MIXED H^2/H^∞ ESTIMATION: PRELIMINARY ANALYTIC CHARACTERIZATION AND A NUMERICAL SOLUTION

B. Halder, B. Hassibi and T. Kailath [1]

Information Systems Lab.
Stanford University
Stanford, CA 94305

Abstract. We introduce and motivate the problem of mixed H^2/H^∞ estimation by studying the stochastic and deterministic approaches of H^2 and H^∞ estimation. Mixed H^2/H^∞ estimators have the property that they have the best average performance over all estimators that achieve a certain worst-case performance bound. They thus allow a tradeoff between average and worst-case performances. In the finite horizon case, we obtain a numerical solution (based on convex optimization methods) for the *optimal* mixed H^2/H^∞ estimator. We also give some analytic characterizations, both on this optimal solution, and on the set of all estimators achieving a guaranteed worst-case bound. A numerical example is also provided.

Keywords. Robust estimation, H^2/H^∞ theory, Convex optimization.

1. INTRODUCTION

Classical methods in estimation theory (such as least-squares, maximum-likelihood, and maximum entropy) and the more recent robust methods in estimation theory (such as H^∞) can be regarded as two extremes in terms of their requirements regarding the statistical properties of the exogenous signals, as well as in terms of their goals. In classical estimation methods optimality of the *average* (or expected) performance of the estimator under some assumptions regarding the statistical nature of the signals is the key issue and hence their performance heavily depends upon the validity of these assumptions. On the other hand, robust estimation methods, or so-called *minimax* estimation strategies, safeguard against the *worst-case* disturbances and therefore make no assumptions on the (statistical) nature of the signals.

Among the classical methods, the widespread use of the linear least-squares (or H^2) estimation technique is mainly due to the following facts: (i) the optimal H^2 problem has a simple closed-form solution, and (ii) the solution has various desirable optimality properties under certain statistical assumption on the signals. However, in many applications, due to the model uncertainties and lack of statistical information H^2 methods are not directly applicable and the behavior of such estimation schemes is uncertain. Recently, following some pioneering work in robust control theory (Zames, 1981), H^∞ estimation theory has been developed to address such problems.

The mixed estimation problem was introduced as a compromise between these two extreme point of views (Khargonekar and Rotea, 1992; Zhou *et al.*, 1994; Bernstein and Haddad, 1989; Yeh *et al.*, 1992; Sznaier, 1994). The mixed H^2/H^∞ problem allows one to trade off between the best average performance of the H^2 estimator and the best guaranteed worst-case performance of the H^∞ estimator. As a result, the optimal mixed H^2/H^∞

[1] This research was supported by the Advanced Research Projects Agency of the Department of Defense monitored by the Air Force Office of Scientific Research under Contract F49620-93-1-0085.

estimators achieve the best average performance, not over the set of all estimators, but over a restricted set of estimators that achieve a certain worst case performance bound. We note that the suboptimal (and even optimal) H^∞ estimators are highly non-unique and the mixed H^2/H^∞ approach attempts to exploit this non-uniqueness by choosing the estimator that has the best average performance. Unlike optimal H^2 and suboptimal H^∞ problems, the question of finding the optimal mixed estimator is still open. In this paper we present partial analytic characterizations of the optimal solution and in the finite horizon, case present an efficient algorithm for numerical solution.

2. THE DATA MODEL

A general framework for estimation problems is shown in Fig. 1 which is capable of incorporating almost all estimation problems, such as Wiener, Kalman and adaptive filtering. The *causal* linear transfer operators $\mathcal{H}$ and $\mathcal{L}$ are assumed to be known. In the finite horizon case $\mathcal{H}$ and $\mathcal{L}$ can be represented by finite *lower triangular* matrices, and in the infinite horizon case they are infinite (or semi-infinite) lower triangular matrices. In the infinite horizon case when $\mathcal{H}$ and $\mathcal{L}$ are time-invariant transfer operators they can be represent by transfer functions $H(z)$ and $L(z)$, respectively. The model considered below is general and applies to all of the above cases. In what follows we shall denote sequences such as $\{u_j\}$ by u, and simply write $z = \mathcal{L}u$, to denote that $\mathcal{L}$ maps the input sequence $\{u_j\}$ to the output sequence $\{z_j\}$.

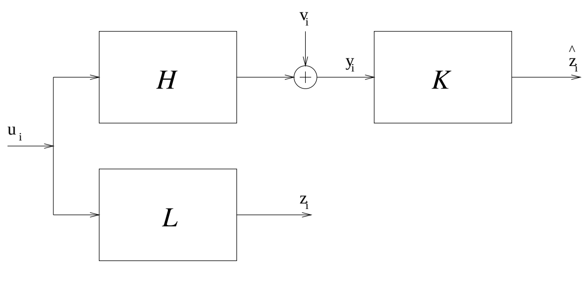

Fig. 1. *A general estimation problem.*

The sequences $\{u_j\}$ and $\{v_j\}$ are assumed to be *unknown*. [$\{u_j\}$ may be considered as a driving disturbance and $\{v_j\}$ as a measurement disturbance. In general, both may include modeling errors resulting from lack of knowledge of the "true" $\mathcal{H}$ and $\mathcal{L}$.] The goal is to design a causal transfer operator (or filter) $\mathcal{K}$ that estimates z_i, the output of $\mathcal{L}$, using the observations $\{y_j, j \le i\}$. The estimates are denoted by $\hat{z}_i$ and the estimation errors by $\tilde{z}_i \triangleq z_i - \hat{z}_i$. From Fig. 1, we can write the induced transfer operator that maps the disturbances $\{u_j\}$ and $\{v_j\}$ to the estimation errors $\{\tilde{z}_j\}$ as

$$\mathcal{T}_K = \left[\, \mathcal{L} - \mathcal{K}\mathcal{H} \quad -\mathcal{K} \,\right]. \tag{1}$$

3. CLASSICAL H^2 METHODS

The problem of estimation is to select $\mathcal{K}$, and thereby the estimates $\hat{z}_i$, based on some performance criterion. The most widely used of such criteria is the H^2 norm of the transfer operator, i.e., $\|\mathcal{T}_K\|_2$. In the finite horizon case, $\|\mathcal{T}_K\|_2$ is simply the Frobenius norm of the matrix $\mathcal{T}_K$, $\|\mathcal{T}_K\|_2 = (\text{trace } [\mathcal{T}_K\mathcal{T}_K^*])^{1/2}$.

The solution to the H^2 estimation problem is well known and depending on the nature of the transfer operators $\mathcal{H}$ and $\mathcal{L}$, the H^2 optimal solution takes on various forms. When $\mathcal{H}$ and $\mathcal{L}$ have state-space structure the optimal solution yields the Kalman filter. When they have transfer function representations $H(z)$ and $L(z)$, the solution is the Wiener filter, and in the adaptive filtering case it corresponds to the recursive-least-squares (RLS) algorithm. As mentioned earlier, H^2 estimators have the following properties: (i) If the $\{u_j\}$ and $\{v_j\}$ are assumed to be zero-mean, uncorrelated and temporally white random variables, then they minimize the expected estimation error energy, and (ii) If, in addition to the assumptions of part (i), the $\{u_j\}$ and $\{v_j\}$ are assumed to be jointly Gaussian, then they yield the maximum-likelihood estimate of the $\{z_i\}$.

4. THE H^∞ APPROACH

Since, in practice we may not always know the statistics of the disturbances we cannot always guarantee the validity of the assumptions required of H^2 estimators. Therefore, the question that begs itself is: *is it possible that* **small** *disturbances and modeling errors may lead to* **large** *estimation errors?*

Intuitively, a non-robust algorithm would be one for which the above is true, and a robust algorithm would be one for which small disturbances lead to small estimation errors.

The problem of robust estimation is thus an important one and the H^∞ estimation formulation is an *attempt* at addressing this question. The idea is to come up with estimators that minimize (or in the suboptimal case, bound) the maximum energy gain (or so-called H^∞ norm) from the disturbances to the estimation errors. This will guarantee that if the disturbances are small (in energy) then the estimation errors will be as small as possible (in energy), *no matter what the disturbances are.*

Definition 1. (The H^∞ Norm). The H^∞ norm of a transfer operator $\mathcal{T}$ is defined as

$$\|\mathcal{T}\|_\infty = \sup_{x\in h^2, x\neq 0} \frac{\|\mathcal{T}x\|_2}{\|x\|_2} \qquad (2)$$

where h^2 denotes the space of all square-summable causal sequences.

Note that if $\mathcal{T}$ is a finite matrix, then $\|\mathcal{T}\|_\infty$ is simply $\bar{\sigma}(\mathcal{T})$, the maximum singular value of $\mathcal{T}$. When $\mathcal{T}$ is represented by a transfer matrix $T(z)$, then

$$\|\mathcal{T}\|_\infty = \sup_{0\leq\omega\leq 2\pi} \bar{\sigma}\left(T(e^{j\omega})\right).$$

Ideally, one would like to find the $\mathcal{K}$ that minimizes $\|\mathcal{T}_K\|_\infty$ over all possible disturbances. However, unlike in H^2 estimation, there are very few cases where a closed-form solution to this optimal H^∞ problem can be found (see (Hassibi *et al.*, 1996b)), and in general one relaxes the minimization and settles for a suboptimal solution.

Problem 1. (Suboptimal H^∞ Estimation Problem). Given $\gamma > 0$, determine whether a causal $\mathcal{K}$ that guarantees

$$\|\mathcal{T}_K\|_\infty = \sup_{u,v\in h^2, u,v\neq 0} \frac{\left\| \mathcal{T}_k \begin{bmatrix} u \\ v \end{bmatrix} \right\|_2}{(\|u\|_2^2 + \|v\|_2^2)^{1/2}} < \gamma, \quad (3)$$

exists, and if so, find one such estimator, $\mathcal{K}$.

4.1 *Solution to the Suboptimal H^∞ Problem*

Note that $\|\mathcal{T}_K\|_\infty < \gamma$ means that $\mathcal{T}_K\mathcal{T}_K^* < \gamma^2 I$, where I is the identity operator that maps input sequences to themselves. This inequality can be rewritten as

$$(\mathcal{L} - \mathcal{K}H)(\mathcal{L} - \mathcal{K}H)^* + \mathcal{K}\mathcal{K}^* - \gamma^2 I < 0,$$

or equivalently,

$$\begin{bmatrix} \mathcal{K} & I \end{bmatrix} \begin{bmatrix} I + \mathcal{H}\mathcal{H}^* & -\mathcal{H}\mathcal{L}^* \\ -\mathcal{L}\mathcal{H}^* & -\gamma^2 I + \mathcal{L}\mathcal{L}^* \end{bmatrix} \begin{bmatrix} \mathcal{K}^* \\ I \end{bmatrix} < 0. \quad (4)$$

Now it can be shown (Kwakernaak, 1986; Limebeer and Shaked, 1991; Grimble, 1993; Hassibi *et al.*, 1996b) that a causal $\mathcal{K}$ that guarantees the inequality (4) can be found iff the center block operator (or matrix) in (4) admits the following factorization

$$\begin{bmatrix} I + \mathcal{H}\mathcal{H}^* & -\mathcal{H}\mathcal{L}^* \\ -\mathcal{L}\mathcal{H}^* & -\gamma^2 I + \mathcal{L}\mathcal{L}^* \end{bmatrix} = \begin{bmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} \\ \mathcal{L}_{21} & \mathcal{L}_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} \mathcal{L}_{11}^* & \mathcal{L}_{21}^* \\ \mathcal{L}_{12}^* & \mathcal{L}_{22}^* \end{bmatrix}, (5)$$

where $\mathcal{L}_{11}$, $\mathcal{L}_{21}$ and $\mathcal{L}_{22}$ are causal and causally invertible, and $\mathcal{L}_{12}$ is strictly causal.

Once the factorization (5) has been performed we may rewrite (4) as follows

$$(\mathcal{K}\mathcal{L}_{11} + \mathcal{L}_{21})(\mathcal{K}\mathcal{L}_{11} + \mathcal{L}_{21})^* < (\mathcal{K}\mathcal{L}_{12} + \mathcal{L}_{22})(\mathcal{K}\mathcal{L}_{12} + \mathcal{L}_{22})^*.$$

Since both $(\mathcal{K}\mathcal{L}_{11} + \mathcal{L}_{21})$ and $(\mathcal{K}\mathcal{L}_{12} + \mathcal{L}_{22})$ are causal, this implies that we must have

$$(\mathcal{K}\mathcal{L}_{11} + \mathcal{L}_{21}) = (\mathcal{K}\mathcal{L}_{12} + \mathcal{L}_{22})\mathcal{Q},$$

for some *causal contractive* $\mathcal{Q}$ (*i.e.* , a causal $\mathcal{Q}$ such that $\mathcal{Q}\mathcal{Q}^* \leq I$). We can now solve the above equation for $\mathcal{K}$ and obtain the following result.

Theorem 1. (H^∞ Suboptimal Estimators). An H^∞ estimator that achieves level γ exists iff the factorization (5) (with the $\mathcal{L}_{ij}$ having the aforementioned properties) exists. If this is the case, then all possible H^∞ estimators of level γ are given by

$$\mathcal{K} = (\mathcal{L}_{22}\mathcal{Q} - \mathcal{L}_{21})(\mathcal{L}_{11} - \mathcal{L}_{12}\mathcal{Q})^{-1}, \qquad (6)$$

where $\mathcal{Q}$ is casual and $\mathcal{Q}\mathcal{Q}^* \leq I$. An important choice results from taking $\mathcal{Q} = 0$, so that $\mathcal{K}_{cen} = -\mathcal{L}_{21}\mathcal{L}_{11}^{-1}$, which is the so-called "central" filter.

5. MIXED H^2/H^∞ ESTIMATION

Unlike the H^2 estimation problem, the solution to the H^∞ estimation problem is highly nonunique. This is quite obvious for the suboptimal problem, but even the *optimal* H^∞ problem has, except for a few special cases, more than one solution. The main motivation of the mixed approach is to effectively exploit this non-uniqueness in order to improve some other aspects of the estimator besides robustness. In particular, the mixed H^2/H^∞ criterion attempts to improve the average performance of the estimator among the class of estimators having the same guaranteed level of robustness.

It is worthwhile to point out that, in application, the mixed approach makes sense only if the average performance varies significantly over the set of sub-optimal H^∞ estimators. Although, only a handful of results are available on the average performance of sub-optimal H^∞ estimators, as we show later (see Sections 7 and 8), the change in the average performance over the solution set may be quite significant even for small problems.

A large number of different variations to the mixed H^2/H^∞ problem have been discussed in the literature (see references (Khargonekar and Rotea, 1992) (Sznaier, 1994). However, for the sake of simplicity and brevity we consider the general mixed problem as stated below.

Problem 2. (Mixed H^2/H^∞ Estimation Problem) Given $\gamma > 0$, find a causal estimator $\mathcal{K}$ that minimizes the H^2

norm of the transfer operator $\mathcal{T}_\mathcal{K} = [\mathcal{L} - \mathcal{K}\mathcal{H} \quad -\mathcal{K}]$, subject to the H^∞ norm of $\mathcal{T}_\mathcal{K}$ being less than γ. In other words, find a causal $\mathcal{K}$ that satisfies

$$\min_{\mathcal{K}} \|\mathcal{T}_\mathcal{K}\|_2 \quad \text{subject to} \quad \|\mathcal{T}_\mathcal{K}\|_\infty \leq \gamma. \tag{7}$$

This optimization problem may be interpreted in two different ways: (i) conversion of the unconstrained H^2 optimization problem to one with an H^∞ constraint, or (ii) conversion of a suboptimal H^∞ feasibility problem to an optimization problem by selecting an H^2 norm as the optimization criterion. In either case, this conversion leads to a tradeoff between the H^2 and H^∞ norms of the transfer operator. To explicitly show this tradeoff consider the following form of the mixed problem

$$\min_{\mathcal{K}} \quad \alpha\|\mathcal{T}_\mathcal{K}\|_2 + (1-\alpha)\|\mathcal{T}_\mathcal{K}\|_\infty, \tag{8}$$

where the $\alpha \in [0 \quad 1]$, determines the relative weight of the two norms in the cost function. The forms (7) and (8) are equivalent in the sense that given a feasible γ, there is a corresponding $\alpha \in [0 \quad 1]$ (and vise verse) such that both problems have the same optimizer. For such γ, $\alpha/(\alpha-1)$ is the slope of the tangent to the tradeoff curve at the intersection of the curve with the line $\|\mathcal{T}_\mathcal{K}\|_\infty = \gamma$ (see Fig. (2)). All these results are direct consequences of the underlying convexity of the tradeoff curve.

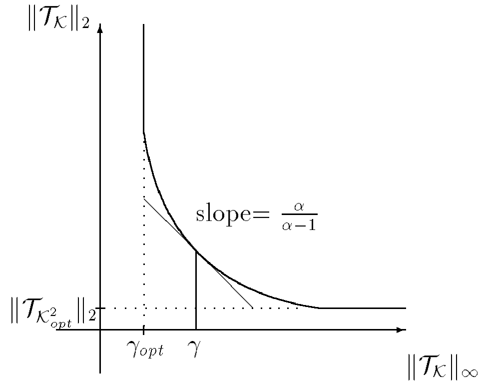

Fig. 2. $\|\mathcal{T}_\mathcal{K}\|_\infty$ *versus* $\|\mathcal{T}_\mathcal{K}\|_2$ *tradeoff curve.*

When the transfer operators $\mathcal{H}$ and $\mathcal{L}$ have state-space structure, unlike the H^2 and central H^∞ estimators, it is not known whether the optimal mixed estimator has state-space structure as well. (See (Khargonekar and Rotea, 1992), (Bernstein and Haddad, 1989) and (Foias*et al.*, 1995).)

6. CONVEX FORMULATION

In this section, for simplicity, we shall concentrate on the finite dimensional general mixed problem (7) and shall develop a semidefinite programming (SDP) formulation for it (Boyd *et al.*, 1994). This will then allow for efficient numerical solutions. We begin by noting that the finite horizon mixed problem can be restated as follows. *Given a feasible γ and two lower triangular matrices $\mathcal{H}$ and $\mathcal{L}$ find a lower triangular matrix $\mathcal{K}$ satisfying*

$$\min_{\mathcal{K}} \quad \text{trace}[\mathcal{T}_\mathcal{K}\mathcal{T}_\mathcal{K}^*] \quad \text{subject to} \quad \|\mathcal{T}_\mathcal{K}\|_\infty \leq \gamma, \tag{9}$$

where $\mathcal{T}_\mathcal{K} = [\mathcal{L} - \mathcal{K}\mathcal{H} \quad -\mathcal{K}]$. As mentioned earlier, the objective function is a convex (quadratic) function of $\mathcal{K}$ and the constraint set is a compact convex set. Hence, the mixed problem, is a convex programming problem. However, for efficient numerical solutions we would like to reformulate the problem as a SDP. An obvious SDP formulation can be obtained using a scalar slack variable t as follows.

Problem 3. (First SDP Formulation of the General Mixed Problem) The finite horizon mixed problem (9) is equivalent to the following SDP:

$$\min_{t,\mathcal{K}} \quad t$$
$$\text{subject to} \quad \begin{bmatrix} t & \text{vec}(\mathcal{T}_\mathcal{K})^* & & \\ \text{vec}(\mathcal{T}_\mathcal{K})^* & \mathcal{I} & & \\ & & \gamma^2\mathcal{I} & \mathcal{T}_\mathcal{K} \\ & & \mathcal{T}_\mathcal{K}^* & \mathcal{I} \end{bmatrix} \geq 0 \tag{10}$$

where $\mathcal{T}_\mathcal{K} = [\mathcal{L} - \mathcal{K}\mathcal{H} \quad -\mathcal{K}]$, $\mathcal{K}$ is lower triangular and $\text{vec}(\cdot)$ is the vectorization operation.

Note that we have converted problem (9) into the minimization of a linear objective subject to a linear matrix inequality (LMI) constraint. Applying Schur complements to the LMI it is easy to see that the first block guarantees that t is an upper bound on the H^2 norm of the transfer operator and that the second block guarantees that γ is an upper bound on the H^∞ norm of the transfer operator. Hence, minimizing the variable t subject to the LMI is is equivalent to the original problem (7). However, it turns out that from a computational point of view this is not the most efficient SDP formulation. A more efficient SDP formulation results using a matrix slack variable and is given below.

Problem 4. (SDP Formulation of the General Mixed Problem) The finite horizon mixed problem (9) is equivalent to the following SDP:

$$\min_{\mathcal{W},\mathcal{K}} \quad -trace\mathcal{W}$$
$$\text{subject to} \quad \begin{bmatrix} \gamma^2\mathcal{I} - \mathcal{W} & \mathcal{T}_\mathcal{K} & \\ \mathcal{T}_\mathcal{K}^* & \mathcal{I} & \\ & & \mathcal{W} \end{bmatrix} \geq 0 \tag{11}$$

where $\mathcal{T}_\mathcal{K} = [\mathcal{L} - \mathcal{K}\mathcal{H} \quad -\mathcal{K}]$ and $\mathcal{K}$ is lower triangular.

In this formulation, $\mathcal{W}$ can be interpreted as a lower bound on the gap between the two matrices $\gamma^2\mathcal{I}$ and $\mathcal{T}_{\mathcal{K}}\mathcal{T}_{\mathcal{K}}^*$. Hence, $\mathcal{W} \geq 0$ implies that the H^∞ norm of the transfer operator is upper bounded by γ. Moreover, the sum of the trace of $\mathcal{W}$ and the H^2 norm of the transfer matrix is upper bounded by $\sqrt{n}\gamma$ (n is the size of the problem). As a result, maximization of the trace of $\mathcal{W}$ is in effect equivalent to minimization of the H^2 norm of the transfer matrix. Note that, as before, the set of lower triangular matrices $\mathcal{K}$ and the positive semidefinite matrices $\mathcal{W}$ which satisfy the LMI is a compact convex set. This SDP can be solved using efficient algorithms such as the primal-dual method (Boyd *et al.*, 1994).

7. GEOMETRIC INTERPRETATION AND SOME RELATED RESULTS

Geometrically, the constrained problem (9) is equivalent to finding the point in the set $\mathbf{S}_\gamma = \{\mathcal{K} : \|\mathcal{T}_{\mathcal{K}}\|_\infty \leq \gamma\}$, that is *closest* to the optimal H^2 estimator, $\mathcal{K}^2_{opt}$ (see Fig (3)). Hence, except for the trivial case where $\mathcal{K}^2_{opt}$

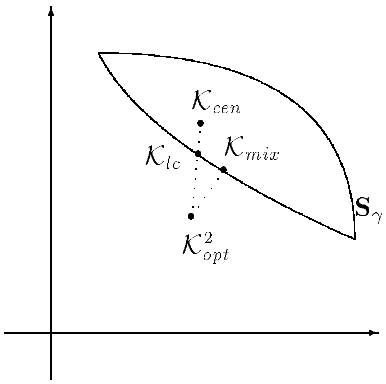

Fig. 3. *The set* $\mathbf{S}_\gamma$ *and different estimators.*

satisfies the H_∞ norm constraint, the solution to the mixed problem lies on the boundary of the set $\mathbf{S}_\gamma$ as stated in the following result.

Result 1. Depending on the value of γ the location of the solution to Problem 2 can be classified as follows:

- For $\gamma < \|\mathcal{T}_{\mathcal{K}^2_{opt}}\|_\infty$ the solution to the mixed problem lies on the boundary of the set $\mathbf{S}_\gamma$. Therefore the corresponding H_∞ norm of the transfer operator is equal to γ.
- For $\gamma \geq \|\mathcal{T}_{\mathcal{K}^2_{opt}}\|_\infty$, the optimal H^2 estimator, $\mathcal{K}^2_{opt}$, is the solution to the mixed problem.

Note that this result also follows from the convexity of the tradeoff curve. Next, we characterize the boundary in terms of the causal contractive $\mathcal{Q}$.

Result 2. For any feasible $\gamma > \gamma_{opt}$, the following statements are equivalent.

- $\|\mathcal{T}_{\mathcal{K}}\|_\infty = \gamma$.
- $\mathcal{K}$ is on the boundary of the set $\{\mathcal{K} : \|\mathcal{T}_{\mathcal{K}}\|_\infty \leq \gamma\}$.
- $\mathcal{K} = (\mathcal{L}_{22}\mathcal{Q} - \mathcal{L}_{21})(\mathcal{L}_{11} - \mathcal{L}_{12}\mathcal{Q})^{-1}$, and $\bar{\sigma}(\mathcal{Q}) = 1$.

From last two results it directly follows that for $\gamma < \|\mathcal{T}_{\mathcal{K}^2_{opt}}\|_\infty$ the central solution $\mathcal{K}_{cen}$ is not optimal, since it corresponds to $\mathcal{Q} = 0$.

Another important question is whether the optimal mixed estimator $\mathcal{K}$ has state-space structure when both $\mathcal{H}$ and $\mathcal{L}$ have state-space structure. Although, this is still an open problem, preliminary results indicate that this is not the case (Foias*et al.*, 1995). In general, however, $\mathcal{K}$ will have state-space structure if and only if $\mathcal{Q}$ has state-space structure. It readily follows that the central solution $\mathcal{K}_{cen}$ will have state-space structure (since $\mathcal{Q} = 0$ has state-space structure). Thus, if we restrict the estimator to a convex combination of $\mathcal{K}_{cen}$ and $\mathcal{K}^2_{opt}$, i.e., $\mathcal{K}_{lc} = \beta\mathcal{K}_{cen} + (1-\beta)\mathcal{K}^2_{opt}$, $\beta \in [0 \quad 1]$, then a state-space realization for the resulting estimator is guaranteed. It can be shown that for such an estimator we have, $\|\mathcal{T}_{\mathcal{K}_{lc}}\|_2 = \beta^2\|\mathcal{T}_{\mathcal{K}_{cen}}\|_2 + (1-\beta^2)\|\mathcal{T}_{\mathcal{K}^2_{opt}}\|_2$. Hence, for any $\beta < 1$, we get an estimator which has improved average performance compared to the central estimator. Moreover, the $\mathcal{K}_{lc}$ that lies on the boundary corresponds to the smallest permissible β and provides a suboptimal solution to the mixed problem with guaranteed state-space structure.

The last result of this section is related to one of our earlier comments about the spread of the H^2 norm over the set of all γ suboptimal H^∞ estimators.

Result 3. The estimators that maximize the H^2 norm of $\mathcal{T}_{\mathcal{K}}$ over the set of γ suboptimal H^∞ estimators, are given by $\mathcal{K} = (\mathcal{L}_{22}\mathcal{Q} - \mathcal{L}_{21})(\mathcal{L}_{11} - \mathcal{L}_{12}\mathcal{Q})^{-1}$, and $\mathcal{Q}\mathcal{Q}^* = \mathcal{I}$. The maximum value of H^2 norm is $\sqrt{n}\gamma$, where n is the size of the transfer matrix $\mathcal{T}_{\mathcal{K}}$

8. NUMERICAL EXAMPLE

To illustrate some properties of different estimators, we present a numerical example where $\mathcal{H} = \mathcal{L}$ are derived from adaptive filtering. The data model corresponds to a simple adaptive filter with a single scalar weight and the size of the transfer operator is set to 15. The selected γ is 10% larger than the optimal, γ_{opt}. For this γ, we numerically computed the central estimator $\mathcal{K}_{cen}$, the (near) optimal mixed estimator (using the SDP (11)), the optimal linearly combined estimator $\mathcal{K}_{lc}$, and an H^2 norm maximizing estimator using Result 3. The distribution of the (squared) singular values of the transfer matrices corresponding to these four estimators and the optimal H^2 estimator are plotted in Fig. 4. Clearly, the average performance of the norm maximizing estimator is far worst than the optimal mixed estimator. This shows that the spread of the H^2 norm over the suboptimal H^∞ estimator set is quite large. Moreover, there is a significant performance difference between the central and the optimal mixed estimator, which shows that the solution to the mixed problem has significant practical implications besides theoretical importance. Finally, although the best linearly combined estimator has better average performance than the central estimator, it is not as good as the best mixed H^2/H^∞ estimator.

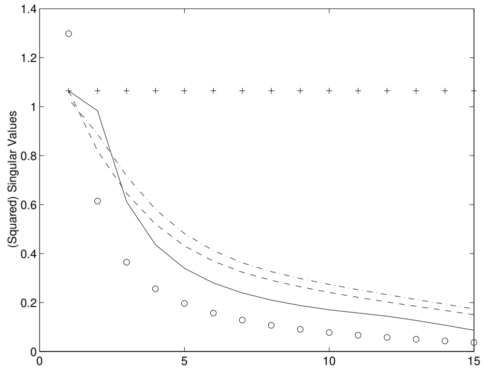

Fig. 4. *Distribution of (squared) singular values of $\mathcal{T}_{\mathcal{K}}$ for different estimators.*
'o': H^2-optimal estimator,
'–': H^2/H^∞ optimal estimator,
'- -': Optimal linearly combined estimator,
'-.': Central estimator,
'+': H^2 norm maximizing estimator.

9. CONCLUSION

We introduced and motivated the mixed H^2/H^∞ estimation problem that exploits the non-uniqueness of the suboptimal H^∞ estimators to obtain the estimator with the best average performance, and essentially combines the purely stochastic and purely deterministic approaches of H^2 and H^∞ estimation. In this paper we attempted to attack the problem of obtaining such estimators and to study their properties. For finite-horizon problems we presented an SDP formulation of the problem that can be numerically solved using efficient convex optimization methods, and we also obtained preliminary analytic characterizations of the solution.

10. REFERENCES

G. Zames (1981). Feedback and optimal sensitivity: model reference transformations, multiplicative seminorms, and approximate inverses, In *Proc. IEEE Trans. on Automatic Control*, AC-26, pp. 301-320.

U. Shaked and Y. Theodor (1992). $H^\infty-$optimal estimation: A tutorial. In *Proc. IEEE Conference on Decision and Control*, pages 2278–2286, Tucson, AZ, Dec. 1992.

P.P. Khargonekar and K. M. Nagpal (1991). Filtering and smoothing in an $H^\infty-$ setting. *IEEE Trans. on Automatic Control*, AC-36:151–166.

T. Basar (1991). Optimum performance levels for minimax filters, predictors and smoothers. *Systems and Control Letters*, 16:309–317.

H. Kwakernaak (1986). A polynomial approach to mimimax frequency domain optimization of multivariable feedback systems. *Int. J. of Control*, 44:117–156.

D. J. Limebeer and U. Shaked (1991). New results in h^∞-filtering. In *Proc. Int. Symp. on MTNS*, pages 317–322, June 1991.

M. J. Grimble (1993). Polynomial matrix solution of the H^∞ filtering problem and the relationship to Riccati equation state-space results. *IEEE Trans. on Signal Processing*, 41(1):67–81, January 1993.

B. Hassibi, A. H. Sayed, and T. Kailath (1996). Recursive linear estimation in Krein spaces - parts I and II. In *Proc. IEEE Transactions on Automatic Control*, Jan. 1996.

B. Hassibi, A. H. Sayed, and T. Kailath (1996). H^∞ optimality of the LMS algorithm. To appear in *IEEE Trans. on Signal Processing*, Feb. 1996.

P. P. Khargonekar and M. A. Rotea (1992). Mixed H^2/H^∞ filtering. In *Proc. IEEE Conference on Decision and Control*, Tucson, AZ, Dec. 1992, pp. 2299-304.

K. Zhou, K. Glover, B. Bodenheimer and J. C. Doyle (1994). Mixed H^2 and H^∞ performance objectives I and II. *IEEE Transactions on Automatic Control*, vol.39, no.8, pp. 1564-1587.

D. S. Bernstein and W. M. Haddad (1989). LQG control with an H^∞ performance bound: A Riccati equation approach. *IEEE Transactions on Automatic Control*, vol.34, no.3, pp. 293-305.

H. H. Yeh, S. S. Banda and B. C. Chang (1992). Necessary and sufficient conditions for mixed H^2 and H^∞ optimal control. *IEEE Transactions on Automatic Control*, vol.37, no.3, pp. 355-358.

M. Sznaier (1994). An exact solution to general SISO mixed H^2/H^∞ problems via convex optimization. *IEEE Transactions on Automatic Control*, vol.39, no.3, pp. 2511-2517.

S. Boyd, L. El Ghaoui, E. Feron and V. Balakrishnan (1994). *Linear Matrix Inequalities in System and Control Theory.* SIAM Studies in Applied Mathematics.

C. Foias, A. E. Frazho and W. S. Li (1995) On H^2 minimization for the Caratheodory-Schur interpolation problem, *Integral Equations and Operator Theory*, vol.21, no.1, pp. 24-32.