Optimal Disturbance Rejection and Robustness for Infinite Dimensional LTV Systems

Seddik M. Djouadi*

Abstract

In this paper, we consider the optimal disturbance rejection problem for possibly infinite dimensional linear time-varying (LTV) systems using a framework based on operator algebras of classes of bounded linear operators. This approach does not assume any state space representation and views LTV systems as causal operators. After reducing the problem to a shortest distance minimization in a space of bounded linear operators, duality theory is applied to show existence of optimal solutions, which satisfy a “time-varying” allpass or flatness condition. Under mild assumptions the optimal TV controller is shown to be essentially unique. Next, the concept of M-ideals of operators is used to show that the computation of time-varying (TV) controllers reduces to a search over compact TV Youla parameters. This involves the norm of a TV compact Hankel operator defined on the space of causal trace-class 2 operators and its maximal vectors. Moreover, an operator identity to compute the optimal TV Youla parameter is provided. These results are generalized to the mixed sensitivity problem for TV systems as well, where it is shown that the optimum is equal to the operator induced of a TV mixed Hankel-Toeplitz. The final outcome of the approach developed here is that it leads to two tractable finite dimensional convex optimizations producing estimates to the optimum within desired tolerances, and a method to compute optimal time-varying controllers.

* S.M. Djouadi is with the Electrical Engineering & Computer Science Department, University of Tennessee, Knoxville, TN 37996-2100. djouadi@eecs.utk.edu
Mathematical Preliminaries and Notation

- $\mathcal{B}(E, F)$ denotes the space of bounded linear operators from a Banach space $E$ to a Banach space $F$, endowed with the operator norm
  \[ \|A\| := \sup_{x \in E, \|x\| \leq 1} \|Ax\|, \quad A \in \mathcal{B}(E, F) \]

- $\ell^2$ denotes the usual Hilbert space of square summable sequences with the standard standard norm
  \[ \|x\|_2^2 := \sum_{j=0}^{\infty} |x_j|^2, \quad x := (x_0, x_1, x_2, \cdots) \in \ell^2 \]
  and inner product
  \[ \langle x, y \rangle = \sum_i \bar{y}_i x_i, \quad y := (y_0, y_1, y_2, \cdots) \in \ell^2 \]

- $P_k$ the usual truncation operator for some integer $k$, which sets all outputs after time $k$ to zero, i.e.,
  \[ P_k x = (x_0, x_1, x_2, \ldots, x_k, 0, 0, \cdots) \]

- An operator $A \in \mathcal{B}(\ell^2, \ell^2)$ is said to be causal if it satisfies the operator equation:
  \[ P_k AP_k = P_k A, \quad \forall k \text{ positive integers} \]
  and strictly causal if it satisfies
  \[ P_{k+1} AP_k = P_{k+1} A, \quad \forall k \text{ positive integers} \quad (1) \]

- A partial isometry is an operator $A$ on a Hilbert space which preserves the norm, i.e., $\|Au\| = \|u\|$ for all vectors $u$ in the orthogonal complement of the null space of $A$ \[6\].

- A bounded linear functional $f$ on a Banach space $E$ is a real or complex valued linear map such that there exists $c \geq 0$ with $| \langle f, x \rangle | \leq c \|x\|_E$, $\forall x \in E$, where $\langle f, x \rangle$ denotes the
image of $x$ under $f$, and $\| \cdot \|_E$ is the norm on $E$. The dual pace of $E$, denoted $E^*$, is the space of bounded linear functionals on $E$ under the norm $\| f \|_{E^*} := \sup_{x \in E, \| x \|_E} | < f, x > |$. Note that $E^*$ is itself a Banach space [6].

- A sequence $\{ x_n \}$ in $E$ is said to converge weakly to $x \in E$ if $| < f, x_n > - < f, x > | \to 0$, $\forall f \in E^*$. A sequence $\{ f_n \} \subset E^*$ is said to converge in the weak* topology to $f \in E^*$ if $| < f_n, x > - < f, x > | \to 0$, $\forall x \in E$.

- A sequence of operators $\{ T_n \} \subset \mathcal{B}(E, F)$ converges to an operator $T \in \mathcal{B}(E, F)$ in the operator topology if $\| T_n - T \| \to 0$. It is said to converge in the strong operator topology (SOT) if $\| T_n x - T x \|_F \to 0$, $\forall x \in E$. $\mathcal{K}$ denotes the space of compact operators in $\mathcal{B}(E, F)$. An operator $T \in \mathcal{B}(E, F)$ is compact if there exists a sequence of finite rank (i.e., with finite dimensional ranges) operators $\{ T_n \} \subset \mathcal{B}(E, F)$, such that, $\| T - T_n \| \to 0$ as $n \to \infty$ [6].

The subscript "c" denotes the restriction of a subspace of operators to its intersection with causal (see [33, 11] for the definition) operators. "*" stands for the adjoint of an operator or the dual space of a Banach space depending on the context. $\ominus$ denotes the orthogonal complement. The symbol "\bigwedge" denotes the closed linear span, and "\bigcap" the intersection. "\rightharpoonup*" denotes convergence in the weak* topology [6, 9]. Right and left hand sides are abbreviated as RHS and LHS, respectively.

## 1 Introduction

There have been numerous attempts in the literature to generalize ideas about robust control theory [44, 46] to time-varying (TV) systems (for e.g. [12, 13, 11, 37, 10, 36, 39, 40, 34] and references therein). In [12, 13] and more recently [11] the authors studied the optimal weighted sensitivity minimization problem, the two-block problem, and the model-matching problem for LTV systems using inner-outer factorization for positive operators. They obtained abstract solutions involving the computation of norms of certain operators, which is difficult since these are infinite dimensional problems. Moreover, no indication on how to compute optimal linear time-varying (LTV) controllers is provided. The operator theoretic methods in [12, 13, 11] are difficult to implement and do not provide algorithms which even compute approximately TV controllers. In [4] a state space extension of the Nehari Theorem to a time-varying system-theoretic setting is developed with parametrization of suboptimal solutions. In [40] the authors rely on state space techniques which lead to algorithms based on infinite dimensional operator inequalities. These methods lead
to suboptimal controllers but are difficult to solve and are restricted to finite dimensional systems. Moreover, they do not allow the degree of suboptimality to be estimated. An extension of these results to uncertain systems is reported in [41] relying on uniform stability concepts.

In [8] both the sensitivity minimization problem in the presence of plant uncertainty, and robust stability for LTV systems in the $\ell^\infty$ induced norm is considered. However, their methods could not be extended to the case of systems operating on finite energy signals.

We believe that it is important to point to the fact that the lack of calculation of the degree of suboptimality in these methods, can result in arbitrary poor overall closed-loop performance. This fact is explored in [18], where the authors, although in the linear time-invariant (LTI) case, studied a "two-arc" counter example that shows that in the limit suboptimal controllers can lead to arbitrary poor performance in the presence of plant uncertainty. It seems that this counter example has not yet received enough attention.

Analysis of time-varying control strategies for optimal disturbance rejection for known time-invariant plants has been studied in [16, 2]. A robust version of these problems was considered in [15, 7] in different induced norm topologies. All these references showed that for time-invariant nominal plants and weighting functions, time-varying control laws offer no advantage over time-invariant ones.

In the book [35], the authors consider various interpolation problems in the stationary (LTI) and nonstationary (LTV) cases. In particular, Nevanlinna-Pick, Hermite-Fejer, Sarason, and Nudelman interpolations and Nehari extension problems are studied. In the TV case the unit disk is replaced by the set of diagonal matrices such as the weighted shift with spectral radius less than 1. The interpolation condition is replaced by an identity involving diagonal matrices. Two methods are used to solve these problems, a method based on the reduction to the time-invariant interpolation case but with operator valued functions, and another method based on a TV commutant lifting theorem developed by the authors of the book. The solutions for these problems are provided in terms of infinite dimensional operator identities.

In this paper, we are interested in optimal disturbance rejection for (possibly infinite-dimensional, i.e., systems with an infinite number of states) LTV systems. These systems have been used as models in computational linear algebra and in a variety of computational and communication networks as described in the excellent book [17]. This allows variable number of states which is predominant in networks which can switch on or off certain parts of the system [17], and infinite number of
states as in distributed parameter systems. We study questions such as existence and uniqueness of optimal LTV controllers under specific conditions. What is the corresponding notion for the Hankel operator which solves the optimal $H^\infty$ problem? What does allpass mean for LTV systems? Is the optimal Youla parameter compact and if so under what conditions? Is it possible to compute optimal LTV controllers within desired accuracy? Note that compactness is important since it allows approximations by finite rank operators (matrices), and therefore makes the problem amenable to numerical computation.

Using inner-outer factorizations as defined in [3, 11] with respect of the nest algebra of lower triangular (causal) bounded linear operators defined on $\ell^2$ we show that the problem reduces to a distance minimization between a special operator and the nest algebra. The inner-outer factorization used here holds under weaker assumptions than [12, 13], and in fact, as pointed in ([3] p. 180), is different from the factorization for positive operators used there. Duality structure and predual formulation of the problem showing existence of optimal LTV controllers is provided. The optimum is shown to satisfy a “TV” allpass condition quantified in the form of a partial isometry of an operator, therefore, generalizing the flatness or allpass condition [48, 24, 25]. The optimal controller and the corresponding dual operator are shown to be essentially unique under mild assumptions.

With the use of M-ideals of operators, it is shown that the computation of time-varying (TV) controllers reduces to a search over compact TV Youla parameters. Furthermore, the optimum is shown to be equal to the norm of a compact time-varying Hankel operator defined on the space of causal Hilbert-Schmidt operators. The latter is a “natural” analogous to the Hankel operator used in the LTI case. An operator equation to compute the optimal TV Youla parameter is also derived. The results obtained here lead to a pair of dual finite dimensional convex optimizations which approach the real optimal disturbance rejection performance from both directions not only producing estimates within desired tolerances, but allowing the computation of optimal time-varying controllers. The numerical computation involve solving a semi-definite programming problem, and a search over lower triangular matrices.

The results are generalized to the mixed sensitivity problem for TV systems as well, where it is shown that the optimum is equal to the operator induced of a TV mixed Hankel-Toeplitz operator generalizing analogous results known to hold in the LTI case [45, 38, 24].

Our approach is purely input-output and does not use any state space realization, therefore the results derived here apply to infinite dimensional LTV systems, i.e., TV systems with an infinite
number of state variables. Although the theory is developed for causal stable system, it can be extended in a straightforward fashion to the unstable case using coprime factorization techniques for LTV systems discussed in [13, 11]. The framework developed can also be applied to other performance indexes, such as the optimal TV robust disturbance attenuation problem considered in [18, 19, 20]. For continuous LTV systems subject to time-varying unstructured uncertainty the problem was considered in [5], where it is shown that for causal LTV systems, it is equivalent to finding the smallest fixed point of a two-disc type optimization problem under TV feedback control laws. For the discrete version of the same problem, the duality structure of the problem and a solution based on a new bilinear map was provided in terms of an infinite dimensional identity in [23].

A related problem involving the computation of the gap metric for LTV systems has been considered in [22]. In particular, the computation of the gap metric for TV systems using a time-varying generalization of normalized coprime factorization as an iterative scheme involving the norm of an operator with a TV Hankel plus Toeplitz structure is provided. The duality structure is characterized by computing the preannihilator and exploiting the particular structure of the problem.

Part of the results presented here were announced in [31, 21] without proofs.

The rest of the paper is organized as follows. Section 2 contains the problem formulation. The duality structure of the problem is worked out in section 3. In section 4, the optimal solution is shown to satisfy a TV allpass property. Section 5 shows that under mild assumptions the TV optimal $Q$ parameter is compact. Section 6 discusses the uniqueness of the optimal solution. In section 7, a solution based on a TV Hankel operator is derived together with an extremal identity for the optimum. Section 8 presents a numerical solution based on duality theory. A Generalization to the TV mixed sensitivity problem is carried out in section 9. Section 10 contains concluding remarks.

2 Problem formulation

In this paper we consider the problem of optimizing performance for causal linear time varying systems. The standard block diagram for the optimal disturbance attenuation problem that is considered here is represented in Fig. 1, where $u$ represents the control inputs, $y$ the measured outputs, $z$ is the controlled output, $w$ the exogenous perturbations. $P$ denotes a causal stable linear time varying plant, and $K$ denotes a time varying controller. The closed-loop transmission from $w$ to $z$ is denoted by $T_{zw}$. Using the standard Youla parametrization of all stabilizing controllers the
closed loop operator $T_{zw}$ can be written as \cite{11},

$$T_{zw} = T_1 - T_2 QT_3$$  \hspace{1cm} (2)

where $T_1$, $T_2$ and $T_3$ are stable causal time-varying operators, that is, $T_1$, $T_2$ and $T_3 \in B_c(\ell^2, \ell^2)$. In this paper we assume without loss of generality that $P$ is stable, the Youla parameter $Q := K(I + PK)^{-1}$ is then an operator belonging to $B_c(\ell^2, \ell^2)$, and is related univocally to the controller $K$ \cite{33}. Note that $Q$ is allowed to be *time-varying*. If $P$ is unstable it suffices to use the coprime factorization techniques in \cite{39, 11} which lead to similar results. The magnitude of the signals $w$ and $z$ is measured in the $\ell^2$-norm. Two problems are considered here optimal disturbance rejection which, and the mixed sensitivity problem for LTV systems which includes a robustness problem in the gap metric studied in \cite{11, 36}. Note that for the latter problem $P$ is assumed to be unstable and we have to use coprime factorizations. The performance index can be written in the following form

$$\mu := \inf \{ \| T_{zw} \| : K \text{ being robustly stabilizing linear time } - \text{ varying controller} \}$$

$$= \inf_{Q \in B_c(\ell^2, \ell^2)} \| T_1 - T_2 QT_3 \|$$  \hspace{1cm} (3)

The performance index (3) will be transformed into a distance minimization between a certain operator and a subspace to be specified shortly. To this end, define a nest $\mathcal{N}$ as a family of closed subspaces of the Hilbert space $\ell^2$ containing $\{0\}$ and $\ell^2$ which is closed under intersection and closed span. Let $Q_n := I - P_n$, for $n = -1, 0, 1, \cdots$, where $P_{-1} := 0$ and $P_{\infty} := I$. Then $\tilde{Q}_n$ is a projection, and we can associate to it the following nest $\mathcal{N} := \{ \tilde{Q}_n \ell^2, n = -1, 0, 1, \cdots \}$. The
triangular or nest algebra $\mathcal{T}(\mathcal{N})$ is the set of all operators $T$ such that $TN \subseteq N$ for every element $N$ in $\mathcal{N}$. That is

$$\mathcal{T}(\mathcal{N}) = \{ A \in \mathcal{B}(\ell^2, \ell^2) : P_n A(I - P_n) = 0, \forall n \}$$

(4)

Note that the Banach space $\mathcal{B}_c(\ell^2, \ell^2)$ is identical to the nest algebra $\mathcal{T}(\mathcal{N})$. Define

$$N^- = \bigvee \{ N' \in \mathcal{N} : N' < N \}, \quad N^+ = \bigwedge \{ N' \in \mathcal{N} : N' > N \}$$

(5)

where $N' < N$ means $N' \subset N$, and $N' > N$ means $N' \supset N$. The subspaces $N \cap N^-$ are called the atoms of $\mathcal{N}$. Since in our case the atoms of $\mathcal{N}$ span $\ell^2$, then $\mathcal{N}$ is said to be atomic [3].

Following [3] we introduce inner-outer factorizations for the operators in $\mathcal{T}(\mathcal{N})$ as follows:

An operator $A$ in $\mathcal{T}(\mathcal{N})$ is called outer if the range projection, denoted $P(R_A)$, $R_A$ being the range of $A$ and $P$ the orthogonal projection onto $R_A$, commutes with $\mathcal{N}$ and $AN$ is dense in the intersection $N \cap R_A$ for every $N \in \mathcal{N}$. A partial isometry $U$ is called inner in $\mathcal{T}(\mathcal{N})$ if $U^*U$ commutes with $\mathcal{N}$ [1] [3] [11]. In our case, $A \in \mathcal{T}(\mathcal{N}) = \mathcal{B}_c(\ell^2, \ell^2)$ is outer if $P$ commutes with each $\tilde{Q}_n$ and $A\tilde{Q}_n\ell^2$ is dense in $\tilde{Q}_n\ell^2 \cap A\ell^2$. $U \in \mathcal{B}_c(\ell^2, \ell^2)$ is inner if $U$ is a partial isometry and $U^*U$ commutes with every $\tilde{Q}_n$. Applying these notions to the time-varying operator $T_2 \in \mathcal{B}_c(\ell^2, \ell^2)$, we get $T_2 = T_{2i}T_{2o}$, where $T_{2i}$ and $T_{2o}$ are inner outer operators in $\mathcal{B}_c(\ell^2, \ell^2)$, respectively. Similarly, outer-inner factorization can be defined and the operator $T_3 = T_{3o}T_{3i}$, where $T_{3i} \in \mathcal{B}_c(\ell^2, \ell^2)$ is inner $T_{3o} \in \mathcal{B}_c(\ell^2, \ell^2)$ is outer.

The performance index $\mu$ in [3] can then be written as

$$\mu = \inf_{Q \in \mathcal{B}_c(\ell^2, \ell^2)} \| T_1 - T_{2i}T_{2o}QT_{3o}T_{3i} \|$$

(6)

Following classical robust control theory [24] [25] [46] assume: (A1) that $T_{2o}$ and $T_{3o}$ are invertible both in $\mathcal{B}_c(\ell^2, \ell^2)$.

Assumption (A1) can be somewhat relaxed by assuming instead that the outer operators $T_{2o}$ and $T_{3o}$ are bounded below (see Lemma [1] p. 220).

Assumption (A1) guarantees that the map $Q \rightarrow T_{2o}B_c(\ell^2, \ell^2)T_{3co}$ is bijective. In the time-invariant case this assumption means essentially that the outer factor of the plant $P$ is invertible [25]. Under this assumption $T_{2i}$ becomes an isometry and $T_{3i}$ a co-isometry in which case $T_{2i}^*T_{2i} = I$ and $T_{3i}T_{3i}^* = I$. By "absorbing" the operators $T_{2o}$ and $T_{3co}$ into the "free" operator $Q$, expression (6) is
then equivalent to
\[
\mu = \inf_{Q \in \mathcal{B}_c(\ell^2, \ell^2)} \| T_{2i}^* T_1 T_{3ci}^* - Q \| \quad (7)
\]
Expression (7) is the distance from the operator \( T_{2i}^* T_1 T_{3ci}^* \in \mathcal{B}(\ell^2, \ell^2) \) to the nest algebra \( \mathcal{B}_c(\ell^2, \ell^2) \).

In the next section we study the distance minimization problem (7) in the context of M-ideals and the operator algebra setting discussed above.

### 3 Duality

Let \( X \) be a Banach space and \( X^* \) its dual space, i.e., the space of bounded linear functionals defined on \( X \). For a subset \( J \) of \( X \), the annihilator of \( J \) in \( X^* \) is denoted \( J^\perp \) and is defined by \[ J^\perp := \{ \Phi \in X^* : \Phi(f) = 0, \ f \in J \}, \]
i.e., \( J^\perp \) is the set of bounded linear functionals on \( X \) which vanish on \( J \). Similarly, if \( K \) is a subset of \( X^* \) then the preannihilator of \( K \) in \( X \) is denoted \( \perp K \), and is defined by \[ \perp K := \{ x \in X : \Phi(x) = 0, \ \Phi \in K \}. \]
The existence of a preannihilator implies that the following identity holds \[ \min_{y \in K} \| x - y \| = \sup_{k \in \perp K, \|k\| \leq 1} | < x, k > | \quad (8) \]
where \( < \cdot, \cdot > \) denotes the duality product. Let us apply these results to the problem given in (7) by putting
\[
X = \mathcal{B}(\ell^2, \ell^2), \ x = T_{2i}^* T_1 T_{3ci}^* \in \mathcal{B}(\ell^2, \ell^2), \ J = \mathcal{B}_c(\ell^2, \ell^2) \quad (9)
\]
Introduce the class of compact operators on \( \ell^2 \) called the trace-class or Schatten 1-class, denoted \( \mathcal{C}_1 \), such that if \( \{ \xi_i \} \) is a basis of \( \ell^2 \), and \( T \in \mathcal{C}_1 \), then \( \sum_i < T \xi_i, \xi_i > < \infty \). Moreover, the sum \( \sum_i < T \xi_i, \xi_i > \) is independent of the choice of the basis \[ [14, 3] \]. The operator \( T^* T \), \( T^* \) being the adjoint operator of \( T \), is positive (i.e., \( < T x, x > > 0, \ x \in \ell^2 \) and \( x \neq 0 \)) and compact and has a unique positive square root, denoted \( (T^* T)^{\frac{1}{2}} \). The trace-class norm is defined as \[ [14, 3] \],
\[
\| T \|_1 := tr(T^* T)^{\frac{1}{2}} = \sum_i < (T^* T)^{\frac{1}{2}} \xi_i, \xi_i > \quad (10)
\]
where \( tr \) denotes the Trace.

We identify \( \mathcal{B}(\ell^2, \ell^2) \) with the dual space of \( \mathcal{C}_1 \), \( \mathcal{C}_1^* \), under trace duality [14], that is, every operator \( A \) in \( \mathcal{B}(\ell^2, \ell^2) \) induces a bounded linear functional on \( \mathcal{C}_1 \) as follows: \( \Phi_A \in \mathcal{C}_1^* \) is defined by \( \Phi_A(T) = tr(AT) \), and we write \( \mathcal{B}(\ell^2, \ell^2) \cong \mathcal{C}_1^* \) to express that \( \mathcal{B}(\ell^2, \ell^2) \) is isometrically isomorphic to \( \mathcal{C}_1^* \).

Every trace-class operator \( T \) in turn induces a bounded linear functional on \( \mathcal{B}(\ell^2, \ell^2) \), namely, \( \Phi_T(A) = tr(AT) \) for all \( A \) in \( \mathcal{B}(\ell^2, \ell^2) \).

To compute the preannihilator of \( \mathcal{B}_c(\ell^2, \ell^2) \) define the subspace \( S \) of \( \mathcal{C}_1 \) by

\[
S := \{ T \in \mathcal{C}_1 : (I - \tilde{Q}_n)T\tilde{Q}_{n+1} = 0, \ \text{for all} \ n \} \tag{11}
\]

In the following Lemma we show that \( S \) is the preannihilator of \( \mathcal{B}_c(\ell^2, \ell^2) \).

**Lemma 1** The preannihilator of \( \mathcal{B}_c(\ell^2, \ell^2) \) in \( \mathcal{C}_1 \), \( \mathcal{B}_c(\ell^2, \ell^2) \), is isometrically isomorphic to \( S \).

**Proof.** By Lemma 16.2 in [3] the preannihilator of \( \mathcal{T}(\mathcal{N}) \) is given by \( \Phi_T \in \mathcal{B}_c(\ell^2, \ell^2) \) if and only \( T \) belongs to the subspace

\[
\{ T \in \mathcal{C}_1 : P(N^-)^\perp TP(N) = 0 \ \text{for all} \ N \in \mathcal{N} \}
\]

where \( P(N) \) denotes the orthogonal projection on \( N \), likewise for \( P(N^-) \). \( P(N^-)^\perp \) the complementary projection of \( P(N^-) \), that is, \( P(N^-)^\perp = I - P(N^-) \).

In our case \( \mathcal{N} \) is atomic, and for any \( N \in \mathcal{N} \) there exists \( n \) such that \( N = \tilde{Q}_{n+1}\ell^2 \), i.e., \( P(N) = \tilde{Q}_{n+1} \).

The immediate predecessor of \( N, N^- \), is then given by \( \tilde{Q}_n\ell^2 \), i.e., \( N^- = \tilde{Q}_n\ell^2 \). The orthogonal complement of \( N^- \) is then \( \ell^2 \oplus \tilde{Q}_n\ell^2 \). So \( P(N^-)^\perp = I - \tilde{Q}_n \). Therefore,

\[
S = \{ T \in \mathcal{C}_1 : (I - \tilde{Q}_n)T\tilde{Q}_{n+1} = 0, \ \forall \ n \}
\]

is isometrically isomorphic to the preannihilator of \( \mathcal{B}_c(\ell^2, \ell^2) \), and the Lemma is proved. ■

The existence of a predual \( \mathcal{C}_1 \) and a preannihilator \( S \) implies the following Theorem which is a consequence of Theorem 2 in [9] (Chapter 5.8).

**Theorem 1** Under assumption (A1) there exists at least one optimal \( Q_o \) in \( \mathcal{B}_c(\ell^2, \ell^2) \) achieving
optimal performance $\mu$ in \cite{1}, moreover the following identities hold

$$
\mu = \inf_{Q \in \mathcal{B}_c(\ell^2, \ell^2)} \| T_{2i}^* T_{1i}^* - Q \| = \| T_{2i}^* T_{1i}^* - Q_o \| = \sup_{T \in S, \| T \| \leq 1} | tr(T T_{2i}^* T_{1}^* T_{3i}^*) | \tag{12}
$$

Theorem 1 not only shows the existence of an optimal LTV controller, but plays an important role in its computation by reducing the problem to primal and dual of finite dimensional convex optimizations. Under a certain condition in the next section it is shown that the supremum in (12) is achieved. Theorem 1 also leads to a solution based on a Hankel type operator, which parallels the Hankel operator known in the $H^\infty$ control theory.

4 TV Allpass Property of the Optimum

The dual space of $\mathcal{B}(\ell^2, \ell^2)$, $\mathcal{B}(\ell^2, \ell^2)^*$, is given by the space, $\mathcal{B}(\ell^2, \ell^2)^* \simeq C_1 \oplus_1 \mathcal{K}^\perp$, where $\mathcal{K}^\perp$ is the annihilator of $\mathcal{K}$, the space of compact operators on $\ell^2$, and the symbol $\oplus_1$ means that if $\Phi \in C_1 \oplus_1 \mathcal{K}^\perp$ then $\Phi$ has a unique decomposition as follows \cite{14}

$$
\Phi = \Phi_o + \Phi_T, \quad \| \Phi \| = \| \Phi_o \| + \| \Phi_T \| \tag{13}
$$

where $\Phi_o \in \mathcal{K}^\perp$, and $\Phi_T$ is induced by the operator $T \in C_1$, i.e., $\Phi_T(A) = tr(AT)$, $A \in \mathcal{B}(\ell^2, \ell^2)$. Banach space duality states that \cite{9}

$$
\inf_{y \in J} \| x - y \| = \max_{\Phi \in J^\perp, \| \Phi \| \leq 1} | \Phi(x) | \tag{14}
$$

In our case $J = \mathcal{B}_c(\ell^2, \ell^2)$. Since $\mathcal{B}(\ell^2, \ell^2)^*$ contains $C_1$ as a subspace, then $\mathcal{B}_c(\ell^2, \ell^2)^\perp$ contains the preannihilator $S$, i.e., the following expression for the annihilator of $J$, $J^\perp$, is deduced

$$
J^\perp := \mathcal{B}_c(\ell^2, \ell^2)^\perp = S \oplus_1 \left( \mathcal{K} \cap \mathcal{B}_c(\ell^2, \ell^2) \right)^\perp \tag{15}
$$

A result in \cite{3} asserts that if a linear functional $\Phi$ belongs to the annihilator $J^\perp$ and $\Phi$ decomposes as $\Phi = \Phi_o + \Phi_T$, where $\Phi_o \in \left( \mathcal{K} \cap \mathcal{B}_c(\ell^2, \ell^2) \right)^\perp$ and $\Phi_T \in S$, then both $\Phi_o \in \mathcal{B}_c(\ell^2, \ell^2)^\perp$ and $\Phi_T \in$
\[ \min_{Q \in \mathcal{B}_c(\ell^2, \ell^2)} \| T_{2i}^* T_1 T_{3ci}^* - Q \| = \max_{\Phi_o \in (\mathcal{K} \cap \mathcal{B}_c)^\perp, T \in S} \| \Phi_o(T_{2i}^* T_1 T_{3ci}^*) + tr(T T_{2i}^* T_1 T_{3ci}^*) \| \] (16)

If \( \Phi_{opt} = \Phi_{opt,a} + \Phi_{T_{opt}} \) achieves the maximum in the RHS of (16), and \( Q_o \) the minimum on the LHS, then the alignment condition in the dual is immediately deduced from (16)

\[ |\Phi_{opt,a}(T_{2i}^* T_1 T_{3ci}^*) + tr(T_{opt} T_{2i}^* T_1 T_{3ci}^*)| = |T_{2i}^* T_1 T_{3ci}^* - Q_o| (\| \Phi_{opt,a} \| + \| T_{opt} \|_1) \] (17)

Assume further that (A2): \( T_{2i}^* T_1 T_{3ci}^* \) is a compact operator. This is the case, for example, if \( T_1 \) is compact, then \( \Phi_{opt,a}(T_{2i}^* T_1 T_{3ci}^*) = 0 \) and the maximum in (16) is achieved on \( S \), that is the supremum in (16) becomes a maximum. Compact operators include, for example, systems with impulse responses that have measurable and square integrable kernels.

It is instructive to note that in the LTI case assumption (A2) is the analogue of the assumption that \( T_{2i}^* T_1 T_{3ci}^* \) is the sum of two parts, one part continuous on the unit circle and the other in \( H^\infty \), in which case the optimum is allpass [24, 18]. We would like to find the allpass equivalent for the optimum in the linear time varying case. This may be formulated by noting that flatness or allpass condition in the LTI case means that the modulus of the optimum \( |(T_{2i}^* T_1 T_{3ci}^* - Q_o)(e^{i\theta})| \) is constant at almost all frequencies (equal to \( \mu \)). In terms of operator theory, note that the optimum viewed as a multiplication operator acting on \( L^2 \) or \( H^2 \), changes the norm of any function in \( L^2 \) or \( H^2 \) by multiplying it by a constant (\( = \mu \)). That is, the operator achieves its norm at every \( f \in L^2 \) of unit \( L^2 \)-norm. This interpretation is generalized to the LTV case in the following Theorem, which part of it first appeared in [31] without a proof. In fact, the optimal cost (3) is shown to be a partial isometry.

**Theorem 2** Under assumptions (A1) and (A2) there exists at least one optimal linear time varying \( Q_o \in \mathcal{B}_c(\ell^2, \ell^2) \) that satisfies the following

i) the duality expression

\[ \mu = \| T_{2i}^* T_1 T_{3ci}^* - Q_o \| tr(T_o T_{2i}^* T_1 T_{3ci}^*) \| = \max_n \{(I - \tilde{Q}_n) T_{2i}^* T_1 T_{3ci}^* \tilde{Q}_n \} \] (18)
holds, where \( T_o \) is some operator in \( S \) and \( \| T_o \|_1 = 1 \).

ii) and if \( \mu > \mu_{oo} := \inf_{Q \in K} \| T_2^* T_1 T_{3ci}^* - Q \| \), i.e. when the causality constraint is removed, the following operator

\[
\frac{T_2^* T_1 T_{3ci}^* - Q_o}{\| T_2^* T_1 T_{3ci}^* - Q_o \|}
\]

is a partial isometry. That is, the optimum is an isometry on the range space of the operator \( T_o \) in i). This is the time-varying counterpart of flatness. Note the condition \( \mu > \mu_{oo} \) is sharp, in the sense that if it does not hold there exist \( T_2^* T_1 T_{3ci}^* \) and \( Q_o \) such that \( \frac{T_2^* T_1 T_{3ci}^* - Q_o}{\| T_2^* T_1 T_{3ci}^* - Q_o \|} \) is not a partial isometry.

**Proof.**

i) Identity (18) is implied by the previous argument that the supremum in (12) is achieved by some \( T_o \) in \( S \) with trace-class norm equal to 1. Combining this result with Corollary 16.8 in [3] (see also [1]), which asserts that

\[
\| T_2^* T_1 T_{3ci}^* - Q_o \| = \sup_n \| (I - \tilde{Q}_n) T_2^* T_1 T_{3ci}^* \tilde{Q}_n \|
\]

shows in fact that the supremum w.r.t. \( n \) is achieved proving that (18) holds.

ii) The operator \( T_o^* T_o \) is self-adjoint and compact, and therefore admits a spectral representation \( T_o^* T_o = \sum_j \lambda_j \phi_j \otimes \phi_j \), where \( \{ \phi_j \} \) form an orthonormal basis for \( \ell^2 \) consisting of eigenvectors of \( T_o^* T_o \), and \( \lambda_j \) are its necessarily positive real eigenvalues [14]. It follows that

\[
\sum_j < T_o^* T_o \phi_j, \phi_j > = tr(T_o^* T_o)^{\frac{1}{2}} = \sum_j \lambda_j = \| T_o \|_1
\]

where we used the fact that \( T_o^* T_o \phi_j = \lambda_j \phi_j \neq 0 \). Let the polar decomposition of \( T_o \) be \( T_o = U(T_o^* T_o)^{\frac{1}{2}} \), where \( U \) is an isometry on the set \( \{ \phi_j \} \), and \( (T_o^* T_o)^{\frac{1}{2}} \) the "square root" of \( T_o^* T_o \). Now \( \sqrt{\lambda_j} \) being the non-zero singular values of \( T_o \), are also the eigenvectors of \( (T_o^* T_o)^{\frac{1}{2}} \). It follows that for \( \lambda_j \neq 0 \), we have

\[
\frac{1}{\sqrt{\lambda_j}} T_o \phi_j = \frac{1}{\sqrt{\lambda_j}} U(T_o^* T_o)^{\frac{1}{2}} \phi_j = \frac{1}{\sqrt{\lambda_j}} U \sqrt{\lambda_j} \phi_j = U \phi_j
\]
so \( \{ U\phi_j \} \) is an orthonormal set which spans the range of \( T_o \). Call \( \psi_j := U\phi_j \), then \( T_o \) can be written as 
\[
T_o = \sum_j \lambda_j \phi_j \otimes \psi_j
\]
and
\[
\|T_o\|_1 = \sum_j \lambda_j = 1 \quad (23)
\]

\[
\mu = \left| \text{tr} \left( (T_{2i}^* T_1 T_{3ci}^* - Q_o) T_o \right) \right| = \left| \sum_j \lambda_j < \phi_j, (T_{2i}^* T_1 T_{3ci}^* - Q_o) \psi_j > \right| \quad \text{by definition of the trace}
\]
\[
\leq \sum_j \lambda_j \|\phi_j\|_2 \| (T_{2i}^* T_1 T_{3ci}^* - Q_o) \psi_j \|_2 \leq \sum_j \lambda_j \| T_{2i}^* T_1 T_{3ci}^* - Q_o \| \| \psi_j \|_2 \quad (24)
\]
\[
\leq \sum_j \lambda_j \| T_{2i}^* T_1 T_{3ci}^* - Q_o \| \leq \| T_o \|_1 \| T_{2i}^* T_1 T_{3ci}^* - Q_o \| = \mu \quad (25)
\]

where the first inequality in (24) follows from the Cauchy-Schwartz inequality, the second from a standard property of the induced norm, and the fact that \( \{ \phi_j \} \) and \( \{ \psi_j \} \) are orthonormal sets. Hence equality must hold throughout yielding
\[
\| (T_{2i}^* T_1 T_{3ci}^* - Q_o) \psi_j \|_2 = \| T_{2i}^* T_1 T_{3ci}^* - Q_o \| \| \psi_j \|_2 \quad (26)
\]
for each \( \psi_j \), that is, \( T_{2i}^* T_1 T_{3ci}^* - Q_o \) attains its norm on each \( \psi_j \), it must then attain its norm everywhere on the span of \( \{ \psi_j \} \), so that \( \frac{T_{2i}^* T_1 T_{3ci}^* - Q_o}{\| T_{2i}^* T_1 T_{3ci}^* - Q_o \|} \) is an isometry on the range of \( T_o \).

Note that if condition \( \mu > \mu_o \) does not hold it is straightforward to find a counter example such that the optimal is not a partital isometry. Choose for example \( T_{2i}^* T_1 T_{3ci}^* \) to be a strict upper triangular which is not a partial isometry, then an optimal \( Q_o \) is the zero matrix. The optimum is equal to \( \| T_{2i}^* T_1 T_{3ci}^* \| \) which does not correspond to a partial isometry. ■

Identity (19) represents the allpass condition in the time-varying case, since it corresponds to the allpass or flatness condition in the time-invariant case for the standard optimal \( H^\infty \) problem. In the next section, we show that the search over \( Q \) can be restricted to compact operators. This is achieved by introducing a Banach space notion known as M-ideals [3].

### 5 M-Ideals and Compact Youla TV Parameters

In this section we show that the optimal TV \( Q \) is compact. This will allow approximations by finite rank operators to any desired accuracy, and consequently bring in finite dimensional convex
optimizations to carry out the computations. In order to achieve our objective we rely on the concept of \( M \)-ideals which is discussed next.

Following [3] we say that a closed subspace \( M \) of a Banach space \( B \) is an \( M \)-ideal if there exists a linear projection \( \Pi : B^* \to M^\perp \) from the dual space of \( B \), \( B^* \) onto the annihilator of \( M \), \( M^\perp \) in \( B^* \), such that for all \( b^* \in B^* \), we have

\[
\|b^*\| = \|\Pi b^*\| + \|b^* - \Pi b^*\| \quad (27)
\]

In this case \( M^\perp \) is called an \( L \)-summand of \( B^* \). The range \( N \) of \((I - \Pi)\) is a complementary subspace of \( M^\perp \), and \( B^* = M^\perp \oplus 1 \cdot N \). A basic property of \( M \)-ideals is that they are proximinal, that is, for every \( b \in B \), there is an \( m_o \) in \( M \) such that \( \inf_{m \in M} \|b - m\| = \|b - m_o\| \). Under assumption (A2) we generalize Lemma 1.6. in [32] to causal LTV systems, i.e., the space \( B_c(\ell^2, \ell^2) \). Recall that Lemma 1.6 states that if \( f \) is a function continuous on the unit circle, i.e., \( f \in C \), then

\[
\inf_{g \in H^\infty} \|f - g\|_{\infty} = \inf_{g \in A} \|f - g\|_{\infty} \quad (28)
\]

where \( A \) is the disk algebra, i.e., the space of analytic and continuous function on the unit disk \( A = H^\infty \cap C \). That is it suffices to restrict the search to functions in \( A \). To generalize (28) to causal LTV systems put \( B := K \), \( M := B_c(\ell^2, \ell^2) \), and show that for \( b \in K \) we have

\[
\inf_{m \in B_c(\ell^2, \ell^2)} \|b - m\| = \inf_{m \in B_c(\ell^2, \ell^2) \cap K} \|b - m\| \quad (29)
\]

By Theorem 3.11 in [3], \( B_c(\ell^2, \ell^2) \cap K \) is weak* dense in \( B_c(\ell^2, \ell^2) \). By Theorem 11.6 and Corollary 11.7 in [3], \( B_c(\ell^2, \ell^2) \cap K \) is an \( M \)-ideal in \( B_c(\ell^2, \ell^2) \), and the quotient map \( q_1 : B_c(\ell^2, \ell^2) / (B_c(\ell^2, \ell^2) \cap K) \to B_c(\ell^2, \ell^2) \) is isometric. Likewise, the quotient map \( q_2 : K / (B_c(\ell^2, \ell^2) \cap K) \to B_c(\ell^2, \ell^2) \) is isometric. And the identity (29) holds. In our case under assumption (A2) \( b = T_{2i}^*T_1T_{3ci}^* \in K \), and \( m = Q \) yields

\[
\inf_{Q \in B_c(\ell^2, \ell^2)} \|T_{2i}^*T_1T_{3ci}^* - Q\| = \inf_{Q \in B_c(\ell^2, \ell^2) \cap K} \|T_{2i}^*T_1T_{3ci}^* - Q\| = \|T_{2i}^*T_1T_{3ci}^* - Q_o\|
\]

for some optimal \( Q_o \in B_c(\ell^2, \ell^2) \cap K \). That is under (A2) the optimal \( Q \) is compact, and thus it suffices to restrict the search in [30] to causal and compact parameters \( Q \).

Now we turn our attention to studying questions regarding the uniqueness of the optimal solution.
6 On the Uniqueness of the Optimal TV Controller

In this section under assumptions (A1), (A2), and \( \mu > \mu_{oo} \), we prove that the optimal TV controller is essentially unique. Since there is an one-to-one onto correspondence between optimal controllers and Youla parameters, it suffices to show that the latter is essentially unique. Suppose by way of contradiction that there are two optimal compact Youla parameters \( Q_1 \) and \( Q_2 \) such that

\[
\mu = \| T_{2i}^*T_1^*T_{3ci}^* - Q_1 \| = \| T_{2i}^*T_1^*T_{3ci}^* - Q_2 \| \tag{30}
\]

and \( T_{2i}^*T_1^*T_{3ci}^* - Q_1 \) and \( T_{2i}^*T_1^*T_{3ci}^* - Q_1 \) are both isometries on the range of \( T_o \), where \( T_o \) is the optimal dual operator in Theorem 2. In this case, any convex combination of \( Q_1 \) and \( Q_2 \) is also optimal since

\[
\| \lambda(T_{2i}^*T_1^*T_{3ci}^*) + (1 - \lambda)(T_{2i}^*T_1^*T_{3ci}^*) + \lambda Q_1 + (1 - \lambda)Q_2 \| \geq \mu, \quad \forall \lambda \in (0, 1), \tag{31}
\]

implying that

\[
\| T_{2i}^*T_1^*T_{3ci}^* + \lambda Q_1 + (1 - \lambda)Q_2 \| = \mu, \quad \forall \lambda \in [0, 1] \tag{32}
\]

In particular for \( Q^* = \frac{Q_1 + Q_2}{2} \), let \( \Gamma := T_{2i}^*T_1^*T_{3ci}^* - Q^* \), and note that

\[
\| \Gamma \| = \left\| \Gamma \pm \frac{Q_1 - Q_2}{2} \right\| = \mu \tag{33}
\]

By the parallelogram law we have for all \( x \in \ell^2 \),

\[
\left\| \left( \Gamma - \frac{Q_1 - Q_2}{2} \right) T_o x \right\|_2^2 + \left\| \left( \Gamma + \frac{Q_1 - Q_2}{2} \right) T_o x \right\|_2^2 = 2\| (T_o x) \|_2^2 + 2 \left\| \frac{Q_1 - Q_2}{2} (T_o x) \right\|_2^2 \tag{34}
\]

Since each of the operators \( \Gamma \) and \( \Gamma \pm \frac{Q_1 - Q_2}{2} \) is an isometry on the range of \( T_o \), we have

\[
\left\| \left( \Gamma \pm \frac{Q_1 - Q_2}{2} \right) T_o x \right\|_2 = \| T_o x \|_2 = \| T_o x \|_2
\]

which implies that \( \| \frac{Q_1 - Q_2}{2} (T_o x) \| = 0, \forall x \in \ell^2 \), that is, we have necessarily \( Q_1 \equiv Q_2 \), showing uniqueness on the range of \( T_o \). \( \blacksquare \)
Next, we show that the operator $T_o$ in the dual maximization is unique. To see this, again suppose that there exists another operator $T'_o$ such that

$$
\|T^*_o T_i T^*_{3ci} - Q_o\| = tr \left( (T^*_o T_i T^*_{3ci} - Q_o)T_o \right) = tr \left( (T^*_o T_i T^*_{3ci} - Q_o)T'_o \right)
$$

By the proof of Theorem 2 there exist orthonormal sequences $\{\phi_i\}$ and $\{\psi_j\}$ such that

$$
T_o = \sum_j \lambda_j \phi_j \otimes \psi_j, \quad \|T_o\|_1 = \sum_j \lambda_j = 1.
$$

Moreover since $(T^*_o T_i T^*_{3ci} - Q_o)T_o = \sum_j \lambda_j \psi \otimes (T^*_o T_i T^*_{3ci} - Q_o)\psi_j$, we have then

$$
tr \left( (T^*_o T_i T^*_{3ci} - Q_o)T_o \right) = \sum_j \lambda_j < \phi_j, (T^*_o T_i T^*_{3ci} - Q_o)\psi_j >, \quad \text{and} \quad \|T^*_o T_i T^*_{3ci} - Q_o\| = \mu \Rightarrow < \phi_j, (T^*_o T_i T^*_{3ci} - Q_o)\psi_j >= \mu, \quad \forall j.
$$

That is, $(T^*_o T_i T^*_{3ci} - Q_o)\psi_j = \mu \phi_j, \quad \forall j.$ The latter shows that $(T^*_o T_i T^*_{3ci} - Q_o)T_o = (T^*_o T_o)^{1/2}$. Similarly, $(T^*_o T_i T^*_{3ci} - Q_o)T'_o = (T^*_o T'_o)^{1/2}$.

Now suppose that $R := T_o - T'_o \neq 0, R \in S$, i.e., $R$ is strictly causal, then $\|T_o\|_1 = \|R + T'_o\|_1 = \|T'_o\|_1 = 1$, and

$$
tr \left( (T^*_o T_i T^*_{3ci} - Q_o)(T'_o + R) \right) = tr \left( (T^*_o T_i T^*_{3ci} - Q_o)T'_o \right).
$$

Therefore, $\|T'_o\|_1 = \|T'_o + R\|_1 = tr \left( (T^*_o T_i T^*_{3ci} - Q_o)T'_o \right)$, and then

$$
(T^*_o T_i T^*_{3ci} - Q_o)R = (T^*_o T_i T^*_{3ci} - Q_o)(T'_o + R) - (T^*_o T_i T^*_{3ci} - Q_o)T'_o = ((T'_o + R)^* (T'_o + R))^{1/2} - (T'_o T'_o)^{1/2}
$$

showing that $(T^*_o T_i T^*_{3ci} - Q_o)R$ is self-adjoint. But $(T^*_o T_i T^*_{3ci} - Q_o)R$ is strictly-causal, that is, belongs to $S$, likewise for its adjoint $((T^*_o T_i T^*_{3ci} - Q_o)R)^* \in S$. This implies that

$$
((T^*_o T_i T^*_{3ci} - Q_o)R)^* ((T^*_o T_i T^*_{3ci} - Q_o)R) \in S
$$

Since $B_c(\ell^2, \ell^2)$ is the annihilator of $S$ it follows that

$$
tr \left( ((T^*_o T_i T^*_{3ci} - Q_o)R)^* ((T^*_o T_i T^*_{3ci} - Q_o)R) \right) = 0
$$
Thus, \((T_{2i}T_1T_{3ci}^* - Q_o)R = 0\), and from \([35]\) we have \(((T_o' + R)(T_o' + R))^{\frac{1}{2}} = (T_oT_o')^{\frac{1}{2}}\). Since \(T_{2i}T_1T_{3ci}^* - Q_o\) is an isometry on the ranges of \(T_o'\) and \(T_o = T_o' + R\), we have for every \(x \in \ell^2\),

\[
\|(T_o' + R)x\|_2 = \|(T_{2i}T_1T_{3ci}^* - Q_o)(T_o' + R)x\|_2 \tag{41}
\]

\[
= \|(T_{2i}T_1T_{3ci}^* - Q_o)T_o'x\|_2 = \|T_o'x\|_2 \tag{42}
\]

In the same manner, we get \(\|(T_o' - R)x\|_2 = \|T_o'x\|_2\). Next, apply the parallelogram law to get the equality \(\|(T_o' + R)x\|_2^2 + \|(T_o' - R)x\|_2^2 = 2\|T_o'x\|_2^2 + 2\|Rx\|_2^2\), from which we deduce that \(\|Rx\|_2 = 0\) for all \(x \in \ell^2\), that is, we must have \(R \equiv 0\) and then \(T_o \equiv T_o'\), showing uniqueness of \(T_o\). 

In the next section, we relate our problem to an LTV operator analogous to the Hankel operator, which is known to solve the standard optimal \(H^\infty\) control problem \([24, 46]\).

### 7 Triangular Projections and Hankel Forms

Let \(C_2\) denote the class of compact operators on \(\ell^2\) called the Hilbert-Schmidt or Schatten 2-class \([14, 3]\) under the norm,

\[
\|A\|_2 := \left(\text{tr}(A^*A)\right)^{\frac{1}{2}}, \quad A \in C_2
\]  

Define the space \(A_2 := C_2 \cap B_c(\ell^2, \ell^2)\), then \(A_2\) is the space of causal Hilbert-Schmidt operators. The orthogonal projection \(P\) of \(C_2\) onto \(A_2\) is the lower triangular truncation, and is analogous to the standard positive Riesz projection (for functions on the unit circle). Following \([30]\) an operator \(X\) in \(B(\ell^2, \ell^2)\) determines a Hankel operator \(H_X\) on \(A_2\) if

\[
H_XA = (I - P)XA, \quad \text{for } A \in A_2
\]  

In the sequel we show that \(\mu\) is equal to the norm of a particular LTV Hankel operator. To this end, we need first to characterize all atoms, denoted \(\Delta_n\), of \(B_c(\ell^2, \ell^2)\) as \(\Delta_n := \tilde{Q}_{n+1} - \tilde{Q}_n, \quad n = 0, 1, 2, \cdots\). Write \(C^+\) for the set of operators \(A \in B_c(\ell^2, \ell^2)\) for which \(\Delta_n A \Delta_n = 0\), \(n = 0, 1, 2, \cdots\) and let \(C_2^+ := C_2 \cap C^+\). In \([30, 29]\) it is shown that any operator \(A\) in \(B_c(\ell^2, \ell^2) \cap C_1\) admits a Riesz
factorization, that is, there exist operators $A_1$ and $A_2$ in $A_2$ such that $A$ factorizes as

$$A = A_1 A_2, \quad \text{and} \quad \|A\|_1 = \|A_1\|_2 \|A_2\|_2$$

(45)

A Hankel form $[\cdot, \cdot]_B$ associated to a bounded linear operator $B \in \mathcal{B}(\ell^2, \ell^2)$ is defined by [30, 29]

$$[A_1, A_2]_B = tr(A_1 B A_2), \quad A_1, A_2 \in A_2$$

(46)

Since any operator in the preannihilator $S$ belongs also to $\mathcal{B}_c(\ell^2, \ell^2) \cap \mathcal{C}_1$, then any $A \in S$ factorizes as in (45). And if $\|A\|_1 \leq 1$, as on the RHS of (12), $A \in S$, the operators $A_1$ and $A_2$ both in $A_2$ can be chosen such that $\|A_1\|_2 \leq 1$, $\|A_2\|_2 \leq 1$, and $\Delta_n A_1 \Delta_n = 0$, $n = 0, 1, 2, \cdots$, that is, $A_1 \in C_2^+$ [29].

Now write $P_+$ for the orthogonal projection with range $A_2 \cap \mathcal{C}_+$. Introducing the notation $(B_1, B_2) = tr(B_2^* B_1)$, and the Hankel form associated to $B := T_{2i}^* T_1 T_{3ci}^*$, we have by a result in [29],

$$[A_1, A_2]_B = tr(B A_2 A_1) = (A_1, (BA_2)^*) = (P_+ A_1, (BA_2)^*) = (A_1, P_+(BA_2)^*) = (A_1, H_B^* A_2)$$

(47)

where $H_B$ is the Hankel operator $(I - P)B P$ associated with $B := T_{2i}^* T_1 T_{3ci}^*$. The Hankel operator $H_B$ belongs to the Banach space of bounded linear operators on $C_2$. Furthermore, we have

$$\|H_{T_{2i}^* T_1 T_{3ci}^*}\| = \sup_{\|A_2\|_2 \leq 1, A_2 \in C_2} \|H_B A_2\|_2 = \sup_{\|A_2\|_2 \leq 1, A_2 \in C_2} (A_1, H_B^* A_2)$$

(48)

We have then the following Theorem which relates the optimal performance $\mu$ to the induced norm of the Hankel operator $H_{T_{2i}^* T_1 T_{3ci}^*}$. The Theorem was announced in [31] without a proof.

**Theorem 3** Under assumptions (A1) and (A2) the following hold

$$\mu = \|H_{T_{2i}^* T_1 T_{3ci}^*}\| = \|(I - P)T_{2i}^* T_1 T_{3ci}^* P\|$$

(49)
Proof. Since by the previous discussion any operator $T \in S$ can be factored as $T = A_1A_2$, where $A_1 \in C_2^+$, $A_2 \in A_2$, $\|A_1\|_2 = \|A_2\|_2 \leq 1$, and $\|T\|_1 = \|A_1\|_2 \|A_2\|_2$, the duality identity (12) yields

$$
\mu = \sup_{T \in S, \|T\|_1 \leq 1} |tr(TT_i^*T_1^*T_3^{aci})| = \sup_{\|A_2\|_2 \leq 1, A_2 \in A_2} |tr(T_2^*TT_1^*T_3^{aci}A_2A_1)|
$$

$$
= \|H_{T_2^*T_1^*T_3^{aci}}\|, \quad \text{by (18)}
$$

By Theorem 2.1. [20] the Hankel operator is a compact operator if and only if $B$ belongs to the space $\mathcal{B}_c(\ell^2, \ell^2) + \mathcal{K}$. It follows in our case that under assumption (A2) $H_{T_2^*T_1^*T_3^{aci}}$ is a compact operator on $A_2$. A basic property of compact operators on Hilbert spaces is that they have maximizing vectors, that is, there exists an $A \in A_2$, $\|A\|_2 = 1$ such that $\|H_{T_2^*T_1^*T_3^{aci}}\| = \|H_{T_2^*T_1^*T_3^{aci}}A\|_2$, that is, a $A$ achieves the norm of $H_{T_2^*T_1^*T_3^{aci}}$. We can then deduce from (12) an expression for the optimal TV Youla parameter $Q_o$ as follows

$$
\|H_{T_2^*T_1^*T_3^{aci}}\| = \|H_{T_2^*T_1^*T_3^{aci}}A\|_2 = \|H_{T_2^*T_1^*T_3^{aci} - Q_oA}\|_2 = \|(I - \mathcal{P})(T_2^*T_1^*T_3^{aci} - Q_oA)\|_2
$$

$$
\leq \|T_2^*T_1^*T_3^{aci} - Q_oA\|_2 \leq \|T_2^*T_1^*T_3^{aci} - Q_o\|A\|_2 \leq \|T_2^*T_1^*T_3^{aci} - Q_o\| = \|H_{T_2^*T_1^*T_3^{aci}}\|
$$

All terms must be equal, and then

$$
\|(I - \mathcal{P})(T_2^*T_1^*T_3^{aci} - Q_oA)\|_2 = \|T_2^*T_1^*T_3^{aci} - Q_oA\|_2
$$

(50)

Since $T_2^*T_1^*T_3^{aci} - Q_oA = (I - \mathcal{P})(T_2^*T_1^*T_3^{aci} - Q_oA) + \mathcal{P}(T_2^*T_1^*T_3^{aci} - Q_oA)$ and $\mathcal{P}(T_2^*T_1^*T_3^{aci} - Q_oA) = 0$. The TV optimal $Q_o$ can then be computed from the following operator identity

$$
Q_oA = T_2^*T_1^*T_3^{aci}A - H_{T_2^*T_1^*T_3^{aci}}
$$

(51)

The upshot of these methods is that they lead to the computation of $\mu$ within desired tolerances by solving two finite dimensional convex optimizations.
8 Numerical Computation of the Optimal Solution

In this section a numerical solution based on duality theory is developed. If \( \{ e_n : n = 0, 1, 2, \cdots \} \) is the standard orthonormal basis in \( \ell^2 \), then \( \widetilde{Q}_n \ell^2 \) is the linear span of \( \{ e_k : k = n + 1, n + 2, \cdots \} \). The matrix representation of \( \textbf{A} \in \mathcal{B}_c(\ell^2, \ell^2) \) w.r.t. this basis is lower triangular. Note \( P_n = I - \widetilde{Q}_n \rightarrow I \) as \( n \rightarrow \infty \) in the strong operator topology (SOT). If we restrict the minimization in (12) over \( Q \in \mathcal{B}_c(\ell^2, \ell^2) \) to the span of \( \{ e_n : n = 0, 1, 2, \cdots, N \} \), that is, \( P_N \ell^2 =: \ell^2_N \), this yields a finite dimensional convex optimization problem in lower triangular matrices \( Q_N \) of dimension \( N \), that is,

\[
Q_N = \begin{pmatrix}
Q_{11} & 0 & 0 & \cdots & \cdots \\
Q_{21} & Q_{22} & 0 & \cdots & \cdots \\
\vdots & \vdots & \ddots & \ddots & \cdots \\
Q_{N1} & Q_{N2} & Q_{N3} & \cdots & Q_{NN}
\end{pmatrix}
\] (52)

and the optimization

\[
\mu_N := \inf_{Q_N \in \mathcal{B}_c(P_N \ell^2, P_N \ell^2)} \| T_{2i}^* T_1 T_{3ci}^* - Q_N \|
\] (53)

where

\[
T_{2i}^* T_1 T_{3ci}^* - Q_N = \begin{pmatrix}
T_{11} - Q_{11} & T_{12} & T_{13} & \cdots & T_{1N} & T_{1(N+1)} & \cdots \\
T_{21} - Q_{21} & T_{22} - Q_{22} & T_{23} & \cdots & T_{2N} & T_{2(N+1)} & \cdots \\
T_{31} - Q_{31} & T_{32} - Q_{32} & T_{33} - Q_{33} & \cdots & T_{3N} & T_{3(N+1)} & \cdots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots \\
T_{N1} - Q_{N1} & T_{N2} - Q_{N2} & T_{N3} - Q_{N3} & \cdots & T_{NN} - Q_{NN} & T_{N(N+1)} & \cdots \\
T_{(N+1)1} & T_{(N+1)2} & T_{(N+1)3} & \cdots & T_{(N+1)N} & T_{(N+1)(N+1)} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots 
\end{pmatrix}
\] (54)

where \( T_{ij} \) are fixed and correspond to the entries of \( T_{2i}^* T_1 T_{3ci}^* \), and \( Q_{ij} \) are variable. The optimization \( \mu_N \) overestimates \( \mu \), but \( \mu_N \downarrow \mu \), and results in upper bounds and suboptimal TV parameters \( Q_{N,0} \) and control laws, since \( Q \) is restricted to a proper subspace of \( \mathcal{B}_c(\ell^2, \ell^2) \). By Arveson distance formula [1] the minimization (53) is equal to

\[
\mu_N = \max_{1 \leq n \leq N} \| (I - \widetilde{Q}_n)(T_{2i}^* T_1 T_{3ci}^*)\widetilde{Q}_n \|
\] (55)
The degree of suboptimality can be computed explicitly as follows: Applying the same argument to the dual optimization on the RHS of (12), by restricting $S$ to the finite dimensional subspace

$$S_N := \{ T_N \in C_1(P_N \ell^2, P_N \ell^2) : (I - \tilde{Q}_n)T_N\tilde{Q}_{n+1} = 0, \text{ for all } 0 \leq n \leq N - 1 \}$$

(56)

In fact, with respect to the canonical basis $\{e_i, i = 0, 1, \cdots \}$ for $P_N \ell^2$ the subspace $S_N$ is nothing but the space of $N \times N$ strictly lower triangular matrices.

The dual optimization becomes

$$\mu'_N := \sup_{T_N \in S_N, \|T_N\|_1 \leq 1} |tr(T_N(T_{2i}^*T_1^*T_{3ci}))|_N$$

(57)

where $(T_{2i}^*T_1^*T_{3ci})|_N$ is the restriction of the operator $T_{2i}^*T_1^*T_{3ci}$ to $P_N \ell^2$. The supremum in (57) is in fact a maximum, since (57) is a maximization of a continuous linear functional over a compact set.

The optimization (57) is a finite variable constrained convex optimization in the (strictly lowers entries of $T_N$), which yields lower bounds for $\mu$ since the dual optimization involves a supremum rather than an infimum, i.e., $\mu'_N \uparrow \mu$, and suboptimal $T_{N,o}$.

The optimization problem (57) is in fact a semi-definite programming problem (SDP) since $\|T_N\|_1 \leq 1$ if and only if there exist matrices $Y, Z \in \mathbb{R}^{N \times N}$ such that [42, 43]

$$\left( \begin{array}{cc}
Y & T_N \\
T_N^T & Z
\end{array} \right) \succeq 0, \quad trY + trZ \leq 2$$

(58)

The optimization (57) becomes then

$$\sup_{T_N \in S_N} \quad tr(T_N(T_{2i}^*T_1^*T_{3ci}))|_N$$

subject to (58)

$$P_nT_N(I - P_{n+1}) = 0, \quad n = 0, 1, \cdots, N - 1$$

where $P_n$ is the $N \times N$ truncation matrix having as its first $n$ columns the standards basis vectors $\{e_i\}, i = 1, 2, \cdots, n$ and the remaining columns as $N \times 1$ zero vectors, that is,

$$P_n = (e_1, e_2, \cdots, e_n, 0, \cdots, 0)$$
Under assumption (A2) $T_2^* T_1 T_3 c_i$ and the optimal $Q_o$ are compact, and by a compactness argument we have that $Q_{N,o} \rightarrow Q_o$ in the operator topology, i.e., $\|Q_o - Q_{N,o}\| \rightarrow 0$ as $N \rightarrow \infty$. Likewise $T_{N,o} \rightarrow T_o$ as $N \rightarrow 0$ in the trace class topology. Since $P_N \rightarrow I$ as $N \rightarrow \infty$ in the SOT. It is straightforward to show that these upper and lower bounds $\mu_N$ and $\mu'_N$ converge to the optimum $\mu$ as $N \rightarrow \infty$. These optimizations estimate $\mu$ within known tolerance and compute the corresponding LTV operators $Q_N$, which in turn allow the computation of LTV controllers $K$ through the Youla parametrization. Solving such problems are then applications of finite variable convex programming techniques.

Note the size of the truncations depends on the degree of accuracy desired. However, the size required for a particular accuracy can be exactly estimated by looking at the difference between the optimization (55) and the dual optimization (57). For periodic systems [40], say of period $q$, it suffices to take the first $q$ vectors $\{e_n\}_{n=1}^q$ of the standard basis, i.e., $N = q$. In this case and for finite horizon problems the finite dimensional convex optimizations yield exactly the optimal corresponding TV Youla parameter $Q_o$ and hence the optimal TV controller.

9 The Mixed Sensitivity Problem for LTV Systems

The mixed sensitivity problem for stable plants [44, 46] involves the sensitivity operator $T_1 := \begin{pmatrix} W \\ 0 \end{pmatrix}$, the complementary sensitivity operator $T_2 = \begin{pmatrix} W \\ V \end{pmatrix}$ $P$ and $T_3 := I$ which are all assumed to belong to $B_c(\ell^2, \ell^2 \times \ell^2)$, and is given by the optimization

$$\mu_o = \inf_{Q \in B_c(\ell^2, \ell^2)} \left\| \begin{pmatrix} W \\ 0 \end{pmatrix} - \begin{pmatrix} W \\ V \end{pmatrix} PQ \right\|$$

where $\| \cdot \|$ stands for the operator norm in $B(\ell^2, \ell^2 \times \ell^2)$ given by

$$\|B\| = \sup_{\|x\|_{\ell^2} \leq 1, x \in \ell^2} \left( \|B_1 x\|_{\ell^2}^2 + \|B_2 x\|_{\ell^2}^2 \right)^{\frac{1}{2}}, \quad B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$$

The optimization problem (60) can be expressed as a distance problem from the operator $T_1$ to the subspace $\mathcal{S} = T_2 P B_c(\ell^2, \ell^2)$ of $B(\ell^2, \ell^2 \times \ell^2)$.

To ensure closedness of $\mathcal{S}$, we assume that $W^* W + V^* V > 0$, i.e., $W^* W + V^* V$ as an operator acting on $\ell^2$ is a positive operator. Then there exists an outer spectral factorization $\Lambda_1 \in B_c(\ell^2, \ell^2)$,
invertible in $\mathcal{B}_c(\ell^2, \ell^2)$ such that $\Lambda_1^* \Lambda_1 = W^* W + V^* V \quad [11]$. Consequently, $\Lambda_1 P$ as a bounded linear operator in $\mathcal{B}_c(\ell^2, \ell^2)$ has an inner-outer factorization $U_1 G$, where $U_1$ is inner and $G$ an outer operator defined on $\ell^2 \quad [3]$.

Next we assume (A3) $G$ is invertible, so $U_1$ is unitary, and the operator $G$ and its inverse $G^{-1} \in \mathcal{B}_c(\ell^2, \ell^2)$. (A3) is satisfied when, for e.g., the outer factor of the plant is invertible. Let $R = T_2 \Lambda_1^{-1} U_1$, assumption (A3) implies that the operator $R^* R \in \mathcal{B}(\ell^2, \ell^2)$ has a bounded inverse, this ensures closedness of $\mathcal{S}$. According to Arveson (Corollary 2, [1]), the self-adjoint operator $R^* R$ has a spectral factorization of the form: $R^* R = \Lambda^* \Lambda$, where $\Lambda$, $\Lambda^{-1} \in \mathcal{B}_c(\ell^2, \ell^2)$. Define $R_2 = R \Lambda^{-1}$, then $R_2^* R_2 = I$, and $\mathcal{S}$ has the equivalent representation, $\mathcal{S} = R_2 \mathcal{B}_c(\ell^2, \ell^2)$. After ”absorbing” $\Lambda$ into the free parameter $Q$, the optimization problem (60) is then equivalent to:

$$\mu_o = \inf_{Q \in \mathcal{B}_c(\ell^2, \ell^2)} \| T_1 - R_2 Q \|$$

(62)

To solve the TV optimization (62) it suffices to apply the duality results of section $\mathbb{3}$. The latter yields the predual space of $\mathcal{B}(\ell^2, \ell^2 \times \ell^2)$ under trace duality, as the Banach space isometrically isomorphic to

$$\mathcal{C}_1 := \left\{ B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} : B_i \in \mathcal{C}_1 \right\}$$

(63)

under the norm

$$\| B \|_{11}^2 := tr(B_1^* B_1 + B_2^* B_2)^{\frac{1}{2}}$$

(64)

The preannihilator $\perp \mathcal{S}$ of $\mathcal{S}$ is characterized in the following Lemma.

**Lemma 2** The preannihilator $\perp \mathcal{S}$ can be computed as

$$\perp \mathcal{S} = R_2 \mathcal{S} \oplus (I - R_2 R_2^*) \mathcal{C}_1$$

(65)

where $\oplus$ denotes the direct sum of two subspaces.

**Proof.** To show (65) notice that $tr(\Phi^* T) = 0$, $\forall T$ in $\perp \mathcal{S}$ for $\Phi \in \mathcal{B}(\ell^2, \ell^2 \times \ell^2)$, is equivalent to $\Phi^* (I - RR^*) = 0$ and $\Phi^* R = A^*$ for some $A \in \mathcal{B}_c(\ell^2, \ell^2)$. This implies that $\Phi^* = A^* R^*$. By taking the adjoints we get $\Phi = R_2 A \in \mathcal{S}$.

24
The following lemma is a consequence of (8).

**Lemma 3** Under assumption (A3) there exists at least one optimal TV operator \( Q_o \in B_c(\ell^2, \ell^2) \) s.t.

\[
\mu_o = \|T_1 - R_2 Q_o\| = \sup_{s \in \mathfrak{S}, \|s\|_{\ell^1} \leq 1} |\operatorname{tr} T_1^* s| \tag{66}
\]

Note that Lemma yields 3 not only shows existence of an optimal TV controller \( K \) through the Youla parameter \( Q_o \) under assumption (A3), but also leads to two dual finite variable convex programming problems using the same argument as in section 8, which solutions yield the optimal \( Q_o \) within desired tolerance. A TV allpass property holds also for the TV mixed sensitivity problem under the assumption that \( T_1 \) is a compact operator in \( B_c(\ell^2, \ell^2 \times \ell^2) \) and \( \mu_o > \mu'_oo \), where \( \mu'_oo := \inf_{Q \in \mathcal{K}} \|T_1 - R_2 Q\| \), i.e., when the causality condition of \( Q \) is removed. This generalizes the allpass property of section 4 and is summarized in the following lemma.

**Lemma 4** If \( T_1 \) is a compact operator and \( \mu_o > \mu'_oo \) then the optimal mixed sensitivity operator \( \frac{T_1 - R_2 Q_o}{\mu_o} \) is a partial isometry from \( \ell^2 \) onto \( \ell^2 \times \ell^2 \). The condition \( \mu_o > \mu'_oo \) is sharp, in the sense that if it does not hold there exists \( T_1 \) and \( R_2 \) such that the optimum is not allpass.

**Proof.** The proof is similar to the proof of Theorem 1 by working in the predual maximization. The details are omitted. If \( \mu_o > \mu'_oo \) does not hold, take for example \( V = 1 \) and \( W = \frac{1 - z^2}{2} \), in which case \( Q_o = 0 \) showing that the optimum is not allpass.

Lemma 4 is the TV version of the same notion known to hold in the LTI case as shown earlier in [27] using broadband theory and [47].

Next we show that the mixed sensitivity optimization is equal to the norm of a certain TV Hankel-Topelitz operator. In order to do this we use a standard trick in [11]. Since \( R_2 = \begin{pmatrix} R_{21} \\ R_{22} \end{pmatrix} \) is an

25
isometry, so is $U := \begin{pmatrix} R_2^* \\ I - R_2 R_2^* \end{pmatrix}$. Thus,

$$
\mu_o = \inf_{Q \in \mathcal{B}_c(l^2, l^2)} \| U (T_1 - R_2 Q) \| = \inf_{Q \in \mathcal{B}_c(l^2, l^2)} \left\| \begin{pmatrix} R_{21}^* W - Q \\ (I - R_{21} R_{21}^*) W \\ - R_{22} R_{21}^* W \end{pmatrix} \right\|
$$

Call $\Omega := \begin{pmatrix} (I - R_{21} R_{21}^*) W \\ - R_{22} R_{21}^* W \end{pmatrix}$. By analogy with the TV Hankel operator defined in section 7, we define the TV Toeplitz operator $T_{\Omega^* \Omega}$ associated to $\Omega^* \Omega$ as:

$$
T_{\Omega^* \Omega} A = \mathcal{P} \Omega^* \Omega A, \text{ for } A \in \mathcal{A}_2
$$

(67)

The following Theorem solves the problem in terms of a TV mixed Hankel-Toeplitz operator.

**Theorem 4** Under assumption (A3) the following holds

$$
\mu_o^2 = \| H_{R_{21}^* W} H_{R_{21}^* W} + T_{\Omega^* \Omega} \| \tag{68}
$$

where $H_{R_{21}^* W}$ is the Hankel operator associated to $R_{21}^* W$, i.e., $H_{R_{21}^* W} A = (I - \mathcal{P}) R_{21}^* W A$, for $A \in \mathcal{A}_2$.

**Proof.** To prove (68) define the following operator:

$$
\Gamma := \Pi_{\mathcal{A}_2 \times \mathcal{A}_2 \ominus R_2 \mathcal{A}_2} \begin{pmatrix} W \\ 0 \end{pmatrix} \tag{69}
$$

where $\Pi$ is the orthogonal projection from $\mathcal{A}_2 \times \mathcal{A}_2$ into the orthogonal complement of $\mathcal{A}_2$, $\mathcal{A}_2 \times \mathcal{A}_2 \ominus R_2 \mathcal{A}_2$. The orthogonal projection $\Pi$ can be shown to be given explicitly by [23]

$$
\Pi = I - R_2 \mathcal{P} R_2^* \tag{70}
$$

It follows from the Commutant Lifting Theorem for nest algebra (Theorem 20.22 [3]) that

$$
\mu_o = \inf_{Q \in \mathcal{B}_c(l^2, l^2)} \| T_1 - R_2 Q \| = \| \Gamma \| \tag{71}
$$
Since \( U \) is an isometry and thus preserves norms, operator pre-composition with \( U \) and using the explicit expression \((70)\) of \( \Pi \), straightforward computations show that \( \mu_o = \left\| \begin{pmatrix} H_{R^*} \circ \mathbf{w} \\ T_\Omega \end{pmatrix} \right\|_2 \), and therefore \((68)\) holds.

Theorem \(4\) generalizes the solution of the mixed sensitivity problem in terms of a mixed Hankel-Toeplitz operator in the LTI case \([27, 45, 47]\) to the TV case. This result also applies to solve the robustness problem of feedback systems in the gap metric \([38]\) in the TV case as outlined in \([11, 36]\), since the latter was shown in \([11]\) to be equivalent to a special version of the mixed sensitivity problem \((60)\).

10 Conclusion

The optimal disturbance rejection problem for LTV systems involves solving a shortest distance minimization problem in the space of bounded linear operators. Dual and predual representations show existence of optimal TV controllers through the Youla parametrization. Under specific conditions, the optimal solution is compact and satisfies a “time-varying” allpass or flatness condition. The proposed approach leads “naturally” to a numerical solution based on finite variable convex programming, which involves solving a SDP problem for the predual and a search over lower triangular matrices for the dual problem. Moreover, a solution based on a TV compact Hankel operator is proposed. The latter leads to an operator identity for the optimal TV Youla parameter \( Q_o \). A generalization to the mixed-sensitivity problem is carried out that draws certain analogies with the LTI case.

Future work includes investigation of the numerical solutions by semi-define programming and finding efficient algorithms to take advantage of the special structure (lower triangular) of \( Q_o \) to solve numerically the dual problem.

References

[1] Arveson W. Interpolation problems in nest algebras, Journal of Functional Analysis, 4 (1975) pp. 67-71.
[2] Chapellat H., Dahleh M. Analysis of time-varying control strategies for optimal disturbance rejection and robustness, IEEE Transactions on Automatic Control, 37 (1992) 1734-1746.
[3] Davidson K.R. Nest Algebras, Longman Scientific & Technical, UK, 1988.
[4] G. Tadmor and M. Verma, Factorization and the Nehari Theorem in Time-Varying Systems, Math. Control Signals Systems, Math. of Control Signals and Systems, no. 5, pp. 419-452, 1992.
[5] Djouadi S.M. Optimal Robust Disturbance Attenuation for Continuous Time-Varying Systems, International journal of robust and non-linear control, vol. 13, (2003) 1181-1193.

[6] Douglas R.G. *Banach Algebra Techniques in Operator Theory*, Academic Press, NY, 1972.

[7] Khammash M., Dahleh M. Time-varying control and the robust performance of systems with structured norm-bounded perturbations, *Proceedings of the IEEE Conference on Decision and Control*, Brighton, UK, 1991.

[8] Khammash M., J.B. Pearson J.B. Performance robustness of discrete-time systems with structured uncertainty, *IEEE Transactions on Automatic Control*, 36 (1991) 398-412.

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

[10] A. Ichikawa, Quadratic Games and $H^\infty$-Type Problems for Time-Varying Systems, Int. J. Control, vol. 54, pp. 1249-1271, 1991.

[11] Feintuch A. *Robust Control Theory in Hilbert Space*, Springer-Verlag, vol. 130, 1998.

[12] Feintuch A., Francis B.A. Uniformly Optimal Control of Linear Time-Varying Systems, Systems & Control Letters, vol. 5, (1984) 67-71.

[13] Feintuch A., Francis B.A. Uniformly Optimal Control of Linear Feedback Systems, Systems & Control Letters, vol. 21, (1985) 563-574.

[14] Schatten R. *Norm Ideals of Completely Continuous Operators*, Springer-Verlag, Berlin, Gottingen, Heidelberg, 1960.

[15] Shamma J.S. Robust stability with time-varying structured uncertainty, *Proceedings of the IEEE Conference on Decision and Control*, 3163-3168, 1992.

[16] Shamma J.S., Dahleh M.A. Time-varying versus time-invariant compensation for rejection of persistent bounded disturbances and robust stabilization, *IEEE Transactions on automatic Control*, 36 (1991) 838-847.

[17] P. DeWilde A-J. Van der Veen, *Time-Varying Systems and Computations*, Springer-Verlag, 1998.

[18] Zames G., Owen J.G. Duality theory for MIMO robust disturbance rejection, *IEEE Transactions on Automatic Control*, 38 (1993) 743-752.

[19] S.M Djouadi, MIMO Disturbance and Plant Uncertainty Attenuation by Feedback, *IEEE Transactions on Automatic Control*, Vol. 49, No. 12, pp. 2099-2112, December 2004.

[20] S.M. Djouadi, Operator Theoretic Approach to the Optimal Two-Disk Problem, *IEEE Transactions on Automatic Control*, Vol. 49, No. 10, pp. 1607-1622, October 2004.

[21] S.M. Djouadi, Disturbance rejection and robustness for LTV Systems, Proceedings of the American Control Conference, pp. 3648-2653-6080, June 14-26, 2006.

[22] S.M. Djouadi and Y. Li, On the computation of the gap metric for LTV systems, Systems & Control Letters, 56, pp. 753–758, 2007.
[23] S.M. Djouadi and C.D. Charalambous, Time-Varying Optimal Disturbance Minimization in Presence of Plant Uncertainty, SIAM J. Control Optimization, Vol. 48, No. 5, pp. 3354 - 3367, 2010.

[24] Francis B.A., Doyle J.C. Linear Control Theory with an $H^\infty$ Optimality Criterion, SIAM J. Control and Optimization, vol. 25, (1987) 815-844.

[25] Francis B.A. A Course in $H^\infty$ Control Theory, Springer-Verlag, 1987.

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

[27] M. Verma and E. Jonckheere, $L^\infty$-Compensation with Mixed Sensitivity as a Broadband Matching Problem, Systems & Control Letters, vol. 4, pp. 125-129, 1984.

[28] Holmes R., Scranton B., Ward J., Approximation from the Space of Compact Operators and other $M$-ideals, Duke Math. Journal, vol. 42, (1975) 259-269.

[29] Power S. Commutators with the Triangular Projection and Hankel Forms on Nest Algebras, J. London Math. Soc., vol. 2, (32), (1985) 272-282.

[30] Power S.C. Factorization in Analytic Operator Algebras, J. Func. Anal., vol. 67, (1986) 413-432.

[31] S.M. Djouadi and C.D. Charalambous, On Optimal Performance for Linear-Time Varying Systems, Proc. of the IEEE 43th Conference on Decision and Control, Paradise Island, Bahamas, pp. 875-880, December 14-17, 2004

[32] Garnett, J.B., Bounded Analytic Functions, Academic Press, 1981.

[33] Feintuch A., Saeks R. System Theory: A Hilbert Space Approach, Academic Press, N.Y., 1982.

[34] Peters M.A. and Iglesias P.A., Minimum Entropy Control for Time-varying Systems, Boston, Birkhuser, 1997.

[35] C. Foias, A.E. Frazho, I. Gohberg, and M.A. Kaashoek, Metric Constrained Interpolation, Commutant Lifting and Systems, Birkhuser, 1998.

[36] Foias C., Georgiou T. and Smith M.C., Robust Stability of Feedback Systems: A Geometric Approach Using The Gap Metric, SIAM J. Control and Optimization, vol. 31, No.6, pp. 1518-1537, 1993.

[37] Ravi R., Nagpal K.M. and Khargonekar P.P., $H^\infty$ Control of Linear Time-Varying Systems: A State Space Approach, SIAM J. Control and Optimization, vol. 29, No.6, pp. 1394-1413, 1991.

[38] T.T. Georgiou and M.C. Smith, Robust Stabilization in the Gap Metric: Controller Design for Distributed Plants, IEEE Trans. on Automatic Control, vol. 37, No. 8, pp. 1133-1143, 1992.

[39] Dale W.N. and Smith M.C., Stabilizability and Existence of System Representation for Discrete-Time-Varying Systems, SIAM J. Control and Optimization, vol. 31, No.6, pp. 1538-1557, 1993.

[40] Dullerud G.E. and Lall S., A new approach for analysis and synthesis of time-varying systems IEEE Trans. on Automatic Control, Vol.: 44 , Issue: 8 , pp. 1486 - 1497, 1999.

[41] C. Pirie and G.E. Dullerud, Robust Controller Synthesis for Uncertain Time-Varying Systems. SIAM J. Control Optimization., Vol. 40, No. 4, pp. 13121331, 2002.
[42] M. Fazel, H. Hindi and S.P. Boyd, A Rank Minimization Heuristic with Application to Minimum Order System Approximation, Proceedings American Control Conference, volume 6, pages 4734-4739, June 2001.

[43] L. Vandenberghe and S. Boyd, Semidefinite Programming, SIAM Review, 38(1): 49-95, March 1996.

[44] J.C. Doyle, B.A. Francis and A.R. Tannenbaum, Feedback Control Theory, Macmillan, NY, 1990.

[45] C. Foias, H. Ozbay and A.R. Tannenbaum, Robust Control of Infinite Dimensional Systems, Springer-Verlag, Berlin, Heidelberg, New York, 1996.

[46] Zhou K. and Doyle J.C., Essentials of Robust Control, Prentice Hall, 1998.

[47] Djouadi S.M. and Birdwell J.D., Duality theory of the optimal two-block $H^\infty$ problem, Int. J. Robust Nonlinear Control, vol. 18, pp. 1302-1316, 2008.

[48] G. Zames and B.A. Francis, Feedback, Minimax Sensitivity, and Optimal Robustness, IEEE Trans. on Automatic Control, Vol.: 46, Issue: 5, pp. 585-601, May 1983.