On Nonlinear Filters for Mixed H^2/H^∞ Estimation*

Babak Hassibi and Thomas Kailath
Information Systems Laboratory
Stanford University, Stanford CA 94305

Abstract

We study the problem of mixed least-mean-squares/H^∞-optimal (or mixed H^2/H^∞-optimal) estimation of signals generated by discrete-time, finite-dimensional, linear state-space models. The major result is that, for finite-horizon problems, and when the stochastic disturbances have Gaussian distributions, the optimal solutions have finite-dimensional (*i.e.* , bounded-order) *nonlinear* state-space structure of order $2n+1$ (where n is the dimension of the underlying state-space model). Being nonlinear, strictly speaking, the filters do not minimize an H^2 norm subject to an H^∞ constraint, but instead minimize the least-mean-squares estimation error (given a certain a priori probability distribution on the disturbances) subject to a given constraint on the maximum energy gain from disturbances to estimation errors. The mixed filters therefore have the property of yielding the best average (least-mean-squares) performance over all filters that achieve a certain worst-case (H^∞) bound.

1. Introduction

Classical methods in estimation theory (such as least-mean-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. All classical estimation methods require some assumption regarding the statistical nature of the signals 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.

The mixed H^2/H^∞ estimation problem was introduced (see *e.g.*, [2, 3, 4, 5] and the references therein) as a compromise between these two extreme point of views. The mixed 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^∞ estimators have the best average performance among all estimators having a guaranteed worst-case performance. Thus, the best average performance is sacrificed to attain a certain level of robustness.

Unlike the unconstrained H^2 and suboptimal H^∞ problems the *pure* mixed H^2/H^∞ problem of minimizing an H^2 norm, subject to an H^∞ norm constraint, has been an open problem. Indeed in [6, 7] it has been shown that for infinite-horizon problems, and when the underlying models are linear-time-invariant (LTI), the *linear* mixed H^2/H^∞-optimal controller (or estimator) is infinite-dimensional (if, of course, the H^∞ constraint is not redundant). [For this reason, recently several related problems with an auxiliary cost (which replaces the H^2 norm) have been considered (see *e.g.*, [2, 3, 4]).]

In this paper we expand the domain and allow for nonlinear estimators. Of course, once we have a nonlinear estimator we cannot really speak of the H^2 norm (or the H^∞ norm for that matter). Therefore instead of minimizing the H^2 norm, subject to an H^∞ norm constraint, these estimators minimize the expected estimation error energy (given a certain probablity distribution on the disturbances), subject to a bound on the worst-case energy gain from the disturbances to the estimation errors.

A major result of this paper is that in the mixed H^2/H^∞ problem, even when the underlying model is linear, nonlinear filters offer an improvement over linear ones. This may appear surprising given the fact that for linear plants the optimal H^2 and optimal (central) H^∞ estimators are both linear. In other words, in H^2 and in H^∞ estimation there is nothing to be gained by considering nonlinear estimators. However, in the mixed problem it is possible to further reduce the expected estimation error energy by considering nonlinear filters. More important, is the fact that (for finite-horizon problems) the resulting estimator has (bounded) finite-dimensional state-space structure.

The nonlinearity of the optimal mixed estimator arises from the fact that at each iteration we need to solve a nonlinear program (with dimension equal to the number of signals to be estimated). At the present we have not been able to give an explicit solution to this nonlinear program, but we have been able to come up with a (suboptimal) recursive solution, which involves solving a convex quadratic program at each iteration.

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

2. Problem Formulation

Consider the following (possibly) time-variant discrete-time linear state-space model

$$\left\{ \begin{array}{rcl} x_{i+1} & = & F_i x_i + G_i^{(1)} u_i^{(1)} + G_i^{(2)} u_i^{(2)} \\ y_i & = & H_i x_i + v_i^{(1)} + v_i^{(2)} \\ z_i & = & L_i x_i \end{array} \right. \qquad x_0 = x_0^{(1)} + x_0^{(2)}, \tag{1}$$

where $\{F_i, G_i^{(1)}, G_i^{(2)}, H_i, L_i\}$ are known matrices of dimensions $n \times n$, $n \times m_1$, $n \times m_2$, $p \times n$ and $q \times n$, respectively, $\{x_0, u_i^{(1)}, u_i^{(2)}, v_i^{(1)}, v_i^{(2)}\}$ are unknown disturbances, $\{y_i\}$ is the observed output, and $\{z_i\}$ is the signal we intend to estimate. We have separated the unknown initial state $x_0 = x_0^{(1)} + x_0^{(2)}$, the unknown driving disturbance $G_i^{(1)} u_i^{(1)} + G_i^{(2)} u_i^{(2)}$ and the unknown measurement disturbance $v_i^{(1)} + v_i^{(2)}$ into two components since each component has a different nature. The initial condition $x_0^{(1)}$ and the disturbances $\{u_i^{(1)}, v_i^{(1)}\}$ are assumed to be (zero-mean) random variables with known joint probability distribution, and are used to reperesent that component of the disturbances for which we know the underlying probability distribution. The initial condition $x_0^{(2)}$ and the disturbances $\{u_i^{(2)}, v_i^{(2)}\}$ are assumed to be deterministic, but unknown, signals, and are used to represent model uncertainties in (1) and that component of the disturbances for which we have no a priori statistical knowledge.

In this paper we shall be concerned with the problem of estimating the signal, $z_i = L_i x_i$, using the observations, y_i. In particular, we shall be interested in the following two estimates:

$$\hat{z}_{i|i} = \mathcal{F}_{f,i}\{y_0, y_1, \ldots y_i\}, \tag{2}$$

which is referred to as the *a posteriori* or *filtered* estimate, and

$$\hat{z}_i = \mathcal{F}_{p,i}\{y_0, y_1, \ldots y_{i-1}\}, \tag{3}$$

which is referred to as the *a priori* or *predicted* estimate. Corresponding to these two estimates we will also have the following filtered and prediction errors,

$$\tilde{z}_{i|i} = e_{f,i} = z_i - \hat{z}_{i|i} \quad \text{and} \quad \tilde{z}_i = e_{p,i} = z_i - \hat{z}_i.$$

2.1. H^2 Estimation

In this case, we assume that

$$x_0^{(2)} = 0 \ , \ u_i^{(2)} = 0 \ , \ v_i^{(2)} = 0$$

so that the state-space model (1) becomes

$$\left\{ \begin{array}{rcl} x_{i+1} & = & F_i x_i + G_i^{(1)} u_i^{(1)} \\ y_i & = & H_i x_i + v_i^{(1)} \\ z_i & = & L_i x_i \end{array} \right. \qquad x_0 = x_0^{(1)}. \tag{4}$$

In the H^2 estimation problem we would like to find $\hat{z}_{i|i}$ and $\hat{z}_i$ such that the expected estimation error energies

$$E \sum_{i=0}^{N} (z_i - \hat{z}_{i|i})^* (z_i - \hat{z}_{i|i}) \quad \text{and} \quad E \sum_{i=0}^{N} (z_i - \hat{z}_i)^* (z_i - \hat{z}_i)$$

are respectively minimized, where the expectation is taken over the random variables $\{x_0^{(1)}, \{u_i^{(1)}, v_i^{(1)}\}_{i=0}^N\}$.

The solution to the above problem is wellknown and is given by the conditional mean of z_i subject to the observations. When the $\{x_0^{(1)}, \{u_i^{(1)}, v_i^{(1)}\}_{i=0}^N\}$ are independent zero-mean Gaussian random variables with known variances, then the conditional means (*i.e.*, the $\hat{z}_{i|i}$ and $\hat{z}_i$) are readily found via the Kalman filter recursions (which has also has a state-space model of order n).

2.2. H^∞ Estimation

In this case, we assume that

$$x_0^{(1)} = 0 \ , \ u_i^{(1)} = 0 \ , \ v_i^{(1)} = 0$$

so that the state-space model (1) becomes

$$\left\{ \begin{array}{rcl} x_{i+1} & = & F_i x_i + G_i^{(2)} u_i^{(2)} \\ y_i & = & H_i x_i + v_i^{(2)} \\ z_i & = & L_i x_i \end{array} \right. \qquad x_0 = x_0^{(2)}. \tag{5}$$

In the H^∞ estimation problem the initial state $x_0^{(2)}$ and the disturbances $\{u_i^{(2)}, v_i^{(2)}\}_{i=0}^N$ are assumed to be deterministic but unknown. Therefore we have no statistical assumptions and cannot speak of expected values. In this problem, to ensure robustness of the estimator with respect to model uncertainty and lack of statistical knowledge, it is proposed to find the estimates $\hat{z}_{i|i}$ and $\hat{z}_i$ such that the worst-case energy gain from the disturbances to the estimation errors be bounded by given thresholds γ_f and γ_p. In other words, we would like to find, if possible, all $\mathcal{F}_{f,i}(\cdot)$ and $\mathcal{F}_{p,i}(\cdot)$ such that

$$\sup_{x_0^{(2)} \neq 0, \{u_i^{(2)}, v_i^{(2)}\} \neq 0} \frac{\|e_f\|^2}{x_0^{(2)*} \Pi_0^{-1} x_0^{(2)} + \|u\|^2 + \|v\|^2} \leq \gamma_f^2, \tag{6}$$

and

$$\sup_{x_0^{(2)} \neq 0, \{u_i^{(2)}, v_i^{(2)}\} \neq 0} \frac{\|e_p\|^2}{x_0^{(2)*} \Pi_0^{-1} x_0^{(2)} + \|u\|^2 + \|v\|^2} \leq \gamma_p^2, \tag{7}$$

where $\Pi_0 > 0$ is given and $\|a\|^2 \triangleq \sum_{i=0}^N a_i^* a_i$.

The next theorem gives a sufficient condition for the existence of estimators that achieve (6) and (7) (the condition is essentially the necessary and sufficient condition for the existence of estimators that achieve strict inequalities) and then parametrizes all possible solutions.

Theorem 1 (H^∞ Filters) *(a) An estimator $\hat{z}_{i|i} = \mathcal{F}_{f,i}\{y_0, y_1, \ldots y_i\}$ that achieves (6) with strict inequality exists if, and only if, the matrices $R_i = I_p \oplus (-\gamma^2 I_q)$ and*

$$R_{e,i} = R_i + \begin{bmatrix} H_i \\ L_i \end{bmatrix} P_i \begin{bmatrix} H_i^* & L_i^* \end{bmatrix}$$

have the same inertia for all $i = 0, \ldots, N$, where P_i satisfies the Riccati recursion

$$P_{i+1} = F_i P_i F_i^* + G^{(2)} G^{(2)*} - K_{p,i} R_{e,i} K_{p,i}^*, \tag{8}$$

with $P_0 = \Pi_0$ and $K_{p,i} = F_i P_i \begin{bmatrix} H_i^* & L_i^* \end{bmatrix} R_{e,i}^{-1}$. If this is the case then all a posteriori estimators that achieve (6) are characterized by

$$\sum_{j=0}^{i} \begin{bmatrix} y_j - H_j\hat{x}_j \\ \hat{z}_{j|j} - L_j\hat{x}_j \end{bmatrix}^* R_{e,j}^{-1} \begin{bmatrix} y_j - H_j\hat{x}_j \\ \hat{z}_{j|j} - L_j\hat{x}_j \end{bmatrix} \geq 0, \qquad (9)$$

for $i = 0, \ldots, N$, where $\hat{x}_i$ satisfies the recursion

$$\hat{x}_{i+1} = F_i\hat{x}_i + K_{p,i} \begin{bmatrix} y_i - H_i\hat{x}_i \\ \hat{z}_{i|i} - L_i\hat{x}_i \end{bmatrix}, \quad \hat{x}_0 = 0. \qquad (10)$$

(b) An estimator $\hat{z}_i = \mathcal{F}_{p,i}\{y_0, y_1, \ldots y_{i-1}\}$ that achieves (7) with strict inequality exists if, and only if,

$$-\gamma^2 I_q + L_i P_i L_i^* < 0 \;\; and \;\; I_p + H_i(P_i^{-1} - \gamma^{-2}L_i^* L_i)^{-1}H_i^* > 0$$

where P_i satisfies (8). If this is the case then all a priori estimators that achieve (7) are characterized by

$$\sum_{j=0}^{i-1} \begin{bmatrix} y_j - H_j\hat{x}_j \\ \hat{z}_j - L_j\hat{x}_j \end{bmatrix} R_{e,j}^{-1} \begin{bmatrix} y_j - H_j\hat{x}_j \\ \hat{z}_j - L_j\hat{x}_j \end{bmatrix}^* - (\hat{z}_i - L_i\hat{x}_i)^*(\gamma^2 I_q - L_i P_i L_i^*)^{-1}(\hat{z}_i - L_i\hat{x}_i) \geq 0, \qquad (11)$$

for $i = 0, \ldots, N$, where $\hat{x}_i$ satisfies the recursion

$$\hat{x}_{i+1} = F_i\hat{x}_i + K_{p,i} \begin{bmatrix} y_i - H_i\hat{x}_i \\ \hat{z}_i - L_i\hat{x}_i \end{bmatrix}, \quad \hat{x}_0 = 0 \qquad (12)$$

with $R_{e,i}$ and $K_{p,i}$ as in part (a).

The above parametrization of all possible H^∞ estimators is not in the form often cited in the literature, which is given in terms of a linear contraction (see, *e.g.*, [1]). However, it will be most suitable for solving the mixed H^2/H^∞ estimation problem, as done in the next section, since it does not disclude nonlinear estimators.

2.3. Mixed H^2/H^∞ Estimation

Let us now return to the general state-space model, (4). In mixed H^2/H^∞ estimation the goal is to construct an estimate of the signal, z_i, that effectively deals with the combined effects of random disturbances with known probability density functions (which are represented by $\{x_0^{(1)}, \{u_i^{(1)}, v_i^{(1)}\}_{i=0}^N\}$) and unknown (but deterministic) disturbances and modeling errors (which are represented by $\{x_0^{(2)}, \{u_i^{(2)}, v_i^{(2)}\}_{i=0}^N\}$). Therefore, it is proposed to minimize the expected estimation error energy (assuming the deterministic disturbances are zero) over all estimators that achieve a certain worst-case energy gain bound from the disturbances to the estimation errors (assuming the random disturbances are zero).

In other words, we would like to find estimation strategies $\hat{z}_{i|i} = \mathcal{F}_{f,i}\{y_0, y_1, \ldots y_i\}$ and $\hat{z}_i = \mathcal{F}_{p,i}\{y_0, y_1, \ldots y_{i-1}\}$ that respectively minimize the expected estimation error energies

$$E\sum_{i=0}^{N}(z_i - \hat{z}_{i|i})^*(z_i - \hat{z}_{i|i}) \quad \text{and} \quad E\sum_{i=0}^{N}(z_i - \hat{z}_i)^*(z_i - \hat{z}_i)$$

where the expectation is taken over the random variables $\{x_0^{(1)}, \{u_i^{(1)}, v_i^{(1)}\}_{i=0}^N\}$ (given $\{x_0^{(2)} = 0, \{u_i^{(2)}, v_i^{(2)}\} = \{0,0\}\}$), subject to the energy gain constraints (6) and (7) (where it is now assumed that $\{x_0^{(1)} = 0, \{u_i^{(1)}, v_i^{(1)}\} = \{0,0\}\}$).

3. Finite-Horizon Solution

Theorem 2 (Mixed H^2/H^∞ Estimators) *Consider the state-space model (4) and suppose that the disturbances $\{x_0^{(1)}, \{u_i^{(1)}, v_i^{(1)}\}_{i=0}^N\}$ are independent zero-mean Gaussian random variables with known covariances,*

$$E\begin{bmatrix} x_0^{(1)} \\ u_i^{(1)} \\ v_i^{(1)} \end{bmatrix}\begin{bmatrix} x_0^{(1)} \\ u_i^{(j)} \\ v_i^{(j)} \end{bmatrix}^* = \begin{bmatrix} \Pi_0 & 0 \\ 0 & \begin{bmatrix} Q_i & 0 \\ 0 & R_i \end{bmatrix}\delta_{ij} \end{bmatrix}.$$

(a) An estimator $\hat{z}_{i|i} = \mathcal{F}_{f,i}\{y_0, y_1, \ldots y_i\}$ that solves the a posteriori mixed H^2/H^∞ estimation problem (with level γ) is given by the solution to the following nonlinear program: for $i = 0, \ldots N$,

$$\begin{cases} \min_{\hat{z}_{i|i}} & (\hat{z}_{i|i} - \bar{z}_{i|i})^*(\hat{z}_{i|i} - \bar{z}_{i|i}) + \Psi_{i+1}(\tilde{x}_{i+1}, J_{i+1}) \\ \text{subject to} & J_i + N_i \geq 0 \end{cases} \qquad (13)$$

where

$$N_i = \begin{bmatrix} \bar{e}_i - H_i\tilde{x}_i \\ (\hat{z}_{i|i} - \bar{z}_{i|i}) - L_i\tilde{x}_i \end{bmatrix}^* R_{e,i}^{-1} \begin{bmatrix} \bar{e}_i - H_i\tilde{x}_i \\ (\hat{z}_{i|i} - \bar{z}_{i|i}) - L_i\tilde{x}_i \end{bmatrix},$$

and $\bar{z}_{i|i} = L_i(\bar{x}_i + \bar{P}_i H_i^*(R_i + H_i\bar{P}_i H_i^*)^{-1}\bar{e}_i)$ is the H^2 a posteriori estimate, $\bar{e}_i = y_i - H_i\bar{x}_i$ is the innovations, $\tilde{x}_i = \hat{x}_i - \bar{x}_i$, and where $\bar{x}_i$, $\hat{x}_i$ and J_i satisfy the recursions,

$$\begin{cases} \bar{x}_{i+1} &=& F_i\bar{x}_i + \bar{K}_{p,i}(y_i - H_i\bar{x}_i) \\ \hat{x}_{i+1} &=& F_i\hat{x}_i + K_{p,i}\begin{bmatrix} y_i - H_i\hat{x}_i \\ \hat{z}_{i|i} - L_i\hat{x}_i \end{bmatrix} \\ J_{i+1} &=& J_i + \begin{bmatrix} y_i - H_i\hat{x}_i \\ \hat{z}_{i|i} - L_i\hat{x}_i \end{bmatrix}^* R_{e,i}^{-1}\begin{bmatrix} y_i - H_i\hat{x}_i \\ \hat{z}_{i|i} - L_i\hat{x}_i \end{bmatrix} \end{cases}$$

all initialized with zero, where $K_{p,i}$ and $R_{e,i}$ are as in Theorem 1, $\bar{K}_{p,i} = F_i\bar{P}_i H_i^*\bar{R}_{e,i}^{-1}$, $\bar{R}_{e,i} = R_i + H_i\bar{P}_i H_i^*$, and where $\bar{P}_i$ satisfies the Riccati recursion,

$$\bar{P}_{i+1} = F_i\bar{P}_i F_i^* + G_i^{(1)}Q_i G_i^{(1)*} - \bar{K}_{p,i}\bar{R}_{e,i}\bar{K}_{p,i}^*, \quad \bar{P}_0 = \Pi_0.$$

Finally, if we denote the solution to (13) by the function $\Psi_i^1(\tilde{x}_i, J_i, \bar{e}_i)$, then the nonlinear functions $\Psi_i(\cdot,\cdot)$ are given by the following backward functional recursion,

$$\Psi_i(\tilde{x}_i, J_i) = \int_{-\infty}^{\infty} \frac{exp(-\bar{e}_i^*(2\bar{R}_{e,i})^{-1}\bar{e}_i)}{(\sqrt{2\pi})^p\, det(\bar{R}_{e,i})}\Psi_i^1(\tilde{x}_i, J_i, \bar{e}_i)d\bar{e}_i, \qquad (14)$$

initialized with $\Psi_{N+1}(\cdot,\cdot) = 0$.

(b) An estimator $\hat{z}_i = \mathcal{F}_{p,i}\{y_0, y_1, \ldots y_{i-1}\}$ that solves the a priori mixed H^2/H^∞ estimation problem (with level γ) is given by the solution to the following nonlinear program: for $i = 0, \ldots N$,

$$\begin{cases} \min_{\hat{z}_i} & (\bar{z}_i - \hat{z}_i)^*(\bar{z}_i - \hat{z}_i) + \Phi_i(\tilde{x}_i, M_i) \\ \text{subject to} & M_i \geq 0 \end{cases}, \qquad (15)$$

where $\bar{z}_i = L_i\bar{x}_i$ is the H^2 a priori estimate,

$$M_i = J_i - (\hat{z}_i - L_i\hat{x}_i)^*(\gamma^2 I_q - L_i P_i L_i^*)^{-1}(\hat{z}_i - L_i\hat{x}_i), \quad (16)$$

$\tilde{x}_i = \hat{x}_i - \bar{x}_i$, and where $\bar{x}_i$, $\hat{x}_i$ and J_i satisfy the same recursions as part (a), except that $\hat{z}_{i|i}$ is replaced by $\hat{z}_i$. Finally, if we denote the solution to (15) by the function $\Phi_i^1(\tilde{x}_{i-1}, M_{i-1}, \bar{e}_{i-1})$, the nonlinear functions $\Phi_i(\cdot,\cdot)$ are given by the following backward functional recursion,

$$\Phi_i(\tilde{x}_i, M_i) = \int_{-\infty}^{\infty} \frac{exp(-\bar{e}_i^*(2\bar{R}_{e,i})^{-1}\bar{e}_i)}{(\sqrt{2\pi})^p\, det(\bar{R}_{e,i})}\Phi_{i+1}^1(\tilde{x}_i, M_i, \bar{e}_i)d\bar{e}_i, \quad (17)$$

initialized with $\Phi_N(\cdot,\cdot) = 0$.

The H^2/H^∞-optimal estimators of Theorem 2 are nonlinear filters with finite-dimensional state-space structure of order $2n+1$. Indeed it is easy to see that the state variable is $\xi_i = \{\bar{x}_i, \hat{x}_i, J_i\}$.

We should note that, in contrast to both the H^2-optimal and the central H^∞ filters, the above solution is non-recursive in the sense that the solution depends on the horizon N. Indeed the filters obtained for problems with horizon N and $N+1$ are completely different.

4. A Recursive Solution

To obtain the desired filters we must first solve the backwards functional recursions (14) and (17) for $\Psi_i(\cdot,\cdot)$ and $\Phi_i(\cdot,\cdot)$.[1] This appears to be a formidable task and we have not yet been able to compute these functions.

To somewhat alleviate this problem, we can instead define a recursive mixed H^2/H^∞ filtering problem where one attempts to minimize the expected estimation error energies <u>at each time instant i</u> (subject, of course, to the given H^∞ constraints). The solution to this problem is given below.

Theorem 3 (Recursive Mixed H^2/H^∞ Estimators) *Consider the state-space model (4) and the setting of Theorem 2.*
(a) An estimator $\hat{z}_{i|i} = \mathcal{F}_{f,i}\{y_0, y_1, \ldots y_i\}$ that solves the a posteriori recursive mixed H^2/H^∞ estimation problem (with level γ) is given by the solution to the following convex quadratic program

$$\left\{ \begin{array}{lc} \min_{\hat{z}_{i|i}} & (\hat{z}_{i|i} - \bar{z}_{i|i})^*(\hat{z}_{i|i} - \bar{z}_{i|i}) \\ \text{subject to} & J_i + N_i \geq 0 \end{array} \right. , i = 0, \ldots N \quad (18)$$

where $\bar{z}_{i|i}$ and N_i are as in Theorem 2, part (a).
(b) An estimator $\hat{z}_i = \mathcal{F}_{p,i}\{y_0, y_1, \ldots y_{i-1}\}$ that solves the a priori recursive mixed H^2/H^∞ estimation problem (with level γ) is given by the solution to the following convex quadratic program

$$\left\{ \begin{array}{lc} \min_{\hat{z}_i} & (\hat{z}_i - \bar{z}_i)^*(\hat{z}_i - \bar{z}_i) \\ \text{subject to} & M_i \geq 0 \end{array} \right. , i = 0, \ldots N \quad (19)$$

where $\bar{z}_i$ and M_i are as in Theorem 2, part (b).

[1] The functions $\Psi_i(\hat{x}_i - \bar{x}_i, J_i)$ and $\Phi_i(\hat{x}_i - \bar{x}_i, M_i)$ represent the optimal cost-to-go, given the current state of the filter, $\hat{x}_i$, $\bar{x}_i$ and J_i.

The above solution has an interesting structure and effectively combines the H^2 and H^∞ solutions. The reason why the functions $\Psi_i(\cdot,\cdot)$ and $\Phi_i(\cdot,\cdot)$ do not appear in the solution is that recursive estimators attempt to achieve the smallest possible cost at each time instant and are therefore not concerned with the cost-to-go.

The nonlinear optimizations (18) and (19) are *convex* quadratic programs and can be readily solved using convex optimization techniques. Since the number of unknowns in these programs is q, the number of signals to be estimated, the complexity of the solution at each iteration is $O(q^3)$. As q is typically less than the number of states, n, the main computational burden at each iteration is in the propagation of the Riccati variables $\bar{P}_i$ and P_i which requires $O(n^3)$ operations. Thus the computational complexity of the recursive mixed H^2/H^∞ estimator of Theorem 3 is of the same order as that of the Kalman filter and (central) H^∞ filters.

When z_i is a scalar signal (*i.e.*, $q = 1$) one can solve the quadratic programs (18) and (19) in closed form and obtain a much more explicit form for the solution. This solution shows that, depending on the sign of the signals given in (20) and (25) below, the desired estimates are chosen either as the H^2 estimates, $\bar{z}_{i|i}$ and $\bar{z}_i$, or as a convex combination of the H^2 estimates and some other estimates (see (22) and (27)).

Lemma 1 (Solution in the Scalar Case) *Consider the setting of Theorem 3 where now z_i is a scalar signal and define*

$$\begin{aligned} \hat{x}_{i|i} &= \hat{x}_i + P_i H_i^*(I + H_i P_i H_i^*)^{-1}(y_i - H_i\hat{x}_i) \\ A_i &= \left[\gamma^2 I - L_i(P_i^{-1} + H_i^* H_i)^{-1} L_i^*\right]^{-1} \\ B_i &= \left[\gamma^2 I - L_i P_i L_i^*\right]^{-1}. \end{aligned}$$

(a) The convex quadratic program (18) has the following solution: If

$$J_i + \begin{bmatrix} y_i - H_i\hat{x}_i \\ \bar{z}_{i|i} - L_i\hat{x}_i \end{bmatrix}^* R_{e,i}^{-1} \begin{bmatrix} y_i - H_i\hat{x}_i \\ \bar{z}_{i|i} - L_i\hat{x}_i \end{bmatrix} \geq 0, \quad (20)$$

then

$$\hat{z}_{i|i} = \bar{z}_{i|i}. \quad (21)$$

Otherwise,

$$\hat{z}_{i|i} = \theta_{f,i}\bar{z}_{i|i} + (1 - \theta_{f,i})L_i\hat{x}_{i|i}, \quad (22)$$

where

$$\theta_{f,i} = \sqrt{\frac{a_i}{A_i|\bar{z}_{i|i} - L_i\hat{x}_{i|i}|^2}} < 1, \quad (23)$$

and

$$a_i = J_i + (y_i - H_i\hat{x}_i)^*(I + H_i P_i H_i^*)^{-1}(y_i - H_i\hat{x}_i). \quad (24)$$

(b) The convex quadratic program (19) has the following solution: If

$$J_i - (\bar{z}_i - L_i\hat{x}_i)^* B_i(\bar{z}_i - L_i\hat{x}_i) \geq 0, \quad (25)$$

then

$$\hat{z}_i = \bar{z}_i. \tag{26}$$

Otherwise,

$$\hat{z}_i = \theta_{p,i}\bar{z}_i + (1-\theta_{p,i})L_i\hat{x}_i, \tag{27}$$

where

$$\theta_{p,i} = \sqrt{\frac{J_i}{B_i|\bar{z}_i - L_i\hat{x}_i|^2}} < 1. \tag{28}$$

5. Simple Example

We now consider a simple example arising from adaptive filtering, where the goal is to use past and current observations to predict the next output. In order to compare the properties of the various estimators discussed in this paper, we shall use the central H^∞ filter, the optimal linear mixed H^2/H^∞ filter (computed using the method of [8]), and the nonlinear recursive mixed H^2/H^∞ filter of Theorem 3, for various values of γ. [Note that we have not yet been able to explicitly construct the functions $\Psi_i(\cdot,\cdot)$ and $\Phi_i(\cdot,\cdot)$, and so cannot use the optimal solution of Theorem 2.]

The results are given in Fig. 1 where we have plotted the expected estimation error energy as a function of the maximum energy gain, γ, for each of the three aforementioned estimators.[2]

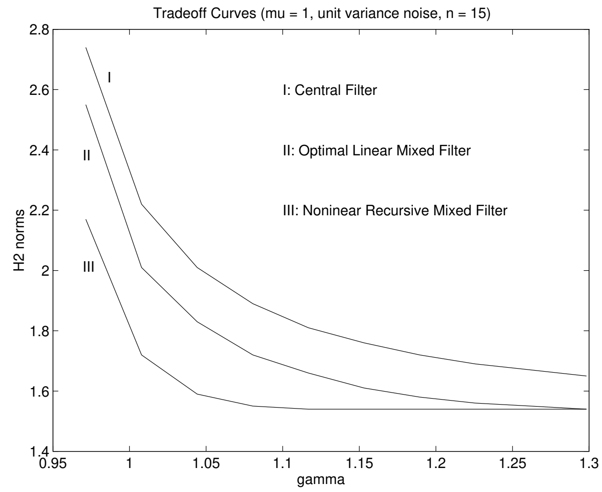

Figure 1: *Expected estimation error energy as a function of maximum energy gain for, (I) the central H^∞ filters, (II) the optimal linear mixed H^2/H^∞ filters, and, (III) the nonlinear recursive mixed H^2/H^∞ filters. [The horizon is $N = 15$.]*

As expected, for each value of γ, the optimal linear mixed H^2/H^∞ filter has an expected estimation error energy that is less than that of the central solution. What is perhaps surprising is that the nonlinear recursive mixed H^2/H^∞ filter outperforms the best linear filter for each value of γ. Since the recursive filter is suboptimal (over the set of nonlinear filters) it would be interesting to see how much further the optimal nonlinear mixed H^2/H^∞ filters of Theorem 2 would reduce the expected estimation error energy.

[2] The horizon has been taken as $N = 15$ since the computation required for finding the optimal linear H^2/H^∞ filter becomes prohibitively large as the horizon increases.

6. Conclusion

In this paper we studied the problem of mixed H^2/H^∞ estimation using nonlinear filters. We essentially show that nonlinear filters offer an improvement over linear filters in minimizing the expected estimation error energy over a given maximum energy gain, and moreover, that in the finite-horizon case, the optimal nonlinear mixed H^2/H^∞ estimator has finite-dimensional state-space structure. The solution involves a nonlinear program at each iteration which we have not been able to solve. We also considered a related suboptimal solution, with the property of being recursive, and that involves a convex quadratic program at each iteration. Also, though not treated here, similar results hold for the closely-related mixed H^2/H^∞ control problem.

References

[1] M. Green and D.J.N. Limebeer. *Linear Robust Control* Prentice Hall, Englewood Cliffs NJ, 1995.

[2] D. S. Bernstein and W. M. Haddad. LQG Control with an H^∞ Performance Bound: A Riccati Equation Approach. *IEEE Transactions on Automatic Control*, vol.34, no.3, pp. 293-305, 1989.

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

[4] D. J. N. Limebeer and B. D. O. Anderson. A Nash game approach to mixed H^2/H^∞ control. *IEEE Transactions on Automatic Control*, vol.39, no.1, pp. 69-82, 1994.

[5] K. Zhou, K. Glover, B. Bodenheimer and J. C. Doyle. Mixed H^2 and H^∞ Performance Objectives I: Robust Performance Analysis. *IEEE Transactions on Automatic Control*, vol.39, no.8, pp. 1564-1574, 1994.

[6] A. Megretski. On the order of controllers in mixed H^2/H^∞ control. In *Proc. IEEE Conference on Decision and Control*, Lake Buena Vista, FL, Dec. 1994, pp. 3173-3174.

[7] C. Foias, A. E. Frazho and W. S. Li, On H^2 Minimization for the Caratheodory-Schur Interpolation Problem, *Integral Equations and Operator Theory*, vol.21, no.1, pp. 24-32, 1995.

[8] B. Halder, B. Hassibi and T. Kailath, Mixed H^2/H^∞ estimation: preliminary analytic characterizations and a numerical solution. In the *Proceedings of the 1996 IFAC World Congress*, San Francisco, CA, July 1996.