A Robust $H_2$ Performance Problem for Discrete-Time Systems

H. T. Toivonen  
Department of Engineering  
Åbo Akademi University  
FIN-20500 Åbo (Turku), FINLAND.  
Fax: Int+358-21-2654479  
htoivone@abo.fi

May 27, 1996

Abstract

A robust performance problem is studied for discrete-time systems with norm-bounded uncertainties. The control performance is characterized in a deterministic setting as the average value of a quadratic cost with respect to disturbances in $l_2$. For linear time-invariant systems, the average cost defined in this way reduces to the $H_2$ norm of the system transfer function. A robust control problem is introduced by minimizing an upper bound on the worst-case performance for an uncertain plant. The control problem involves a particular mixed $H_2/H_\infty$ problem as a subproblem. For this problem, synthesis equations in the form of coupled nonlinear matrix equations are derived, which correspond to similar equations previously obtained for the continuous-time problem.

Keywords: Discrete time systems; mixed $H_2/H_\infty$ control; optimal control; robust control; worst-case design.
1 Introduction

In many control problems, it is natural to characterize the control performance as the $H_2$ norm of the closed-loop transfer function. For example, when the disturbance can be described as a white noise process, the $H_2$ norm defines the sum of the output variances. It is, however, usually not sufficient to consider the nominal control performance only, but robustness against process uncertainties should also be taken into account. Robustness against plant uncertainties which are assumed to be norm-bounded operators on $l_2$ (or, for continuous-time systems, on $L_2$) can be quantified via an $H_\infty$ norm defined for the uncertain system.

The problem of designing controllers which achieve good $H_2$ performance and have good robustness properties has been approach by introducing various mixed $H_2/H_\infty$ problems. In particular, it has been shown that an upper bound on robust $H_2$ performance with respect to norm-bounded plant uncertainties can be characterized in terms of a particular kind of mixed $H_2/H_\infty$ cost (Zhou et al., 1990, 1994; Stoorvogel, 1993; Toivonen and Pensar, 1995). For continuous-time systems, this characterization has been proposed for the design of robust controllers for both linear time-invariant uncertainties (Zhou et al., 1990, 1994; Toivonen and Pensar, 1995), and nonlinear, time-varying uncertainties (Stoorvogel, 1993). A similar approach can be applied to the design of tracking controllers which achieve robust performance (Toivonen and Pensar, 1996).

The upper bound on the worst-case $H_2$ performance, which has been used in the above studies, can be interpreted as a Lagrangian-type function associated with a constrained quadratic cost. The robust performance problem then consists of finding both the minimizing controller and an optimal, possibly frequency dependent, Lagrangian multiplier weight. The mixed $H_2/H_\infty$ cost of the robust performance problem can be regarded as a dual version of a cost due to Bernstein and Haddad (1989), in the sense that it is equal to the Bernstein-Haddad cost of the dual system and vice versa. The structure of the optimal mixed $H_2/H_\infty$ controller is known, and it can be characterized by a set of synthesis equations in the form of coupled matrix equations (Bernstein and Haddad, 1989; Doyle et al., 1989, 1994; Yeh et al., 1992). An alternative, numerically appealing approach is to formulate the problem as a convex optimization problem (Khargonekar and Rotea, 1991).

For discrete-time systems, the robust $H_2$ performance problem and the associated mixed $H_2/H_\infty$ control problem have been studied much less. For the discrete mixed $H_2/H_\infty$ problem, the synthesis of optimal fixed-order controllers (Haddad et al., 1991; Davis et al., 1994) and optimal state-feedback controllers for both the primal problem (Mustafa, 1991) and the dual problem (Peters, 1993) have been studied. Kaminer et al. (1993) have shown that the optimal control performance is achieved by a controller with state dimension which does not exceed the plant dimension, and they also give a procedure for finding an optimal controller via convex optimization. However, the
discrete-time counterparts of the general synthesis equations for the optimal controller are not available, cf. the discussion in (Kaminer et al., 1993; Davis et al., 1994).

In this paper, the discrete-time robust $H_2$ performance problem and the associated mixed $H_2/H_{\infty}$ control problem are studied. One of the main contributions of this paper consists of the derivation of synthesis equations for the discrete mixed $H_2/H_{\infty}$ control problem. These consist of a set of coupled nonlinear matrix equations, and are analogous to the synthesis equations of the continuous-time problem. The robust $H_2$ performance problem studied in this paper consists of minimizing the mixed $H_2/H_{\infty}$ cost with respect to both the controller and a Lagrange multiplier type scaling filter. For this problem, a two stage approach is proposed, in which the cost is alternately reduced with respect to the controller and the scaling filter. The procedure is similar to an approach which has been applied to the continuous-time case (Toivonen and Pensar, 1995). It has also some similarities with two-stage procedures applied in structured singular value synthesis, such as the 'DK-iteration'. The major difference is that in the robust $H_2$ problem, the calculation of the optimal scaling filter involves the minimization of a mixed $H_2/H_{\infty}$ cost.

A second contribution of this paper is a time-domain characterization of the average control performance in a deterministic setting. This is achieved by forming the average of a quadratic cost over all disturbances in $l_2$. For finite-horizon problems, the resulting performance measure is equivalent to the Hilbert-Schmidt norm of the system operator. In the infinite-horizon, time-invariant, case, a limiting argument reduces the performance measure to the standard $H_2$ norm of the transfer function.

The paper is organized as follows. In Section 2, a deterministic, time-domain average performance measure is defined and its relation to the $H_2$ norm is discussed. Upper bounds on the worst-case average performance measure for uncertain plants are introduced in Section 3. The solution of the discrete-time mixed $H_2/H_{\infty}$ problem induced by the robust performance problem is presented in Section 4. In Section 5 an iterative procedure for solving the robust performance control problem is studied, and an illustrative numerical example of the proposed design procedure is presented in Section 6.

2 An average performance measure

In this section an average quadratic performance measure is introduced in a deterministic setting. The measure will first be developed for the finite-horizon case, and is then generalized to the infinite-horizon situation.
2.1 The finite-horizon case

Consider a discrete-time system $z = \mathcal{P} v$ given by

$$
\begin{align*}
    x_{k+1} &= F x_k + G v_k, \quad x_0 = 0 \\
    z_k &= H x_k + J v_k, \quad k = 0, \ldots, N.
\end{align*}
$$

Here $v \in l_2^m(0, N)$ and $z \in l_2^p(0, N)$ are the input disturbance the output variable, respectively. The performance measure will be defined in terms of the $l_2$-norm of the output $z$. Since the output depends on the disturbance $v$, we should specify the disturbance or the set of disturbances for which the output norm is to be evaluated. Typically, the plant may be subject to a range of disturbances, and it is then natural to consider its average response to the set of disturbances. In order to obtain a quantitative measure of the average response of the system $\mathcal{P}$, the following operator norm is introduced.

**Definition 2.1** Let $X$ and $Y$ be Hilbert spaces and consider a linear operator $\mathcal{T} : X \to Y$. Let $\{\eta_i\}$ be any orthonormal basis of $X$. Then the quantity

$$
\|\mathcal{T}\|_{HS}^2 := \sum_i \|\mathcal{P}\eta_i\|_2^2
$$

is basis-independent and $\|\mathcal{T}\|_{HS}$ is called the Hilbert-Schmidt norm of the operator $\mathcal{T}$.

Note that the Hilbert-Schmidt norm is a natural generalization of the Frobenius norm of a matrix to general linear operators.

The average performance for the system $\mathcal{P}$ taken over all disturbances in $l_2(0, N)$ is defined in terms of the Hilbert-Schmidt norm as

$$
\mathcal{J}_{(0, N)}(\mathcal{P}) := \frac{1}{N + 1} \|\mathcal{P}\|_{HS}^2.
$$

In practice the disturbances in the average cost defined according to (3) and (2) should be weighted to reflect their relative importance. This weighting can, however, be absorbed in the system equation (1) via appropriate weighting filters, and hence the definition (3) of the average cost is not restrictive.

The following lemma gives some explicit formulae for the Hilbert-Schmidt norm of the system operator $\mathcal{P}$.

**Lemma 2.1** Consider the system operator $\mathcal{P}$ described by equation (1). The Hilbert-Schmidt norm defined by equation (2) is given by any of the following expressions:

$(a)$

$$
\|\mathcal{P}\|_{HS}^2 = \sum_{k,l=0}^N \text{tr}[G'(k,l)G(k,l)].
$$

4
where $G(\cdot, \cdot)$ denotes the impulse response function of $\mathcal{P}$,

$$
G(k, l) := \begin{cases} 
HF^{k-l-1}G & \text{if } k > l \\
J & \text{if } k = l \\
0 & \text{if } k < l,
\end{cases}
$$

(5)

(b) \quad \|\mathcal{P}\|_{HS}^2 = \sum_{k=0}^{N} \text{tr}(H P_k H' + J J')

(6)

where the matrices $P_k$ are given by the recursive equation

$$
P_{k+1} = F P_k F' + G G', \quad P_0 = 0,
$$

(7)

c) \quad \|\mathcal{P}\|_{HS}^2 = \sum_{k=0}^{N} \text{tr}(G' S_k G + J' J)

(8)

where the matrices $S_k$ are given by the recursive equation

$$
S_k = F' S_{k+1} F + H' H', \quad S_N = 0.
$$

Proof: (a) By standard properties of the Hilbert-Schmidt norm we have that if $(g_{ij})$ is a matrix representation of $\mathcal{P}$, $\|\mathcal{P}\|_{HS}$ can be expressed as

$$
\|\mathcal{P}\|_{HS} = \sum_{i,j} g_{ij}^2
$$

(10)

The result then follows from the fact that the system operator $\mathcal{P}$ has a matrix representation consisting of the blocks $\{G(k, l)\}$.

(b) The basis in the definition (2) applied to the operator $\mathcal{P}$ can be selected as $\{e_{li}\}$, where

$$
e_{li,k} = \begin{cases} 
\epsilon_i & \text{if } l = k \\
0 & \text{otherwise}
\end{cases}
$$

(11)

and $\epsilon_i$ is the $i$th unit vector in $\mathbb{R}^m$. It is then easy to verify that the right-hand side of equation (2) is given by (6).

(c) The result is obtained by applying (b) to the adjoint of $\mathcal{P}$ and the fact that the Hilbert-Schmidt norms of $\mathcal{P}$ and its adjoint are equal.

2.2 The infinite-horizon, time-invariant case

In the infinite-horizon, time-invariant case the Hilbert-Schmidt norm of the system operator is unbounded, and it can therefore not be used directly to define an average
performance measure. The average performance measure in (3) can, however, be generalized in a natural way as follows. Introduce the projection \( \Pi_{l_2(0,N)} \) from the space \( l_2(0, \infty) \) onto \( l_2(0, N) \), and define the limit

\[
J_{\text{average}}(\mathcal{P}) := \lim_{N \to \infty} \frac{1}{N+1} \| \Pi_{l_2(0,N)} \mathcal{P} \Pi_{l_2(0,N)} \|^2_{HS}.
\]  

(12)

**Lemma 2.2** Consider the system \( \mathcal{P} \) defined by equation (1). Assume that the matrix \( F \) is asymptotically stable. Then the limit in (12) exists and is given by

\[
J_{\text{average}}(\mathcal{P}) = \| \mathcal{P} \|^2_2
\]

(13)

where \( \| \mathcal{P} \|_2 \) denotes the \( H_2 \)-norm of \( \mathcal{P} \).

**Proof:** The result follows from (12), Lemma 2.1 and the definition of the \( H_2 \)-norm. \( \square \)

Lemma 2.2 and the relation (12) give a deterministic time-domain characterization of the \( H_2 \) norm as an average performance measure with respect to disturbances in \( l_2 \). It is interesting to study the connection between the characterization in (12) and other commonly used characterizations of the \( H_2 \) norm. In particular, it follows from the time-invariance of the system that average performance measure in (12) can be expressed as follows.

**Lemma 2.3** Consider the system \( \mathcal{P} \) defined by equation (1), and suppose that the matrix \( F \) is asymptotically stable. Then the average performance measure in (12) can be characterized as follows:

(a)

\[
J_{\text{average}}(\mathcal{P}) = \| \mathcal{P} \Pi_{l_2(0,0)} \|^2_{HS},
\]

(14)

where \( \Pi_{l_2(0,0)} \) is the projection from the space \( l_2^n(0, \infty) \) onto \( l_2^n(0, 0) = \mathbb{R}^m \),

(b)

\[
J_{\text{average}} = \sum_{j=1}^m \left[ \| z \|_{l_2}^2 \mid v = e_j \delta_{k,0} \right],
\]

(15)

where \( e_j, j = 1, \ldots, m \) denotes the \( j \)th unit vector in \( \mathbb{R}^m \), and \( \delta_0 \) is the unit impulse at \( k = 0 \),

(c)

\[
J_{\text{average}} = \frac{1}{2\pi} \int_0^{2\pi} \text{tr} [ \hat{G}(e^{i\omega}) \hat{G}(e^{-i\omega})] d\omega
\]

(16)

where \( \hat{G}(e^{i\omega}) := H(e^{i\omega}I - F)^{-1}G + J \).

**Proof:** As the Hilbert-Schmidt norm is independent of the basis used in (2), we can take the set \( \{ \eta_l \} \) as \( \eta_{j+m} = e_j \delta_l, \ l = 0, 1, \ldots, j = 1, \ldots, m \), where \( e_j, j = 1, \ldots, m \), denotes the \( j \)th unit vector in \( \mathbb{R}^m \), and \( \delta_l \) is the unit impulse at time instant \( l \). The characterizations (14) and (15) then follow from (12), the time-invariance of \( \mathcal{P} \) and the definition of the Hilbert-Schmidt norm. The characterization in (16) is the frequency-domain counterpart of (15). \( \square \)
3 A robust performance measure for uncertain plants

In this section the uncertain plant described by the interconnection in Fig. 3.1 is studied. Here $\Delta$ represents a plant uncertainty which is assumed to belong to a norm-bounded set $\mathcal{B}$, but is otherwise unknown. A robust performance measure is defined as the worst-case average performance of the uncertain plant taken over uncertainty blocks $\Delta$ in the set $\mathcal{B}$. An upper bound on the robust performance measure in the form of a mixed $H_2/H_\infty$ cost is then introduced.

The nominal plant is assumed to be stable, linear and time-invariant. It is assumed that the transfer function $\Delta \mathcal{P}_{z_0v}$ is strictly causal, i.e., either $\mathcal{P}_{z_0v}$ or $\Delta$ (or both) contain a time delay. This assumption is made because it is compliant with a large class of standard uncertainty descriptions, (cf. Sections 5 and 6), and it therefore appears to be the more relevant case to study. This property can then be exploited to construct a less conservative measure for robust performance. For clarity, only the strictly causal case will be considered in this paper. The assumption can, however, be relaxed if required.

Two kinds of uncertainties will be considered: the set of bounded linear time-invariant (LTI) operators,

$$
\mathcal{B}_{LTI}^\sigma := \left\{ \Delta \in H_\infty^{r \times r}, \|\Delta\|_\infty \leq \sigma \right\},
$$

and the set of bounded, possibly nonlinear and time-varying (NLTV) causal operators,

$$
\mathcal{B}_{NLTV}^\sigma := \left\{ \Delta : l_2^r \to l_2^r, \Delta \text{ causal, } \|\Delta\| \leq \sigma \right\}.
$$

For the uncertainty set $\mathcal{B}_{LTI}^\sigma$ a worst-case average cost is defined as

$$
I(\mathcal{P}, \mathcal{B}_{LTI}^\sigma) := \sup \left\{ J_{\text{average}}(\mathcal{F}(\mathcal{P}, \Delta)) : \Delta \in \mathcal{B}_{LTI}^\sigma, \Delta \mathcal{P}_{z_0v} \text{ strictly causal} \right\}
$$

where $\mathcal{F}(\mathcal{P}, \Delta)$ denotes the transfer function from the disturbance $v$ to the output $z_0$, $\mathcal{F}(\mathcal{P}, \Delta) := \mathcal{P}_{z_0v} + \mathcal{P}_{z_0w\Delta} \Delta(I - \mathcal{P}_{z_0w\Delta} \Delta)^{-1} \mathcal{P}_{z_0v}$.

The quantity (2) in Definition 2.1 is in general basis-dependent for non-linear operators, and the average performance measure is therefore not directly applicable to the nonlinear case. Therefore, a worst-case cost with respect to the uncertainty set $\mathcal{B}_{NLTV}^\sigma$ is defined via the characterization (15) as

$$
I(\mathcal{P}, \mathcal{B}_{NLTV}^\sigma) := \sup \left\{ \sum_{j=1}^{m} \left\| z_0 \right\|_2^2 : v = e_j \delta_0 \right\} : \Delta \in \mathcal{B}_{NLTV}^\sigma, \Delta \mathcal{P}_{z_0v} \text{ strictly causal} \right\}.
$$

Note that this cost still depends on the choice of the basis $\{e_j\}$ of $R^m$. This will, however, not affect the robust performance analysis of this section, since an upper
bound on the worst-case cost (20) will be introduced which holds for all choices of basis functions.

The evaluation of the robust performance measures (19) and (20) for the uncertainty sets $\mathcal{B}_{LT}^\varepsilon$ and $\mathcal{B}_{NLT}^\varepsilon$ is very hard. It is therefore well motivated to introduce more tractable upper bounds on the robust performance measures. Note that since $w_\Delta = \Delta z_\Delta$, it follows that $\Delta \in \mathcal{B}_{NLT}^\varepsilon$ implies

$$
\sum_{j=1}^m \left[ \| \hat{z}_\Delta \|_2^2 - \sigma^{-2} \| w_\Delta \|_2^2 \mid v = e_j \delta_0 \right] \geq 0,
$$

while $\Delta \in \mathcal{B}_{LT}^\varepsilon$ implies an inequality at each frequency,

$$
\sum_{j=1}^m \left[ \| \hat{z}_\Delta(e^{i\omega}) \|_2^2 - \sigma^{-2} \| \hat{w}_\Delta(e^{i\omega}) \|_2^2 \mid v = e_j \delta_0 \right] \geq 0, \text{ all } \omega \in [0, 2\pi],
$$

where $\hat{z}_\Delta(\cdot)$ and $\hat{w}_\Delta(\cdot)$ denote the Fourier-transforms of $z_\Delta$ and $w_\Delta$, respectively. Upper bounds on the robust performance measures can therefore be constructed by introducing a worst-case cost subject to the quadratic constraints (21) and (22). For this purpose, consider the system in Fig. 3.2, and define the worst-case cost

$$
\mathcal{L}(\mathcal{P}, d) := \sup_{\mathcal{T}} \left\{ \sum_{j=1}^m \left[ \frac{1}{2\pi} \int_0^{2\pi} \left( \| \hat{z}_0(e^{i\omega}) \|_2^2 + |d(e^{i\omega})|^2 (\| \hat{z}_\Delta(e^{i\omega}) \|_2^2 - \sigma^{-2} \| \hat{w}_\Delta(e^{i\omega}) \|_2^2) \right) d\omega \mid v = e_j \delta_0 \right] : w_\Delta = \mathcal{T} v, \mathcal{T} \text{ strictly causal} \right\}
\begin{align*}
&= \sup_{\mathcal{T}} \left\{ \sum_{j=1}^m \left[ \| \hat{z}_0 \|_2^2 - \sigma^{-2} \| \hat{w}_\Delta \|_2^2 \mid v = e_j \delta_0 \right] : w_\Delta = \mathcal{T} v, \mathcal{T} \text{ strictly causal} \right\}.
\end{align*}
$$

The cost (23) can be regarded as Lagrange-function type cost associated with the average cost (15) and the constraints (21) and (22). Here $d$ is a weight parameter which serves as a Lagrange multiplier associated with the constraints in (21) and (22). Depending on the constraint that is relevant, a real-valued or frequency dependent weight is appropriate. A real-valued Lagrange multiplier $d \in \mathbb{R}$ is associated with the constraint (21), whereas with the infinite-dimensional frequency domain constraint (22) one can associate a weighting filter $d$ such that $d, d^{-1} \in H_\infty$.

Note that the supremum in (23) is taken over strictly causal operators $\mathcal{T} : l_2 \to l_2$, such that $(\mathcal{T} v)_k$ is a function of $v_l, l < k$, but does not depend on $v_k$. This is in compliance with the strict causality condition in the performance measures (19) and (20).

The following upper bounds on robust performance are similar to corresponding continuous-time bounds (Zhou et al., 1990, 1994; Stoorvogel, 1993; Toivonen and Pemsar, 1995, 1996).
Theorem 3.1 Consider the system in Fig. 3.1. The worst-case costs (19) and (20) are bounded if the system is stable for all norm-bounded uncertainties \(\|\Delta\| \leq \sigma\). This holds if and only if \(P\) is stable and \(\|P_{z_{\Delta}w_{\Delta}}\|_{\infty} < \sigma^{-1}\).

In this case, there exists \(d \in \mathbb{R}\) such that the transfer function from the signal \(dw_{\Delta}\) to \([z_0', (dz_{\Delta})]'\) has \(H_{\infty}\)-norm less than \(\sigma^{-1}\). Moreover, the worst-case cost (23) defined for the system in Fig. 3.2 is then bounded, and

\[
\mathcal{I}(P, B_{\text{NLT}})^{\sigma} \leq \mathcal{L}(P, d), \quad d \in \mathbb{R}.
\]  

An upper bound for \(\mathcal{I}(P, B_{\text{LT}})^{\sigma}\) is given by

\[
\mathcal{I}(P, B_{\text{LT}})^{\sigma} \leq \mathcal{L}(P, d), \quad d, d^{-1} \in H_{\infty}.
\]

Proof: See Appendix A.

The upper bounds of Theorem 3.1 can be extended to include structured uncertainties in a straightforward way. Introduce the structured uncertainty set

\[
\Delta_s := \left\{ \text{diag}(\Delta_1, \ldots, \Delta_s), \Delta_i : \mathbb{R}_+^i \to \mathbb{R}_+^i, i = 1, \ldots, s \right\}
\]

and the corresponding structured linear and nonlinear uncertainty sets,

\[
B_{\text{LT}, s}^{\sigma} := B_{\text{LT}}^{\sigma} \cap \Delta_s
\]

\[
B_{\text{NLT}, s}^{\sigma} := B_{\text{NLT}}^{\sigma} \cap \Delta_s
\]

Define the set \(D_s\) of diagonal weighting matrices corresponding to the uncertainty structure in (26),

\[
D_s := \left\{ \text{diag}(d_1(z)I_{r_1}, \ldots, d_s(z)I_{r_s}), \quad d_i, d_i^{-1} \in H_{\infty} \right\}
\]

and for \(D(z) \in D_s\), define the worst-case cost function

\[
\mathcal{L}(P, D) := \sup_{\mathcal{T}} \left\{ \sum_{j=1}^{m} \left[ \left\| z_0 - \sigma^{-2} \| Dw_{\Delta} \|_2^2 \right. \right] : v = e_j\delta_0 \right\}.
\]

The following generalization of Theorem 3.1 to the case with structured uncertainties is then obtained.

Theorem 3.2 (a) Consider the system in Fig. 3.1 with norm-bounded structured uncertainties. The system is stable for all \(\Delta \in B_{\text{LT}, s}^{\sigma}\) if \(P\) is stable and there exists a diagonal scaling matrix \(D(z) \in D_s\) such that \(\|DP_{z_{\Delta}w_{\Delta}}D^{-1}\|_{\infty} < \sigma^{-1}\). In this case, there exists \(D(z) \in D_s\) such that the transfer function from the signal \(dw_{\Delta}\) to \([z_0', (Dz_{\Delta})]'\)
in Fig. 3.2 has $H_\infty$-norm less than $\sigma^{-1}$. Moreover, the worst-case cost (30) defined for the system in Fig. 3.2 is then bounded, and

$$
I(\mathcal{P}, \mathcal{B}_{LT,s}^\sigma) \leq \mathcal{L}(\mathcal{P}, D), \ D \in \mathcal{D}_s.
$$

(b) The statements in (a) apply to the set of structured, nonlinear, time-varying uncertainties $\Delta \in \mathcal{B}_{NLTV,s}^\sigma$ if the set $\mathcal{D}_s$ of scaling matrices is replaced by the set of real-valued $D(z) \in \mathcal{D}_s$. In this case,

$$
I(\mathcal{P}, \mathcal{B}_{LTI,s}^\sigma) \leq \mathcal{L}(\mathcal{P}, D), \ D \in \mathcal{D}_s, \ D \text{ real.}
$$

Theorems 3.1 and 3.2 motivate the introduction of the performance measures

$$
\mathcal{I}(\mathcal{P}, \mathcal{B}_{LTI,s}^\sigma) := \inf \left\{ \mathcal{L}(\mathcal{P}, D) : D \in \mathcal{D}_s \right\}
$$

$$
\mathcal{I}(\mathcal{P}, \mathcal{B}_{NLTV,s}^\sigma) := \inf \left\{ \mathcal{L}(\mathcal{P}, D) : D \in \mathcal{D}_s, \ D \text{ real} \right\}
$$

to represent upper bounds on the robust performance measures $\mathcal{I}(\mathcal{P}, \mathcal{B}_{LT,s}^\sigma)$ and $\mathcal{I}(\mathcal{P}, \mathcal{B}_{NLTV,s}^\sigma)$, respectively.

Next an explicit expression for the worst-case cost $\mathcal{L}(\mathcal{P}, D)$ is given. Define the signals

$$
w := Dw_\Delta, \ z := \begin{bmatrix} z_0 \\ Dz_\Delta \end{bmatrix}. \quad (35)
$$

Denote the plant $\mathcal{P}$ augmented with the signals (35) by $\mathcal{P}_D$,

$$
z = \mathcal{P}_D \begin{bmatrix} v \\ w \end{bmatrix}, \ \mathcal{P}_D = \begin{bmatrix} I & 0 \\ 0 & D \end{bmatrix} \mathcal{P} \begin{bmatrix} I & 0 \\ 0 & D^{-1} \end{bmatrix} \quad (36)
$$

and let it have the state-space representation

$$
x_{k+1} = Ax_k + B_0v_k + B_1w_k \\
z_k = Cx_k + D_0v_k + D_1w_k
$$

Introduce the worst-case cost

$$
\mathcal{J}(\mathcal{P}_D) := \sup_{\mathcal{T}} \sum_{j=1}^{\mathcal{m}} \left\{ \|z\|^2_2 - \gamma^2 \|w\|^2_2 \mid v = e_j \delta_0 : w = \mathcal{T}v, \ \mathcal{T} \text{ strictly causal} \right\}. \quad (38)
$$

The cost (30) is then given by $\mathcal{L}(\mathcal{P}, D) = \mathcal{J}(\mathcal{P}_D)$ with $\gamma := \sigma^{-1}$.

The cost (38) is a dual version of the mixed $H_2/H_\infty$ cost considered by Haddad et al. (1991) and Kaminer et al. (1993). It can be evaluated by standard discrete $H_\infty$ control theory (Stoorvogel, 1992) and linear optimal control theory.
Theorem 3.3 Consider the system $\mathcal{P}_D$ in (37) and the associated worst-case cost $\mathcal{J}(\mathcal{P}_D)$. Assume that the system is asymptotically stable. Then the cost $\mathcal{J}(\mathcal{P}_D)$ is bounded if the transfer function from $w$ to $z$ has $H_\infty$-norm less than $\gamma$. This holds if and only if there exists a positive semidefinite matrix $S$ which satisfies the algebraic Riccati equation

$$S = A' SA + (B'_1 SA + D' C')' (\gamma^2 I - D'_1 D_1 - B'_1 SB_1)^{-1} (B'_1 SA + D'_1 C) + C'C$$

such that

$$\gamma^2 I - D'_1 D_1 - B'_1 SB_1 > 0$$

and the matrix

$$A_S := A + B_1 (\gamma^2 I - D'_1 D_1 - B'_1 SB_1)^{-1} (B'_1 SA + D'_1 C)$$

is asymptotically stable.

Moreover, the cost is given by

$$\mathcal{J}(\mathcal{P}_D) = \text{tr}(B'_0 SB_0) + \text{tr}(D'_0 D_0)$$

and it is achieved by the signal $w^0 = \mathcal{T}v$, where $\mathcal{T}$ is the linear strictly causal transfer function defined by the equations

$$x_{k+1} = A_S x_k + B_0 v_k$$

$$w^0_k = (\gamma^2 I - D'_1 D_1 - B'_1 SB_1)^{-1} (B'_1 SA + D'_1 C) x_k.$$  

The calculation of the robust performance measures in (33) and (34) involves the minimization of the worst-case cost (23) with respect to the diagonal weighting filter $D \in \mathcal{D}_s$. The case when there is only a single unstructured block ($s = 1$) with a real scaling parameter $d \in R$ involves a simple one-dimensional search. The case when $d_i, d_i^{-1} \in H_\infty, i = 1, \ldots, s$ can be treated by parameterizing the weighting filters and applying parametric optimization techniques. Suppose that the filters $d_i$ have the state-space parameterizations

$$d_i(z) = C_i (zI - A_i)^{-1} B_i + D_i, \quad i = 1, \ldots, s.$$  

The diagonal weighting matrix $D(z) \in \mathcal{D}_s$ then has the state-space representation

$$D(z) = D_D(zI - A_D)^{-1} B_D + D_D$$

where

$$A_D := \text{diag}(A_1, \ldots, A_1, A_2, \ldots, A_s)$$

$$B_D := \text{diag}(B_1, \ldots, B_1, B_2, \ldots, B_s)$$

$$C_D := \text{diag}(C_1, \ldots, C_1, C_2, \ldots, C_s)$$

$$D_D := \text{diag}(D_1, \ldots, D_1, D_2, \ldots, D_s)$$
where the ith blocks are repeated \( r_i \) times. Assume that the plant \( \mathcal{P} \) has state-space representation

\[
x_{P,k+1} = A_P x_{P,k} + B_P v_k + B_{P1} w_{\Delta,k}
\]
\[
z_{0,k} = C_{P0} x_{P,k} + D_{P0} v_k + D_{P1} w_{\Delta,k}
\]
\[
z_{\Delta,k} = C_{P1} x_{P,k}.
\]

The augmented plant \( \mathcal{P}_D \) in (37) then has the system matrices

\[
A := \begin{bmatrix} A_D - B_D D_D^{-1} C_D & 0 & 0 \\ -B_{P1} D_D^{-1} C_D & A_P & 0 \\ 0 & B_D C_{P1} & A_D \end{bmatrix}
\]
\[
B_0 := \begin{bmatrix} 0 \\ B_{P0} \\ 0 \end{bmatrix}, \quad B_1 := \begin{bmatrix} B_D D_D^{-1} \\ B_{P1} D_D^{-1} \\ 0 \end{bmatrix}
\]
\[
C := \begin{bmatrix} C_0 \\ C_1 \end{bmatrix} := \begin{bmatrix} -D_{P1} D_D^{-1} C_D & C_{P0} & 0 \\ 0 & D_D C_{P1} & C_D \end{bmatrix}
\]
\[
D_0 := \begin{bmatrix} D_{P0} \\ 0 \end{bmatrix}, \quad D_1 := \begin{bmatrix} D_{P1} D_D^{-1} \\ 0 \end{bmatrix}.
\]

Gradient-based methods can be applied to the problem of minimizing \( \mathcal{L}(\mathcal{P}, D) \) in the set of filters \( d_i \) of fixed orders. The required gradients are given as follows.

**Theorem 3.4** Consider the system \( \mathcal{P}_D \) in equation (37) with the associated worst-case cost \( \mathcal{J}(\mathcal{P}_D) \) given by equation (42). Assume that the diagonal weighting filter \( D \) is parametrized according to (44)–(46), so that \( \mathcal{P}_D \) has the system matrices in (48). Suppose that the system is stable and that the \( H_\infty \)-norm bound of Theorem 3.3 holds. The gradients of \( \mathcal{J}(\mathcal{P}_D) \) with respect to the filter parameters are then given by

\[
\frac{\partial \mathcal{J}(\mathcal{P}_D)}{\partial A_D} = 2[I \ 0 \ 0] S(A + B_1 G) Q \begin{bmatrix} I \\ 0 \\ 0 \end{bmatrix} + 2[0 \ 0 \ I] S(A + B_1 G) Q \begin{bmatrix} 0 \\ 0 \\ I \end{bmatrix}
\]
\[
\frac{\partial \mathcal{J}(\mathcal{P}_D)}{\partial B_D} = 2[I \ 0 \ 0] S(A + B_1 G) Q \begin{bmatrix} -(D_D^{-1} C_D)' \\ 0 \\ 0 \end{bmatrix} + (D_D^{-1} G)'
\]
\[
+ 2[0 \ 0 \ I] S(A + B_1 G) Q \begin{bmatrix} 0 \\ C_{P1} \\ 0 \end{bmatrix}
\]
\[
\frac{\partial \mathcal{J}(\mathcal{P}_D)}{\partial C_D} = -2[B_1' S(A + B_1 G) + (D_{P1} D_D^{-1})' (D_{P1} D_D^{-1} G + C_0)] Q \begin{bmatrix} I \\ 0 \\ 0 \end{bmatrix} + 2C_1 Q \begin{bmatrix} 0 \\ 0 \\ I \end{bmatrix}
\]
\[
\frac{\partial \mathcal{J}(\mathcal{P}_D)}{\partial D_D} = 2[B_1' S(A + B_1 G) + (D_{P1} D_D^{-1})' (D_{P1} D_D^{-1} G + C_0)] Q \begin{bmatrix} I \\ 0 \\ 0 \end{bmatrix} (D_D^{-1} C_D)'
\]
\[-2[B'_1'S(A + B_1G) + (DP_1D^{-1}_D)'(DP_1D^{-1}_D G + C_0)]Q(D^{-1}_DG)' + 2C_1Q \begin{bmatrix} 0 \\ I \\ 0 \end{bmatrix} C'_{P_1}\]

where it is assumed that the matrices $S$ and $Q$ are partitioned in compliance with $A$ (equation (48)), $S$ is the matrix defined by equation (39),

\[G := (\gamma^2I - D'_1D_1 - B'_1SB_1)^{-1}(B'_1SA + D_1C_1), \quad (49)\]

and $Q$ is the symmetric positive definite solution of the matrix equation

\[Q = A_SQ A'_S + B_0B'_0. \quad (50)\]

where $A_S := A + B_1G$, cf. equation (41).

**Proof:** Equation (39) can be written in the form

\[S = (A + B_1K)'S(A + B_1K) - (K - M^{-1}D'_1C)'M(K - M^{-1}D'_1C) + C'(I + D_1M^{-1}D'_1)C, \]

\[M = \gamma^2I - D'_1D_1\]

with $K = G$. Then,

\[\frac{\partial J(P_D)}{\partial \Theta} = \left(\frac{\partial J(P_D)}{\partial \Theta}\right)_K + \sum_{k,l} \left(\frac{\partial J(P_D)}{\partial K_{kl}}\right)_0 \left(\frac{\partial K_{kl}}{\partial \Theta}\right) \quad (51)\]

where a subscript is used to indicate a variable which is considered constant in the differentiation. The gradient formulae then follow from the observation that

\[\left(\frac{\partial J(P_D)}{\partial K}\right)_0 = 0\]

at $K = G$, and by direct evaluation of the first term in (51) with $\Theta = A_D, B_D, C_D$ and $D_D$, respectively.

The calculation of an optimal weighting filter is considerably simplified if it can be ensured that the $H_\infty$-norm bound of Theorem 3.3 holds, i.e., the matrix (41) is asymptotically stable. The following result shows that this can be guaranteed under some fairly mild assumptions.

**Lemma 3.1** Consider the system $P_D$ and the associated cost $J(P_D)$. Suppose that $D \in D_1$ is parameterized according to (44)-(46). Assume that the system is stable and that the $H_\infty$ norm from $w$ to $z$ is less than $\gamma$. Define the scaled weighting matrix $D_c := cD, \ c > 0$. Then there exists $c_{\text{min}} < 1$ such that the $H_\infty$ norm of the transfer function from $w$ to $z$ of the system $P_cD$ is less than $\gamma$ if and only if $c > c_{\text{min}}$, and
is equal to $\gamma$ for $c = c_{\text{min}}$. When $c = c_{\text{min}}$, the matrix $A_S$ in (41) has one or more eigenvalues on the unit circle.

Suppose that the pair $(A_S, B_0)$ is (asymptotically) stabilizable. Then there exists $c_0 > c_{\text{min}}$ such that $J(\mathcal{P}_{c_0}) < J(\mathcal{P}_{c_{\text{min}}})$, and $J(\mathcal{P}_D)$ is strictly decreasing on $c \in [c_{\text{min}}, c_0]$.

Proof: Observe that $c_{\text{min}}$ is the number

$$c_{\text{min}} := \min \left\{ c : \left\| \begin{bmatrix} c^{-1}I & 0 \\ 0 & I \end{bmatrix} \mathcal{P}_D \begin{bmatrix} 0 \\ I \end{bmatrix} \right\|_\infty \leq \gamma \right\},$$

and the first part of the lemma follows from standard $H_\infty$ control theory. By a straightforward adaptation of Theorem 3.1 in Stoorvogel (1993) to the discrete-time case it follows from the assumption on $(A_S, B_0)$ that

$$\|D z_\Delta\|^2_2 - \|D w_\Delta\|^2_2 \downarrow -\infty$$

as $c \downarrow c_{\text{min}}$. From the definition of $J(\mathcal{P}_{c_0})$ it is seen that the above quantity equals $\partial J(\mathcal{P}_{c_0}) / \partial c$, and the latter part of the lemma follows.

The significance of the above result is that when the assumptions stated in the lemma hold, the calculation of the optimal weighting filter can essentially be treated as an unconstrained optimization problem as far as the stability boundary of $A_S$ is concerned; one can always move away from the stability boundary by a simple scaling of the weighting filter in such a way that the cost is reduced.

The assumption of the lemma can always be guaranteed by introducing a perturbed problem where the matrix $B_0$ is replaced by $B_{0,\epsilon}$ defined according to

$$B_{0,\epsilon} := B_0 + \epsilon I, \quad \epsilon > 0$$

It is straightforward to see that the cost $J(\mathcal{P}_D)$ is continuous with respect to $\epsilon$. Hence the cost can be evaluated to an arbitrary degree of accuracy subject to the condition of Lemma 3.1.

4 The mixed $H_2/H_\infty$ optimal control problem

In this section the mixed $H_2/H_\infty$ optimal control problem induced by the worst-case cost measure (38) introduced in Section 3 is studied. The main contribution will be to give a set of synthesis equations for the optimal controller which correspond to similar synthesis equations for the continuous-time case.

Consider the feedback system in Fig. 4.1. Assume that the system $\mathcal{G}$ has the state-space representation

$$
\begin{align*}
x_{k+1} &= A x_k + B_0 v_k + B_1 w_k + B_2 u_k \\
z_k &= C_1 x_k + D_{12} u_k \\
y_k &= C_2 x_k + D_{20} v_k + D_{21} w_k
\end{align*}
$$

14
The following assumptions are made on the system (55):

\begin{itemize}
    \item[A1] The pair \((A, B_2)\) is stabilizable.
    \item[A2] The pairs \((C_1, A)\) and \((C_2, A)\) are detectable.
    \item[A3] \(D_{12}D_{12} > 0\).
\end{itemize}

Let \(\mathcal{F}(\mathcal{G}, \mathcal{K})\) denote the transfer function of the closed-loop system in Fig. 4.1,

\[ z = \mathcal{F}(\mathcal{G}, \mathcal{K}) \begin{bmatrix} v \\ w \end{bmatrix}. \quad (56) \]

We consider the problem of finding a linear internally stabilizing controller \(\mathcal{K}\) such that the \(H_\infty\)-norm from \(w\) to \(z\) is less than \(\gamma\), and which minimizes the mixed \(H_2/H_\infty\) cost \(\mathcal{J}(\mathcal{F}(\mathcal{G}, \mathcal{K}))\) in (38) for the closed-loop system. For convenience, we introduce the notation

\[ \mathcal{J}(\mathcal{G}, \mathcal{K}) := \mathcal{J}(\mathcal{F}(\mathcal{G}, \mathcal{K})). \quad (57) \]

To summarize, the control problem can be formulated as follows.

**Mixed \(H_2/H_\infty\) problem:** Find a linear internally stabilizing controller \(u = \mathcal{K}y\) for the system \(\mathcal{G}\) in equation (55) such that the closed loop satisfies the \(H_\infty\)-norm bound

\[ \|\mathcal{F}_{zw}(\mathcal{G}, \mathcal{K})\|_\infty < \gamma \quad (58) \]

and which, if such a controller exists, minimizes the mixed \(H_2/H_\infty\) cost \(\mathcal{J}(\mathcal{G}, \mathcal{K})\).

The solution of the mixed \(H_2/H_\infty\) control problem will be presented in two phases. First, a parameterization of the optimal controller is derived. The result is analogous with a corresponding parameterization given by Kaminer et al. (1993) for the dual version of the problem studied here. In the second phase, synthesis equations for the optimal controller are derived.

### 4.1 Parameterization of optimal controller

The following theorems introduce the optimal mixed \(H_2/H_\infty\) full-information controller, and a variable transformation, which takes the problem into an equivalent mixed \(H_2/H_\infty\) optimal estimation problem. The results of the theorems can be derived in a fairly straightforward way by adapting corresponding discrete \(H_\infty\) control results (Stoorvogel, 1992) in combination with standard linear optimal control theory to the cost expression in Theorem 3.3 and the mixed \(H_2/H_\infty\) problem, cf. also the dual problem studied by Kaminer et al. (1993).

**Theorem 4.1** Consider the system (55), and assume that (A1)-(A3) hold. There exists a stabilizing full-information controller \(u = \mathcal{K}_1x + \mathcal{K}_2w\) such that the \(H_\infty\)-norm
from \( w \) to \( z \) of the closed loop is less than \( \gamma \) if and only if there exists a bounded positive semidefinite matrix \( X \) which satisfies the Riccati equation

\[
X = A'XA - \begin{bmatrix} A'XB_1 & A'XB_2 + C_1'D_{12} \end{bmatrix} M(X)^{-1} \begin{bmatrix} B_1'XA \\ B_2'XA + D_{12}'C_1 \end{bmatrix} + C_1'C_1
\]

(59)

where

\[
M(X) := \begin{bmatrix} B_1' \\ B_2' \\ \end{bmatrix} X \begin{bmatrix} B_1 & B_2 \end{bmatrix} + \begin{bmatrix} -\gamma^2 I & 0 \\ 0 & D_{12}'D_{12} \end{bmatrix}
\]

(60)

such that

\[
R := \gamma^2 I - B_1'(X - XB_2(B_2'XB_2 + D_{12}'D_{12})^{-1}B_2')B_1 > 0
\]

(61)

and the matrix

\[
A_X := A - \begin{bmatrix} B_1 & B_2 \end{bmatrix} M(X)^{-1} \begin{bmatrix} B_1'XA \\ B_2'XA + D_{12}'C_1 \end{bmatrix}
\]

(62)

is asymptotically stable. In this case, the stabilizing full-information which achieves the \( H_{\infty} \) norm bound and which minimizes the cost (57) is given by

\[
u^0_k = -K_1x_k - K_2w_k
\]

(63)

where

\[
K_1 := (B_2'XB_2 + D_{12}'D_{12})^{-1}(B_2'XA + D_{12}'C_1)
\]

(64)

\[
K_2 := (B_2'XB_2 + D_{12}'D_{12})^{-1}B_2'XB_1.
\]

(65)

The minimum value of the cost is

\[
\bar{J}_{fi}(\hat{G}) := \min_{K_1,K_2} J(\hat{G},K) = \text{tr}(B_0'XB_0),
\]

(66)

Introduce the variables

\[
\begin{align*}
\tilde{z}_k & := V^{1/2}(u_k + K_1x_k + K_2w_k) \\
\tilde{w}_k & := \gamma^{-1}R^{1/2}(w_k + F_wx_k)
\end{align*}
\]

(67)

where

\[
\begin{align*}
V & := B_2'XB_2 + D_{12}'D_{12} \\
F_w & := -R^{-1}B_1'(X - XB_2V^{-1}B_2')(A - B_2(D_{12}'D_{12})^{-1}D_{12}'C_1).
\end{align*}
\]

Define a transformed system \( \hat{G} \) by introducing the variable transformation (67) into the system (55),

\[
\begin{align*}
x_{k+1} & = \hat{A}x_k + B_0v_k + \hat{B}_1\tilde{w}_k + B_2u_k \\
\tilde{z}_k & = \hat{C}_1x_k + \hat{D}_{11}\tilde{w}_k + \hat{D}_{12}u_k \\
y_k & = \hat{C}_2x_k + D_{20}v_k + \hat{D}_{21}\tilde{w}_k
\end{align*}
\]

(68)
where
\[
\begin{align*}
\tilde{A} & := A - B_1 F_w \\
\tilde{B}_1 & := \gamma B_1 R^{-1/2} \\
\tilde{C}_1 & := V^{-1/2} (B_2 X \tilde{A} + D_1 C_1) \\
\tilde{D}_{11} & := \gamma V^{-1/2} B_2 X B_1 R^{-1/2} \\
\tilde{D}_{12} & := V^{1/2} \\
\tilde{C}_2 & := C_2 - D_21 F_w \\
\tilde{D}_{21} & := \gamma D_{21} R^{-1/2}
\end{align*}
\] (69)

From standard $H_\infty$ control theory it is well known that a controller $u = Ky$ achieves the $H_\infty$ norm bound for the system $G$ if and only if it achieves the bound for the system $\tilde{G}$, and that, for the latter system, the $H_\infty$ problem is equivalent to an $H_\infty$-optimal estimation problem. For the mixed $H_2/H_\infty$ problem and the cost in Theorem 3.3 a similar result is obtained.

**Theorem 4.2** Consider the systems $G$ and $\tilde{G}$ in equations (55) and (68), respectively, and the associated worst-case costs $\mathcal{J}(G, K)$ and $\mathcal{J}(\tilde{G}, K)$. For any linear causal controller $u = Ky$ the following statements are equivalent:

(a) $K$ stabilizes $G$ and the closed-loop transfer function from $w$ to $z$ has $H_\infty$-norm less than $\gamma$,

(b) the condition of Theorem 4.1 is satisfied, $K$ stabilizes $\tilde{G}$ and the closed-loop transfer function from $\tilde{w}$ to $\tilde{z}$ has $H_\infty$-norm less than $\gamma$.

When these conditions hold, the costs $\mathcal{J}(G, K)$ and $\mathcal{J}(\tilde{G}, K)$ are bounded, and

\[ \mathcal{J}(G, K) = \mathcal{J}(\tilde{G}, K) + J_{\text{fit}}(G). \] (70)

Moreover, assuming that the conditions of Theorem 4.1 hold, there exists a stabilizing controller for the system (68) which achieves the $H_\infty$-norm bound in (b) if and only if there exists a stable causal estimator $\hat{z} = \mathcal{F}_e y$ of the output $\tilde{z}$ of the system

\[
\begin{align*}
x_{k+1} &= \tilde{A} x_k + \tilde{B}_0 v_k + \tilde{B}_1 \tilde{w}_k \\
\tilde{z}_k &= \tilde{C}_1 x_k + \tilde{D}_{11} \tilde{w}_k \\
y_k &= \tilde{C}_2 x_k + D_{20} v_k + \tilde{D}_{21} \tilde{w}_k
\end{align*}
\] (71)

such that the $l_2$-induced norm from the disturbance $\tilde{w}$ to the estimation error $\tilde{z} - \hat{z}$ is less than $\gamma$. In this case that associated worst-case cost

\[
\mathcal{J}_e(\tilde{G}, \mathcal{F}_e) := \sup_{\mathcal{T}} \left\{ \sum_{j=1}^m \left[ \| \tilde{z} - \hat{z} \|^2_2 - \gamma^2 \| \tilde{w} \|^2_2 \right] | v = e_j \delta_0 : \tilde{w} = \mathcal{T} v, \mathcal{T} \text{ strictly causal} \right\}
\] (72)
is bounded, and
\[ \inf_{K} J(\hat{G}, K) = \inf_{F_c} J_e(\hat{G}, F_c), \]  
(73)
where the infimum with respect to \( K \) is taken over stabilizing controllers which achieve the \( H_\infty \)-norm bound in (b), and the infimum with respect to \( F_c \) is taken over stable causal estimators which achieve the norm bound stated above.

The following theorem, which a dual version of Theorem 4.1 in Kaminer et al. (1993), states that a static estimator achieves the optimal performance in Theorem 4.2.

**Theorem 4.3** There exists a stable causal estimator \( \hat{z} = F_c y \) for the system (71) such that the \( l_2 \)-induced norm from the disturbance \( \tilde{w} \) to the estimation error \( \hat{z} - \hat{z} \) is less than \( \gamma \) if and only if there exist real matrices \( L_1 \) and \( L_2 \) such that the matrix \( A + L_1 \hat{C}_2 \) is asymptotically stable, and a positive semidefinite matrix \( Y \) which satisfies

\[ Y = A Y A + (B'_1 Y A + D_{11}' C_1)' (\gamma^2 I - D_{11}' D_{11} - B'_1 Y B_1)^{-1} (B'_1 Y A + D_{11}' C_1) + C_1' C_1 \]  
(74)
where

\[ \bar{A} := \tilde{A} + L_1 \hat{C}_2 \]
\[ \bar{B}_1 := \tilde{B}_1 + L_1 D_{21} \]
\[ \bar{C}_1 := \tilde{C}_1 + L_2 \hat{C}_2 \]
\[ D_{11} := D_{11} + L_2 D_{21} \]

such that

\[ \gamma^2 I - D_{11}' D_{11} - B'_1 Y B_1 > 0 \]
(76)
and the matrix

\[ A_Y := \bar{A} + \bar{B}_1 (\gamma^2 I - D_{11}' D_{11} - B'_1 Y B_1)^{-1} (B'_1 Y \bar{A} + D_{11}' C_1) \]
(77)
is asymptotically stable.

When these conditions hold, the stable causal estimator \( \hat{z} = F_c (L_1, L_2) y \) defined by

\[ \hat{x}_{k+1} = \bar{A} \hat{x}_k - L_1 (y_k - \bar{C}_2 \hat{x}_k), \quad \hat{x}_0 = 0 \]
\[ \hat{z}_k = \bar{C}_1 \hat{x}_k - L_2 (y_k - \bar{C}_2 \hat{x}_k) \]  
(78)
achieves a bounded worst-case cost given by

\[ J_e(\hat{G}, F_c (L_1, L_2)) = \text{tr}[(B_0 + L_1 D_{20})' Y (B_0 + L_1 D_{20})] + \text{tr}(D_{20}' L_2' L_2 D_{20}). \]  
(79)

Moreover, the estimator structure (78) is optimal, i.e.,

\[ \inf_{L_1, L_2} J_e(\hat{G}, F_c (L_1, L_2)) = \inf_{F_c} J_e(\hat{G}, F_c), \]  
(80)
where the infima are taken over stable causal estimators \( F_c (L_1, L_2) \) and \( F_c \) which achieve the norm bound stated in Theorem 4.2.
Combining the above results gives the main result of this section.

**Theorem 4.4** Consider the system (55), and assume that (A1)–(A3) hold. The worst-case cost $J(G, K)$ is bounded for any causal stabilizing controller $u = Ky$ such that the closed-loop transfer function from $w$ to $z$ has $H_\infty$-norm less than $\gamma$. Such a controller exist if and only if the following conditions are satisfied:

(i) there exists a positive semidefinite matrix $X$ which satisfies equations (59)–(62),

(ii) there exist real matrices $L_1$ and $L_2$ and a positive semidefinite matrix $Y$ which satisfy the conditions of Theorem 4.3.

When these conditions hold, the controller $K(L_1, L_2)$ defined by

$$\begin{align*}
\dot{x}_{k+1} &= \tilde{A}x_k + B_2u_k - L_1(y_k - \tilde{C}_2\dot{x}_k), \quad \dot{x}_0 = 0 \\
\dot{u}_k &= -K_1\dot{x}_k + (B_2'X + D_{12}'D_{12})^{-1/2}L_2(y_k - \tilde{C}_2\dot{x}_k)
\end{align*}$$

(81)

stabilizes the system (55) and gives the bounded cost

$$J(G, K(L_1, L_2)) = J_e(G, F_e(L_1, L_2)) + J_f(G)$$

$$\quad = \text{tr}[(B_0 + L_1D_{20})'Y(B_0 + L_1D_{20})]$$

$$+ \text{tr}(D_{20}'L_2D_{20}) + \text{tr}(B_0'X B_0)$$

(82)

Moreover,

$$\inf_{L_1, L_2} J(G, K(L_1, L_2)) = \inf_K J(G, K),$$

(83)

where the infima are taken over stabilizing controllers $K$, which achieve the norm bound stated in Theorem 4.2.

### 4.2 Synthesis equations for optimal controller

Theorem 4.4 shows that the controller (81) performs as well as any controller for the mixed $H_2/H_\infty$ cost (57). Convex optimization methods can be used to find optimal values for the parameters $L_1$ and $L_2$ (Kaminer et al., 1993). An alternative approach is derive a set of synthesis equations which define the optimal controller.

**Theorem 4.5** For the controller $K(L_1, L_2)$ defined in Eq. (81), the gradient of the cost (82) is given by

$$\begin{align*}
\frac{\partial}{\partial L_1} J(G, K) &= 2YA'P(\tilde{C}_2 + \tilde{D}_{21}H_w)' + 2Y(B_0 + L_1D_{20})D_{20}' \\
\frac{\partial}{\partial L_2} J(G, K) &= 2L_2D_{20}D_{20}' + 2(C_1 + D_{11}H_w)P(\tilde{C}_2 + \tilde{D}_{21}H_w),'
\end{align*}$$

where

$$H_w := (\gamma^2 I - \tilde{D}_{11}'\tilde{D}_{11} - \tilde{B}_1'Y\tilde{B}_1)^{-1}(\tilde{B}_1'Y\tilde{A} + \tilde{D}_{11}'\tilde{C}_1)$$

(84)
and \( P \) is the symmetric positive semidefinite solution to

\[
P = A_Y P A_Y' + (B_0 + L_1 D_{20})(B_0 + L_1 D_{20})',
\]

and \( A_Y \) is defined in Eq. (77).

**Proof:** Equation (74) can be written in the form

\[
Y = (\bar{A} + \bar{B}_1 \bar{K})' Y (\bar{A} + \bar{B}_1 \bar{K}) - (\bar{K} - \bar{M}^{-1} \bar{Y}'_1 \bar{C}_1)' \bar{M} (\bar{K} - \bar{M}^{-1} \bar{Y}'_1 \bar{C}_1)
+ \bar{C}_1' (I + D_{11} \bar{M}^{-1} D_{11}') \bar{C}_1,
\]

\[
\bar{M} = \gamma^2 I - \bar{Y}'_1 \bar{D}_{11}
\]

with \( \bar{K} = H_w \). The gradients can be derived in a similar way as in Theorem 3.4, by applying the chain rule and the fact that \( \partial J(\bar{G},\bar{K})/\partial \bar{K} = 0 \) at \( \bar{K} = H_w \), followed by direct evaluation of the gradients involved and some lengthy but straightforward manipulations to yield the given expressions. \( \square \)

**Theorem 4.6** (a) Suppose that the infimum in Theorem 4.4 is achieved for some \( L_1^*, L_2^* \) in the interior of the region where the matrix \( A_Y \) in (77) is asymptotically stable. Then the optimality conditions

\[
\frac{\partial}{\partial L_1} J(\bar{G},\bar{K}) = 0, \quad \frac{\partial}{\partial L_2} J(\bar{G},\bar{K}) = 0
\]

are satisfied for \( L_1^*, L_2^* \).

(b) Suppose that there exist \( L_1^*, L_2^* \) in the interior of the region where the matrix \( A_Y \) in (77) is asymptotically stable, and such that

\[
Y > 0
\]

\[
A_Y P (\bar{C}_2 + \bar{D}_{21} H_w)' + (B_0 + L_1 D_{20}) D_{20}' = 0
\]

\[
L_2 D_{20} D_{20}' + (\bar{C}_1 + \bar{D}_{11} H_w) P (\bar{C}_2 + \bar{D}_{21} H_w)' = 0.
\]

hold at \( L_1 = L_1^*, L_2 = L_2^* \). Then

\[
J(\bar{G},\bar{K}(L_1^*, L_2^*)) = \inf_{\bar{K}} J(\bar{G},\bar{K})
\]

**Proof:** Part (a) follows from the fact that the cost is continuously differentiable in the interior of the stability region. The proof of part (b) is given in Appendix B. \( \square \)

The assumption that the minimum is achieved in the interior of the stability region cannot be ensured in general. Note, however, that in the context of the robust \( H_2 \) problem, it can under some mild assumptions be guaranteed that the assumptions in Theorem 4.6 do hold, cf. Section 5.
The optimality conditions in Theorem 4.6 are the discrete analogs of the continuous-time synthesis equations given in \cite{Bernstein1989, Doyle1989, Yeh1992}. Observe that the synthesis equations (86) or (87) consist of a set of coupled nonlinear matrix equations, whose direct solution is numerically quite demanding. Cf. the continuous-time case, where homotopy methods have been proposed for solving the synthesis equations. However, a numerically much more powerful approach is obtained by solving the equations indirectly via the minimization of the cost $J(\mathcal{G}, \mathcal{K}(L_1, L_2))$ using a convergent gradient-based minimization procedure. There are well-established numerically powerful gradient-based algorithms which converge to a stationary point such that (86) hold, if such a point exists.

\section{Solution of robust $H_2$ performance problem}

The robust performance measure introduced in Section 3 and the mixed $H_2/H_\infty$ control problem in Section 4 can be combined to solve a robust $H_2$ performance problem for the control system in Fig. 5.1. Introducing the robust performance measures in (34) and (33), the problem then consists of finding a diagonal weighting filter $D \in D_s$ and a stabilizing controller such that a mixed $H_2/H_\infty$ cost of the form (30) is minimized for the closed-loop system.

Robust $H_2$ performance problem: Find an internally stabilizing controller $\mathcal{K}$ and a diagonal weighting matrix $D \in D_s$ for the system $\mathcal{G}_0$ such that the closed-loop system from $Dw_\Delta$ to $[z_0', (Dz_\Delta)']$ has $H_\infty$-norm less than $\sigma^{-1}$,

$$\left\| \begin{bmatrix} I & 0 \\ 0 & D \end{bmatrix} \mathcal{F}(\mathcal{G}_0, \mathcal{K}) \begin{bmatrix} 0 \\ D^{-1} \end{bmatrix} \right\|_\infty < \sigma^{-1}$$

and which achieves the robust $H_2$ performance bound

$$L(\mathcal{F}(\mathcal{G}_0, \mathcal{K}), D) < \alpha,$$

where the mixed $H_2/H_\infty$ cost $L(\mathcal{F}(\mathcal{G}_0, \mathcal{K}), D)$ is defined in accordance with equation (30) for the closed-loop transfer function, and $\alpha$ denotes a specified performance level.

Recall from Section 3 that it was assumed that the transfer function from $v$ to $w_\Delta$ lacks a direct feedthrough term. This assumption can now be motivated as follows. Observe that for causal systems, the loop transfer function $\mathcal{G}_{0, yv}\mathcal{K}$ in Fig. 5.1 is strictly causal. For a large class of standard uncertainty descriptions, such as additive and multiplicative uncertainties, this implies that either the uncertainty $\Delta$ or the closed-loop transfer function $\mathcal{F}_{\Delta, v}(\mathcal{G}_0, \mathcal{K})$ has a delay, see Section 6 for an example. This feature was exploited in the definition of the mixed $H_2/H_\infty$ cost (57) so as to make it less conservative as an upper bound on the robust performance measures (19) and (20).
Theorem 4.4 provides a parameterization of the optimal controller in terms of the filter gain matrices $L_1$ and $L_2$. In the present context, this parameterization refers to the augmented plant

$$
\mathcal{G} := \begin{bmatrix} I & 0 & 0 \\
0 & D & 0 \\
0 & 0 & I \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\
0 & D & 0 \\
0 & 0 & I \end{bmatrix}
$$

(91)

The optimal controller parameterization in (81) then depends in a complex way on the weighting filter $D$ via the Riccati equation (59). It is therefore quite hard to minimize the mixed $H_2/H_\infty$ cost $\mathcal{L}(\mathcal{F}(\mathcal{G}_0, \mathcal{K}), D)$ simultaneously with respect to both $D$ and the controller $\mathcal{K}$, using the optimal controller parameterization in Theorem 4.4. One approach is then to apply a two-stage procedure, in which the cost is decreased alternately with respect to the controller $\mathcal{K}$ and the weighting filter $D$.

Algorithm (Robust $H_2$ performance problem):

0. Select a initial diagonal weighting matrix $D_0 \in \mathcal{D}_s$ of specified order, and a stabilizing controller $\mathcal{K}_0$ such that the norm bound (89) is satisfied. Set $k = 0$.

1. Apply the parameterization in Theorem 4.4 and the gradients in Theorem 4.5 to find a controller $\mathcal{K}_{k+1}$ such that $\mathcal{L}(\mathcal{F}(\mathcal{G}_0, \mathcal{K}_{k+1}), D_k) \leq \mathcal{L}(\mathcal{F}(\mathcal{G}_0, \mathcal{K}_k), D_k)$.

2. Apply the gradient expression in Theorem 3.4 to the closed-loop plant $\mathcal{F}(\mathcal{G}, \mathcal{K})$ to find a weighting filter $D_{k+1} \in \mathcal{D}_s$ such that $\mathcal{L}(\mathcal{F}(\mathcal{G}_0, \mathcal{K}_{k+1}), D_{k+1}) \leq \mathcal{L}(\mathcal{F}(\mathcal{G}_0, \mathcal{K}_k), D_k)$.

3. Test for convergence, and continue from step 1 until convergence is achieved.

Observe that for numerical efficiency, the minimization steps 1 and 2 should not be solved to convergence at each iteration. Instead, it is more efficient to solve the optimization steps only approximately to provide a decrease of the cost at each iteration.

In numerical computations it is important to ensure that the $H_\infty$ norm bound (89) holds, or equivalently, that the stability bounds in Theorems 4.1 and 4.3 are satisfied. Lemma 3.1 shows that under some mild conditions which are stated in the lemma the minimization with respect to the weighting filter in stage 2 of the algorithm can always be performed in such a way that these stability bounds are satisfied.

6  Numerical example

In this section a simple numerical example is presented to illustrate the proposed synthesis procedure. Consider the system in Fig. 6.1 with additive uncertainty where the nominal plant $\mathcal{G}_{\text{nom}}$ is described by

$$
x_{k+1} = 0.8x_k + [1 \ 0]v_k + u_k \\
z_{0,k} = x_k \\
y_k = x_k + [0 \ 1]v_k.
$$

(92)
The uncertain system in Fig. 6.1 can be characterized by the connection in Fig. 5.1, where \( G_0 \) is described by the equations

\[
\begin{align*}
x_{k+1} &= 0.8x_k + [1 \ 0]v_k + u_k \\
z_{0,k} &= x_k \\
z_{\Delta,k} &= u_k \\
y_k &= x_k + [0 \ 1]v_k + w_{\Delta,k}.
\end{align*}
\]

Observe that it is assumed that the causality of the uncertain system (92) is such that \( y_k \) is a function of \( u_l, l < k \), but not of \( u_k \). It follows that the uncertainty \( \Delta \) does not have a direct coupling from input to output either. On the other hand, if a direct coupling from \( u \) to \( y \) had been assumed, then causality considerations would require the controller \( \mathcal{K} \) to contain a delay, and in this case, the closed-loop transfer function \( \mathcal{F}(G_0, \mathcal{K}) \) would lack a direct coupling from \( v \) to \( z_{\Delta} \). In both cases, the strict causality assumption introduced in Section 3 holds.

Optimal robust controllers which minimize the robust performance measure in Eq. (34) for nonlinear, time-varying uncertainties were computed for various values of the uncertainty radius \( \sigma \). The problem involves the minimization of the mixed \( H_2/H_\infty \) cost \( \mathcal{L}(\mathcal{F}(G_0, \mathcal{K}), D) \) in equation (30) for the closed loop with respect to both the controller \( \mathcal{K} \) and the weighting parameter \( D \), which in this example is a real scalar, \( D = d \). The problem can be solved using the two-stage algorithm in Section 5. Note that the minimization with respect to the scalar weight \( d \) reduces to a simple one-dimensional search in this example.

Figure 6.2 shows the robust performance measure and the nominal performance of the optimal robust controllers and the \( H_2 \)-optimal control law. Observe that the difference between the mixed \( H_2/H_\infty \) cost and the nominal \( H_2 \) cost gives an upper bound on the conservativeness of the upper bound in Theorem 3.1. The gains \( L_1 \) and \( L_2 \) of the optimal robust controller are shown in Fig. 6.3, and the optimized weight parameter \( d \) corresponding to the optimal controller is shown in Fig. 6.4.

### 7 Conclusion

In this paper a robust \( H_2 \) performance problem for discrete-time systems has been studied. The \( H_2 \) norm has here been used to characterize an average performance in a deterministic setting, in terms of the average quadratic cost with respect to \( l_2 \)-disturbances, averaged over time.

For robust performance, the worst-case \( H_2 \) cost for the uncertain plant should be small. A mixed \( H_2/H_\infty \) cost has been introduced as an upper bound on the worst-case performance with respect to norm-bounded plant uncertainties. The robust performance measure introduced in this way involves a Lagrangian multiplier weighting filter, which should be optimized so as to reduce the conservativeness of the performance
measure. For the calculation of the optimal controller and optimal weights a two-stage procedure has been proposed.

It is interesting to note that in structured singular value synthesis, where the design consists of finding a controller $K$ and a scaling matrix $D$ such that the closed-loop system achieves an $H_\infty$ norm bound, two-stage procedures, such as the 'DK-iteration', are also widely applied. The major difference stems from the fact the robust $H_2$ problem involves a mixed $H_2/H_\infty$ cost, and the synthesis procedure will include optimization steps to minimize this cost.

The mixed $H_2/H_\infty$ control problem studied in this paper is a dual version of a problem which has been studied previously in the literature. In this paper a set of synthesis equations for the optimal controller has been given. These are the discrete-time versions of similar synthesis equations for the continuous-time problem obtained previously.

Some major open problems associated with the control problem studied in this paper are for example concerned with the development of efficient special purpose numerical procedures for the various phases of the design process, and the problem of selecting proper filter orders for the Lagrangian multiplier weighting filters in the mixed $H_2/H_\infty$ cost function. These topics are left for future research.

Acknowledgement

This work was in part supported by the Academy of Finland.

8 Appendix A

Proof of Theorem 3.1: The transfer function from $dw_\Delta$ to $[z_0', (dz_\Delta)]'$ has $H_\infty$-norm less than $\sigma^{-1}$ if and only if

$$\left\| \begin{bmatrix} d^{-1}P_{z_0 w_\Delta} \\ \mathcal{P}_{z_0 w_\Delta} \end{bmatrix} \right\|_\infty < \sigma^{-1}. \quad (94)$$

Thus, (94) implies $\| \mathcal{P}_{z_0 w_\Delta} \|_\infty < \sigma^{-1}$, and conversely, if $\| \mathcal{P}_{z_0 w_\Delta} \|_\infty < \sigma^{-1}$ holds, the norm bound in (94) can be achieved by selecting $d$ equal to a sufficiently large constant.

The norm bound (94) implies that there exists $\delta > 0$ such that

$$\left\| \begin{bmatrix} d^{-1}P_{z_0 w_\Delta} \\ \mathcal{P}_{z_0 w_\Delta} \end{bmatrix} \right\|_\infty < (1-\delta)\sigma^{-1}. \quad (95)$$

The boundedness of $\mathcal{L}(\mathcal{P}, d)$ follows from the fact that for any $v \in l_2$,

$$\left\| \begin{bmatrix} z_0 \\ dz_\Delta \end{bmatrix} \right\|_2^2 = \sigma^{-2} \| dw_\Delta \|_2^2 = \left\| \begin{bmatrix} \mathcal{P}_{z_0 v} + \mathcal{P}_{z_0 w_\Delta} w_\Delta \\ d\mathcal{P}_{z_0 w_\Delta} v + d\mathcal{P}_{z_0 w_\Delta} w_\Delta \end{bmatrix} \right\|_2^2 - \sigma^{-2} \| dw_\Delta \|_2^2 \leq \left( \left\| \begin{bmatrix} \mathcal{P}_{z_0 v} \\ d\mathcal{P}_{z_0 w_\Delta} v \end{bmatrix} \right\|_2^2 + \left\| d^{-1} \mathcal{P}_{z_0 w_\Delta} \right\|_\infty \| dw_\Delta \|_2 \right)^2 - \sigma^{-2} \| dw_\Delta \|_2^2,$$
\[
\frac{1}{1 - (1 - \delta)^2} \left\| \begin{bmatrix} \mathcal{P}_{\Delta v} \\
 d\mathcal{P}_{\Delta v} 
\end{bmatrix} \right\|_2^2 < \infty.
\]

In order to show (24) and (25), consider the uncertain system in Fig. 3.1. Suppose that \( \Delta \) satisfies the norm bound \( \| \Delta \| < \sigma \). Then \( w_\Delta = \mathcal{T} v \), where \( \mathcal{T} = (I - \Delta \mathcal{P}_{\Delta w_\Delta})^{-1} \Delta \mathcal{P}_{\Delta v} \) is strictly causal. The inequalities (24) and (25) then follow from the definition of \( \mathcal{L}(\mathcal{P}, d) \) and the fact that \( \Delta \in \mathcal{B}_{NLTV}^\gamma \) implies (21) and \( \Delta \in \mathcal{B}_{LT}^\gamma \) implies (22).

\[ \square \]

9 Appendix B

The proof of Theorem 4.6 will be based on a convex characterization of the minimization problem which is involved. The convex problem representation which is used here is a dual version of the characterization of Kaminer et al. (1993), and it is summarized in the following lemmas.

Lemma B.1 Consider the system matrices defined in (75), and suppose that the matrix \( \bar{A} \) has all eigenvalues in the open unit disc. Then the following statements are equivalent:

(a) There exists a symmetric positive semidefinite matrix \( Y \) which satisfies equations (74) and (76) such that the matrix \( AY \) in (77) is asymptotically stable.

(b) There exists a symmetric \( Z > 0 \) such that

\[
R(Z) := \bar{A}'Z\bar{A} - Z + (\bar{B}'_1 Z \bar{A} + \bar{D}'_{11} \bar{C}_1)' M(Z)^{-1}(\bar{B}'_1 Z \bar{A} + \bar{D}'_{11} \bar{C}_1) + \bar{C}'_1 \bar{C}_1 < 0
\]

\[
M(Z) := \gamma^2 I - \bar{D}'_{11} \bar{D}_{11} - \bar{B}'_1 Z \bar{B}_1 > 0
\]

Moreover, the conditions in (b) are equivalent to

\[
Q(Z, L_1, L_2) := \begin{bmatrix} \bar{A} \\
\bar{B}'_1 
\end{bmatrix} Z \begin{bmatrix} \bar{A} & \bar{B}_1 \end{bmatrix} + \begin{bmatrix} \bar{C}'_1 \\
\bar{D}'_{11} 
\end{bmatrix} \begin{bmatrix} \bar{C}_1 & \bar{D}_{11} \end{bmatrix} - \begin{bmatrix} Z & 0 \\
0 & \gamma^2 I \end{bmatrix} < 0
\]

For convenience, introduce the notation

\[
g(L_1, L_2) := \text{tr}[(B_0 + L_1 D_{20})'Y(B_0 + L_1 D_{20}) + D'_{20} L_2' L_2 D_{20}] \] (96)

for \( J_e(\bar{G}, \mathcal{F}_e(L_1, L_2)) \), and define the function

\[
f(Z, L_1, L_2) := \text{tr}[(B_0 + L_1 D_{20})'Z(B_0 + L_1 D_{20}) + D'_{20} L_2' L_2 D_{20}] \] (97)

Lemma B.2 Assume that \( L_1, L_2 \) satisfy the conditions in Theorem 4.3. Then the cost \( g(L_1, L_2) \) can be characterized as

\[
g(L_1, L_2) = \inf \left\{ f(Z, L_1, L_2) : Z = Z > 0, \ R(Z) < 0, \ \text{and} \ M(Z) > 0 \right\}. \tag{98}
\]
or, equivalently,
\[ g(L_1, L_2) = \inf \left\{ f(Z, L_1, L_2) : Z = Z' > 0, \ Q(Z, L_1, L_2) < 0 \right\} \tag{99} \]

Introduce the variable \( W := ZL_1 \), and define \( f_W(Z, W, L_2) := f(Z, Z^{-1}W, L_2) \),
\[ f_W(Z, W, L_2) := \text{tr}[(ZB_0 + WD_{20})^{-1}(ZB_0 + WD_{20}) + D_{20}L_2^{-1}L_2D_{20}] \tag{100} \]
and \( Q_W(Z, W, L_2) := Q(Z, Z^{-1}W, L_2) \),
\[ Q_W(Z, W, L_2) := \left[ \begin{array}{c} (Z\tilde{A} + W\tilde{C}_2)' \\ (Z\tilde{B}_1 + WD_{21})' \end{array} \right] Z^{-1} \left[ \begin{array}{cc} Z\tilde{A} + W\tilde{C}_2 & Z\tilde{B}_1 + WD_{21} \end{array} \right] \\
+ \left[ \begin{array}{c} \tilde{C}_1' \\ \tilde{D}_{11} \end{array} \right] [\tilde{C}_1 \tilde{D}_{11}] - \left[ \begin{array}{cc} Z & 0 \\ 0 & \gamma^2I \end{array} \right], \]
where the definitions of \( \tilde{A} \) and \( \tilde{B}_1 \) from equation (75) have been introduced. Introduce the sets
\[ \Omega := \left\{ (Z, W, L_2) \in R^{m \times n} \times R^{n \times p} \times R^{m \times p} : Z = Z' > 0 \right\} \tag{101} \]
and
\[ \Phi := \left\{ (Z, W, L_2) \in \Omega : Q_W(Z, W, L_2) < 0 \right\}. \tag{102} \]

Then the minimum cost can be characterized as follows.

**Lemma B.3** Consider the cost in (96). Then
\[ \inf_{L_1, L_2} g(L_1, L_2) = \inf_{(Z, W, L_2) \in \Phi} f_W(Z, W, L_2) \tag{103} \]
where the infimum with respect to \( L_1, L_2 \) is taken over matrices which satisfy the conditions in Theorem 4.3.

The significance of the above result is that it expresses the minimization problem as a convex optimization problem.

**Lemma B.4** The mapping \( f_W : \Omega \to R^+ \) is a convex function on \( \Omega \), and the mapping \( Q_W : \Omega \to R^{m \times n} \) is convex.

**Proof of Theorem 4.6 (b):** The proof is based on the convex characterization in Lemma B.3 and Lemma B.4. Consider the equations (87), and define \( Z^* := Y > 0 \) and \( W^* := YL_1^* \). Consider any \((\tilde{Z}, \tilde{W}, \tilde{L}_2) \in \Phi \), and define
\[ Z(\alpha) := Z^* + \alpha(\tilde{Z} - Z^*) \]
\[ W(\alpha) := W^* + \alpha(\tilde{W} - W^*) \]
\[ L_2(\alpha) := L_2^* + \alpha(\tilde{L}_2 - L_2^*) \tag{104} \]
By convexity (Lemma B.4), \((Z(\alpha), W(\alpha), L_2(\alpha)) \in \Phi \) on \( \alpha \in (0, 1] \), and
\[ f_W(Z(\alpha), W(\alpha), L_2(\alpha)) \leq f_W(Z^*, W^*, L_2^*) + \alpha[f_W(\tilde{Z}, \tilde{W}, \tilde{L}_2) - f_W(Z^*, W^*, L_2^*)]. \tag{105} \]
Note that \((Z^*, W^*, L^*_2) \notin \Phi\), but by continuity, (105) can be extended to \(\alpha = 0\). On \(\alpha \in [0, 1]\), define \(L_1(\alpha) := Z^{-1}(\alpha)W(\alpha)\). Invertibility of \(Z(\alpha)\) follows from the positive definiteness of \(Z^*\) and \(Z\). By construction, the inequalities in Lemma B.1 (b) hold on \(\alpha \in (0, 1]\). The inequality \(R(Z(\alpha)) < 0\) implies that the matrix \(\tilde{A}\) is asymptotically stable. It follows that the pair \((L_1(\alpha), L_2(\alpha))\) satisfies the conditions of Theorem 4.3 on \(\alpha \in [0, 1]\). Let \(Y(\alpha)\) denote the solution of (74) for \(L_1 = L_1(\alpha), L_2 = L_2(\alpha)\). Then, \(R(Z(\alpha)) < 0\) and \(R(Y(\alpha)) = 0\), and from standard monotonicity properties of the Riccati equation it follows that \(Z(\alpha) \geq Y(\alpha)\). Hence,

\[
g(L_1(\alpha), L_2(\alpha)) \leq f(Z(\alpha), L_1(\alpha), L_2(\alpha)) = f_W(Z(\alpha), W(\alpha), L_2(\alpha)),
\]

and, introducing \(f_W(Z^*, W^*, L^*_2) = g(L^*_1, L^*_2)\), which follows from the definitions of \(Z^*\) and \(W^*\), it follows from (105) that

\[
g(L_1(\alpha), L_2(\alpha)) \leq g(L^*_1, L^*_2) + \alpha[f_W(\bar{Z}, \bar{W}, \bar{L}_2) - g(L^*_1, L^*_2)].
\]

The inequality (107) implies

\[
\frac{d}{d\alpha} g(L_1(\alpha), L_2(\alpha)) \bigg|_{\alpha = 0} \leq f_W(\bar{Z}, \bar{W}, \bar{L}_2) - g(L^*_1, L^*_2)
\]

provided the derivative exists. On the other hand, we have

\[
\frac{d}{d\alpha} g(L_1(\alpha), L_2(\alpha)) = \text{tr} \left[ \left( \frac{\partial g(L_1(\alpha), L_2(\alpha))}{\partial L_1} \right)' \left( \frac{\partial L_1(\alpha)}{\partial \alpha} \right) \right] + \text{tr} \left[ \left( \frac{\partial g(L_1(\alpha), L_2(\alpha))}{\partial L_2} \right)' \left( \frac{\partial L_2(\alpha)}{\partial \alpha} \right) \right].
\]

Here, equations (87) imply that \(\partial g/\partial L_1 = 0\) and \(\partial g/\partial L_2 = 0\) at \(L^*_1, L^*_2\), and as

\[
\left( \frac{\partial L_1(\alpha)}{\partial \alpha} \right) \bigg|_{\alpha = 0} = (Z^*)^{-1}(\bar{W} - \bar{Z}L^*_1)
\]

\[
\left( \frac{\partial L_2(\alpha)}{\partial \alpha} \right) \bigg|_{\alpha = 0} = \bar{L}_2 - L^*_2
\]

are bounded, it follows that

\[
\frac{d}{d\alpha} g(L_1(\alpha), L_2(\alpha)) \bigg|_{\alpha = 0} = 0.
\]

From (108) and (112) it then follows that

\[
f_W(\bar{Z}, \bar{W}, \bar{L}_2) - g(L^*_1, L^*_2) \geq 0.
\]

As \((\bar{Z}, \bar{W}, \bar{L}_2) \in \Phi\) was arbitrary, it follows in view of Lemma B.3 that \(L^*_1, L^*_2\) is a minimizer, which was the result to be proven. \(\square\)
REFERENCES


10 Figures

**Figure 3.1.** Uncertain plant.

**Figure 3.2.** System for robust performance measure.

**Figure 4.1.** Control system for mixed $H_2/H_\infty$ problem.

**Figure 5.1.** Control of uncertain plant.
Figure 6.1. System with additive uncertainty.

Figure 6.2. Control performance as a function of uncertainty radius $\sigma$. Worst-case performance measure of optimal robust controller (solid line), nominal $H_2$ cost of optimal robust controller (dotted line), and worst-case performance measure of $H_2$-optimal controller (dashed line).
Figure 6.3. Gains of optimal robust controller versus uncertainty radius $\sigma$ (solid line: $L_1$, dashed line: $L_2$).

Figure 6.4. Optimized weight parameter $d$ of optimal robust controller versus uncertainty radius $\sigma$. 