_1L^*P_nf\|^2 + \|(I - W_1^*W_1)L^*P_nf\|^2 + \|W_2^*W_1L^*P_nf\|^2\}$$

$$= \sup_{\|f\|=1} ([P_nLW_1^*(I - P_n)W_1L^*P_n + P_nL(I - W_1^*W_1)^2L^*P_n + P_nLW_1^*W_2W_2^*W_1L^*P_n]f, f)$$

$$= \sup_{\|f\|=1} (P_nL[W_1^*(I - P_n)W_1 + (I - W_1^*W_1)^2 + W_1^*W_2W_2^*W_1]L^*P_nf, f).$$

We simplify the expression in the square brackets:

$$W_1^*(I - P_n)W_1 + (I - W_1^*W_1)^2 + W_1^*W_2W_2^*W_1$$

$$= W_1^*(I - P_n)W_1 + I - 2W_1^*W_1 + W_1^*W_1W_1^*W_1 + W_1^*W_2W_2^*W_1$$

$$= W_1^*(I - P_n)W_1 + I - 2W_1^*W_1 + W_1^*(W_1W_1^* + W_2W_2^*)W_1$$

$$= W_1^*W_1 - W_1^*P_nW_1 + I - 2W_1^*W_1 + W_1^*W_1 = I - W_1^*P_nW_1.$$

So,

$$\gamma_n^2 = \sup_{\|f\|=1} (P_nL(I - W_1^*P_nW_1)L^*P_nf, f) = \|P_nLP_n(P_n - W_1^*P_nW_1)P_nL^*P_n\|.$$

To simplify, let $L_n = P_nLP_n,\ M_n = P_nMP_n,\ W_{1n} = P_nW_1P_n,\ W_{2n} = P_nW_2P_n$. Then

$$\gamma_n^2 = \|L_n(P_n - W_{1n}^*W_{1n})L_n^*\|.$$

We compute $P_n - W_{1n}W_{1n}$ as an operator acting on P_nH. By causality, $M = BW_1,\ I = BW_2$ imply that $M_n = B_nW_{1n},\ P_n = B_nW_{2n}$. Also $W_1W_1^* + W_2W_2^* = I$ gives $W_{1n}W_{1n}^* + W_{2n}W_{2n}^* = P_n$. Then

$$(P_n + M_n^*M_n)(P_n - W_{1n}^*W_{2n}) = (P_n + W_{1n}^*B_n^*B_nW_{1n})(P_n - W_{1n}^*W_{1n})$$

$$= P_n + W_{1n}^*B_n^*B_nW_{1n} - W_{1n}^*W_{1n} - W_{1n}^*B_n^*B_nW_{1n}W_{1n}^*W_{1n}$$

$$= P_n + W_{1n}^*(B_n^*B_n - P_n - B_n^*B_nW_{1n}W_{1n}^*)W_{1n}$$

$$= P_n + W_{1n}^*(B_n^*B_{1n}(P_n - W_{1n}W_{1n}^*) - P_n)W_{1n} = P_n + W_{1n}^*(B_n^*B_nW_{2n}W_{2n}^* - P_n)W_{1n}$$

$$= P_n + W_{1n}^*(B_n^*P_nW_{2n} - P_n)W_{1n} = P_n + W_{1n}^*(B_n^*W_{2n} - P_n)W_{1n}.$$

Since B is invertible, $BW_2 = I$ gives $W_2B = I$. Thus $W_{2n}B_n = P_n$ and $B_n^*W_{2n}^* = P_n$, so

$$(P_n + M_n^*M_n)(P_n - W_{1n}^*W_{1n}) = P_n.$$

A similar computation gives $(P_n - W_{1n}^*W_{1n})(P_n + M_n^*M_n) = P_n$, so $P_n - W_{1n}^*W_{1n}$ is the inverse of $P_n + M_n^*M_n$ on P_nH and

$$\gamma_n^2 = \|L_n(P_n + M_n^*M_n)^{-1}L_n^*\| = \|L_n(P_n + M_n^*M_n)^{-\frac{1}{2}}\|^2.$$

Thus $\gamma_{opt} = \sup_n\{\gamma_n\} = \sup_n \|L_n(P_n + M_n^*M_n)^{-\frac{1}{2}}\|$. □

Corollary 1. *The formula for* γ_{opt} *hold for all three cases: finite chains, semi-infinite chains, double infinite chains if* $\dim(P_n - P_{n-1}) < \infty$ *for all* n*.*

Remarks. The finite chain case with $\dim(P_n - P_{n-1}) < \infty$ for all n is Theorem 10.4.2 of [HSK]. This Theorem also gives the formula for the infinite horizon (semi-infinite) time-invariant case. In this case $L,\ M \in H^\infty$ and the factorization

$$[M \quad I] = B[W_1 \quad W_2]$$

is the classical outer∗-inner∗ factorization. In this case

$$\gamma_{opt} = \inf\{\|[L - KH \quad -K]\| \ :\ K \in H^\infty\} = \inf\{\|[LW_1^* - K \quad L((I - W_1^*W_1) \quad -LW_1^*W_2]\| \ :\ K \in H^\infty\}.$$

This is a generalized Nehari problem and is a special case of Theorem 3 of [FF]. If P is the orthogonal projection of $L^2 = H^2_- \oplus H^2$ onto H^2, then

$$\gamma_{opt} = \|[(I - P)LW_1^*P \quad (I - P)L(I - W_1^*W_1) \quad -(I - P)LW_1^*W_2]\|.$$

Using the argument of Theorem 3 we obtain

$$\gamma_{opt} = \|L_-(I + M_-^* M_-)^{-\frac{1}{2}}\|,$$

where $L_- = (I-P)L\,|\,H_-^2,\ M_- = (I-P)M\,|\,H_-^2$. This is formula (10.4.44) of Theorem 10.4.2 of [HSK].

5. Measurement Feedback Control

We now turn to our second problem:

$$\text{Find } \beta_{opt} = \inf\{\|T_Q\| \ :\ Q \in \mathcal{S}\}$$

where

$$T_Q = \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} N \\ I \end{bmatrix} Q[M \quad I]$$

with $L,\ M,\ N \in \mathcal{S}$. By Lemma 1 and Theorem 2

$$[M \quad I] = B[W_1 \quad W_2]$$

where $B,\ W_1,\ W_2 \in \mathcal{S},\ W_1W_1^* + W_2W_2^* = I$ and B is invertible with inverse W_2.

For $\begin{bmatrix} N \\ I \end{bmatrix}$ we state the dual results to Lemma 1 and Theorem 2.

Lemma 5. *For any $N \in \mathcal{S}$, there exist $A,\ V_1,\ V_2 \in \mathcal{S}$ with the following properties:*
(1) $V_1^*V_1 + V_2^*V_2 = I$.
(2) $A(I-P_n)\mathcal{H} = \operatorname{ran} A \cap (I-P_n)\mathcal{H}$ *for each* n.
(3) $R_AP_n = P_nR_A$ for each n.
(4) $\begin{bmatrix} N \\ I \end{bmatrix} = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} A$.

This follows from Example 4.2.11 of [F] applied to the operator $\begin{bmatrix} N & 0 \\ I & 0 \end{bmatrix}$ and the nest algebra determined by the sequence $\left\{\begin{bmatrix} I-P_n & 0 \\ 0 & I-P_n \end{bmatrix}\right\}$.

Theorem 5. *Suppose $\mathcal{P} = \{P_n\}$ is a semi-infinite or doubly infinite chain such that* $\dim(P_n - P_{n-1}) < \infty$ *for all n and let A be the operator from Lemma 5. Then A is invertible.*

The proof of Theorem 5 is obtained from the proof of Theorem 2 by replacing B^* with $A,\ I - P_{n-1}$ by P_n and P_{n-1} by $I - P_n$.

Just as for $W = [W_1 \quad W_2],\ [W^* \quad I - W^*W]$ is a co-isometry, for $V = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$, $\begin{bmatrix} V^* \\ I - VV^* \end{bmatrix}$ is an isometry. Thus

$$\|T_Q\| = \left\| \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} N \\ I \end{bmatrix} Q[M \quad I] \right\| = \left\| \begin{bmatrix} V^* \\ I - VV^* \end{bmatrix} \left\{ \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} + VAQBW \right\} [W^* \quad I - W^*W] \right\|$$

$$= \left\| \begin{bmatrix} V_1^*LW_1^* + AQB & \vdots & V_1^*L(I - W_1^*W_1) & -V_1^*LW_1^*W_2 \\ \cdots\cdots\cdots\cdots\cdots & \vdots & \cdots\cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots\cdots \\ (I - V_1V_1^*)LW_1^* & \vdots & (I - V_1V_1^*)L(I - W_1^*W_1) & -(I - V_1V_1^*)LW_1^*W_2 \\ -V_2V_1^*LW_1^* & \vdots & -V_2V_1^*L(I - W_1^*W_1) & V_2V_1^*LW_1^*W_2 \end{bmatrix} \right\|.$$

Since A and B are invertible, the infimum over all $Q \in \mathcal{S}$ is equal to

$$\inf \left\{ \left\| \begin{bmatrix} R+Q & Y \\ X & Z \end{bmatrix} \right\| \ : \ Q \in \mathcal{S} \right\}$$

where $R = V_1^* L W_1^*$ and $X,\ Y,\ Z$ are the blocks designated in the previous matrix. By Theorem 7.4.2 of [F] this is given by

$$\beta_{opt} = \sup_n \left\| \begin{bmatrix} P_n V_1^* L W_1^* (I - P_n) & P_n V_1^* L (I - W_1^* W_1) & -P_n V_1^* L W_1^* W_2 \\ (I - V_1 V_1^*) L W_1^* (I - P_n) & (I - V_1 V_1^*) L (I - W_1^* W_1) & -(I - V_1 V_1^*) L W_1^* W_2 \\ -V_2 V_1^* L W_1^* (I - P_n) & -V_2 V_1^* L (I - W_1^* W_1) & V_2 V_1^* L W_1^* W_2 \end{bmatrix} \right\|.$$

Theorem 6. *For any chain $\{P_n\}$ such that A and B defined in Lemma 5 and Lemma 1 respectively are invertible,*

$$\beta_{opt} = \sup_n \|(I + N(I - P_n)N^*)^{-\frac{1}{2}} L (I + M^* P_n M)^{-\frac{1}{2}}\|.$$

Proof. We denote the matrix inside the norm in the formula before Theorem 6 for β_{opt} by T_n and use the fact that $\|T_n\| = \|(T_n T_n^*)^{\frac{1}{2}}\|$. We write $T_n T_n^*$ as a 3×3 operator matrix with entries $D_{jk};\ j,\ k = 1,\ 2,\ 3.$ By direct computation, using the formula

$$W_1^*(I - P_n)W_1 + (I - W_1^* W_1)^2 + W_1^* W_2 W_2^* W_1 = I - W_1^* P_n W_1$$

obtained in the proof of Theorem 4, we get

$$D_{11} = P_n V_1^* L (I - W_1^* P_n W_1) L^* V_1 P_n,$$

$$D_{22} = (I - V_1 V_1^*) L (I - W_1^* P_n W_1) L^* (I - V_1 V_1^*),$$

$$D_{33} = V_2 V_1^* L (I - W_1^* P_n W_1) L^* V_1 V_2^*,$$

$$D_{12} = D_{21}^* = P_n V_1^* L (I - W_1^* P_n W_1) L^* (I - V_1 V_1^*),$$

$$D_{13} = D_{31}^* = -P_n V_1 L (I - W_1^* P_n W_1) L^* V_1 V_2^*,$$

$$D_{23} = D_{32}^* = -(I - V_1 V_1^*) L (I - W_1^* P_n W_1) L V_1 V_2^*,$$

or

$$T_n T_n^* = \begin{bmatrix} P_n V_1^* \\ I - V_1 V_1^* \\ -V_2 V_1^* \end{bmatrix} L (I - W_1^* P_n W_1) L^* [V_1 P_n \quad I - V_1 V_1^* \quad - V_1 V_2^*].$$

Lettting

$$S_n = \begin{bmatrix} P_n V_1^* \\ I - V_1 V_1^* \\ -V_2 V_1^* \end{bmatrix} L (I - W_1^* P_n W_1)^{\frac{1}{2}}$$

we note that $T_n T_n^* = S_n S_n^*$ so that $\|T_n\| = \|S_n\|$. Thus

$$\beta_n^2 = \sup_{\|f\|=1} \left\| \begin{bmatrix} P_n V_1^* \\ I - V_1 V_1^* \\ -V_2 V_1^* \end{bmatrix} L (I - W_1^* P_n W_1)^{\frac{1}{2}} f \right\|^2$$

$$= \sup_{\|f\|=1} \{ \|P_n V_1^* L (I - W_1^* P_n W_1)^{\frac{1}{2}} f\|^2 + \|(I - V_1 V_1^*) L (I - W_1 P_n W_1)^{\frac{1}{2}} f\|^2 + \|V_2 V_1^* L (I - W_1^* P_n W_1)^{\frac{1}{2}} f\|^2 \}$$

$$= \sup_{\|f\|=1} ((I - W_1^* P_n W_1)^{\frac{1}{2}} L [V_1 P_n V_1^* + (I - V_1 V_1^*)^2 + V_1 V_2^* V_2 V_1^*] L ((I - W_1^* P_n W_1)^{\frac{1}{2}} f, f).$$

A computation similar to that in the proof of Theorem 4 which gave the formula for $I - W_1^* P_n W_1$ gives

$$V_1 P_n V_1^* + (I - V_1 V_1^*)^2 + V_1 V_2^* V_2 V_1^* = I - V_1(I - P_n)V_1^*.$$

Hence

$$\beta_n^2 = \sup_{\|f\|=1} ([I - W_1^* P_n W_1]^{\frac{1}{2}} L^* [I - V_1(I - P_n)V_1^*] L [I - W_1^* P_n W_1]^{\frac{1}{2}} f, f) = \|[I - V_1(I - P_n)V_1^*]^{\frac{1}{2}} L [I - W_1^* P_n W_1]^{\frac{1}{2}}\|^2.$$

However, as in the computation in the Theorem 4,

$$(I + W_1^* P_n W_1)^{-1} = I = M^* P_n M \text{ and } [I - V_1(I - P_n)V_1^*]^{-1} = I + N(I - P_n)N^*,$$

so

$$\beta_{opt} = \sup_{\|f\|=1} \|(I + N(I - P_n)N^*)^{-\frac{1}{2}} L (I + M^* P_n M)^{-\frac{1}{2}}\|.$$

□

Corollary 2. *The formula of Theorem 5 holds for the finite, semi-infinite and double-infinite chains if* $\dim(P_n - P_{n-1}) < \infty$ *for all* n.

It is of interest to consider the non-causal case i.e. the chain $\{0, I\}$. If $\dim \mathcal{H} = \infty$ then this is not covered by Corollary 2. But, as we saw in Theorem 3, we can factor $[M \;\; I] = B[W_1 \;\; W_2]$ with $B = (I + MM^*)^{\frac{1}{2}}$, $W_1 = (I + MM^*)^{-\frac{1}{2}} M$, $W_2 = (I + MM^*)^{-\frac{1}{2}}$ and $\begin{bmatrix} N \\ I \end{bmatrix} = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} A$ with $A = (I + N^*N)^{\frac{1}{2}}$, $V_1 = N(I + N^*N)^{-\frac{1}{2}}$, $V_2 = (I + N^*N)^{-\frac{1}{2}}$. Applying Theorem 5 we obtain

$$\beta_{opt} = \max\{\|(I + NN^*)^{-\frac{1}{2}} L\|, \;\; \|L(I + M^*M)^{-\frac{1}{2}}\|\}.$$

The first term is the solution to

$$\inf\left\{\left\|\begin{bmatrix} L - NK \\ -K \end{bmatrix}\right\| \;:\; K \in \mathcal{L}(\mathcal{H})\right\},$$

the full information control problem of [HSK], Chapter 11, the dual problem to the estimation problem studied in the first part of this paper, for which the second term is the solution.

Finally, we give the formula for β_{opt} for the time invariant case. Again, the proof of this formula is conceptually the same as that of Theorem 6. We reduce to a 4-block problem using inner-outer factorizations and then compute.

Theorem 7. *For* $L,\ M,\ N \in H^\infty$,

$$\inf\left\{\left\|\begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} N \\ I \end{bmatrix} Q[M \;\; I]\right\| \;:\; Q \in H^\infty\right\} = \|(I + N_+ N_+^*)^{-\frac{1}{2}} L (I + M_-^* M_-)^{-\frac{1}{2}}\|$$

where $N_+ = PN \mid H^2$ *and* $M_- = (I - P)M \mid H_-^2$.

References

[A] W. Arveson, Interpolation problems in nest algebras, *J. Funct. Anal.* , **20** (1975), 208–233.

[F] A. Feintuch, *Robust Control Theory in Hilbert Space*, Springer-Verlag, N. Y., 1998.

[FF] A. Feintuch and B. A. Francis, Uniformly optimal control of linear time-variyng systems, *System Control Lett.*, **5** (1984), 67–71.

[FFGK] C. Foias, A. Frazho, I. Gohberg and M. Kaashoek, *Metric Constrained Interpolation, Commutant Lifting and Systems*, Operator Theory: Advances and Applications, vol. **100**, Birkhäuser, Basel, 1998.

[H] B.Hassibi, *Indefinite Metric Spaces in Estimation Control and Adaptive Filtering*, Ph. D. Thesis, Stanford University, Stanford CA, 1996.

[HK1] B. Hassibi and T. Kailath, On optimal solutions to H^∞ problems, *IEEE Trans. Automat. Control* (submitted).

[HK2] B. Hassibi and T. Kailath, On optimal solutions to two-block H^∞ problems, Proceedings of 1998 A.C.C., Philadelphia (1998), 1975–1979.

[HK3] B Hassibi and T. Kailath, Tracking with an H^∞ criterion, *IEEE Trans. Automat. Control* (submitted).

[HSK] B. Hassibi, A. Sayed, T. Kailath, *Indefinite Quadratic Estimation and Control, A Unified Approach to* H^2 *and* H^∞ *Theories*, SIAM Studies in Applied and Numerical Mathematics, vol. **16**, Philadelphia, 1999.