

MIMO Linear Equalization With an H^∞ Criterion

Babak Hassibi, Alper T. Erdogan, and Thomas Kailath

Abstract—In this paper, we study the problem of linearly equalizing the multiple-input multiple-output (MIMO) communications channels from an H^∞ point of view. H^∞ estimation theory has been recently introduced as a method for designing filters that have acceptable performance in the face of model uncertainty and lack of statistical information on the exogenous signals. In this paper, we obtain a closed-form solution to the square MIMO linear H^∞ equalization problem and parameterize all possible H^∞ -optimal equalizers. In particular, we show that, for minimum phase channels, a scaled version of the zero-forcing equalizer is H^∞ -optimal. The results also indicate an interesting dichotomy between minimum phase and nonminimum phase channels: for minimum phase channels the best causal equalizer performs as well as the best noncausal equalizer, whereas for nonminimum phase channels, causal equalizers cannot reduce the estimation error bounds from their a priori values. Our analysis also suggests certain remedies in the nonminimum phase case, namely, allowing for finite delay, oversampling, or using multiple sensors. For example, we show that H^∞ equalization of nonminimum phase channels requires a time delay of at least l units, where l is the number of nonminimum phase zeros of the channel.

Index Terms— H^∞ estimation, linear equalization, minimum phase channels, multiple-input multiple-output (MIMO) equalization, risk sensitive estimation, robustness.

I. INTRODUCTION

IN THE area of equalization of communication channels, various criteria and the corresponding algorithms have been proposed to recover transmitted data from their filtered and noise corrupted versions. Each method has its own advantages and disadvantages in terms of performance and complexity [1]. There has been some recent work addressing the robustness issue in equalization [7], [8], [12], [13].

Equalization is closely related to estimation theory since the equalization problem can be considered as a special case of an estimation problem. Recently, in estimation theory, a (so-called) H^∞ approach has been proposed and extensively studied (see, e.g., [2]–[6] and the references therein) with the belief that the resulting H^∞ estimators will be more robust with respect to model uncertainties and lack of statistical knowledge of the exogenous signals. Due to its aforementioned properties, the H^∞ approach has also been proposed as an alternative method for channel equalization [7], [12], [13]. Further study of this approach for the multiple-input multiple-output (MIMO) linear

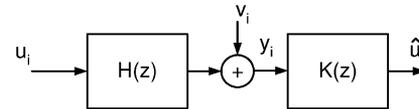


Fig. 1. The linear equalization problem.

equalization problem is the main focus of this paper. As we shall shortly see, the results obtained in this attempt provide us with a new and different perspective for the understanding and analysis of the equalization problem, as well as for H^∞ estimation itself.

The remainder of this paper is organized as follows. In Section II, we introduce the MIMO linear equalization problem. Section III introduces the H^∞ framework for robust estimation and shows how the H^∞ estimation problem can be solved via a certain J-spectral factorization. The description of the H^∞ equalization problem and the smoothing solution are provided in Section IV. Section V contains the main results of this paper. The MIMO H^∞ equalization problem is solved for the square case and all solutions are parameterized. The derivation hinges on the explicit J-spectral factorization of Section V-A1 and shows that the solution depends on whether the channel is minimum phase or not. It is also shown that, for minimum phase channels, appropriately scaled zero-forcing equalizers are H^∞ -optimal. Moreover, the relative merits of various H^∞ -optimal equalizers are discussed. Section V-B deals with the nonsquare case and Section VI studies some remedies for nonminimum phase channels. In particular, it is shown that successful equalization of nonminimum phase channels requires a time delay greater than or equal to the number of nonminimum phase zeros. In Section VII, we provide a comparative simulation example for MIMO linear equalization. This paper concludes in Section VIII.

II. LINEAR EQUALIZATION PROBLEM

The linear equalization problem that will be studied in this paper is depicted in Fig. 1. In this framework $\{u_i\}$ is an unknown sequence (the transmitted sequence), $\{v_i\}$ is an unknown additive disturbance sequence, $\{y_i\}$ is a known observations sequence, and $H(z)$ is a known linear time-invariant (LTI) communications channel. The goal is to design the LTI system $K(z)$ (the so-called equalizer) so as to estimate the unknown transmitted sequence $\{u_i\}$ from the known observations sequence $\{y_i\}$, where the estimated sequence is denoted by $\{\hat{u}_i\}$.

We shall assume that the transmitted vectors u_i are arbitrary vectors in \mathcal{C}^m . The unknown additive disturbance $\{v_i\}$ is typically composed of three components: i) measurement noise, ii) interference from other signals, and iii) modeling errors, due to imperfect knowledge of the true channel. We shall therefore, for the most part, not make any statistical assumptions on the

Manuscript received July 12, 2004; revised April 13, 2005. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Nicholas D. Sidiropoulos.

B. Hassibi is with the California Institute of Technology, Pasadena, CA 91125 USA (e-mail: hassibi@caltech.edu).

A. T. Erdogan is with Koc University, Istanbul, Turkey (e-mail: alperdogan@ku.edu.tr).

T. Kailath is with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: profkailath@yahoo.com).

Digital Object Identifier 10.1109/TSP.2005.861888

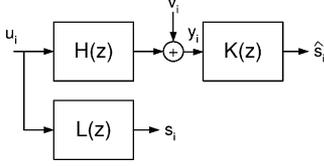


Fig. 2. A general estimation problem.

disturbance sequence $\{v_i\}$ and will simply consider it as an unknown sequence of elements in \mathcal{C}^p . The observations $\{y_i\}$ will therefore also be a sequence of elements in \mathcal{C}^p . The dimensions of $\{u_i\}$ and $\{y_i\}$ determine the input/output dimensions of the channel $H(z)$. The channel is assumed to be causal and stable, which means that it is a $p \times m$ matrix function in z with Laurent series expansion

$$H(z) = h_0 + h_1 z^{-1} + h_2 z^{-2} + \dots \quad (1)$$

that is analytic on and outside the unit circle, $|z| = 1$, where $\{h_i; i \geq 0\}$ denotes the impulse response of $H(z)$. The channel $H(z)$ is often assumed to be rational (of McMillan degree, say, n), or in fact finite impulse response (FIR) of length, say, n . However, almost all our results extend to the nonrational case, and we shall therefore not make this restriction. The equalizer $K(z)$ is assumed to be linear time-invariant since the channel is so, and since $u_i \in \mathcal{C}^m$. Although we shall assume that $K(z)$ is stable, we shall not always assume that it is causal.

III. H^∞ ESTIMATION

A general estimation problem is shown in Fig. 2, where $\{u_i \in \mathcal{C}^m\}$ is an unknown input sequence, $\{v_i \in \mathcal{C}^p\}$ is an unknown additive disturbance sequence, $\{y_i \in \mathcal{C}^p\}$ is a known measurement sequence, and $H(z)$ and $L(z)$ are known causal and stable linear time-invariant systems. The goal is to construct the linear time-invariant $K(z)$ (called the estimator) to estimate the unobservable desired sequence $\{s_i \in \mathcal{C}^q\}$ from the observations $\{y_i\}$. The estimates are denoted by $\{\hat{s}_i\}$. Clearly, the equalization problem is a special case of this formulation with $L(z) = I_m$. The behavior of any estimator $K(z)$ can be captured by the induced transfer matrix, say, $\mathcal{T}_K(z)$, that maps the unknown disturbances $\{u_i\}$ and $\{v_i\}$ to the estimation errors $\{s_i - \hat{s}_i\}$:

$$\mathcal{T}_K : \left\{ \begin{bmatrix} u_i \\ v_i \end{bmatrix} \right\} \rightarrow \{s_i - \hat{s}_i\}. \quad (2)$$

Using Fig. 2, we readily see

$$\mathcal{T}_K(z) = [L(z) - K(z)H(z) \quad -K(z)]. \quad (3)$$

The H^∞ problem can be formulated as follows.

Problem 1: (Optimal H^∞ Estimation Problem): Find a causal estimator $K(z)$ that satisfies

$$\inf_{K(z)} \|\mathcal{T}_K(z)\|_\infty = \inf_{K(z)} \|[L(z) - K(z)H(z) - K(z)]\|_\infty. \quad (4)$$

Moreover, find the resulting $\gamma_{\text{opt}} \triangleq \inf_{K(z)} \|\mathcal{T}_K(z)\|_\infty$.

The approaches to solve problem can be found in various references, e.g., see [17], [2], and [4] or references therein. Our approach in this paper would be through use of the solution to the suboptimal H^∞ problem as outlined by the following theorem [17].

Theorem 1: (Solution to Suboptimal H^∞ Problem): A $K(z)$ that achieves

$$\|\mathcal{T}_K(z)\|_\infty < \gamma \quad (5)$$

exists if, and only if, the Popov function

$$\Sigma(z) = \begin{bmatrix} I + H(z)H^*(z^{-*}) & -H(z)L^{-*}(z^{-*}) \\ -L(z)H^*(z^{-*}) & -\gamma^2 I + L(z)L^*(z^{-*}) \end{bmatrix} \quad (6)$$

admits a canonical J-spectral factorization of the form

$$\Sigma(z) = \begin{bmatrix} L_{11}(z) & L_{12}(z) \\ L_{21}(z) & L_{22}(z) \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \times \begin{bmatrix} L_{11}^*(z^{-*}) & L_{21}^*(z^{-*}) \\ L_{12}^*(z^{-*}) & L_{22}^*(z^{-*}) \end{bmatrix} \quad (7)$$

with $\begin{bmatrix} L_{11}(z) & L_{12}(z) \\ L_{21}(z) & L_{22}(z) \end{bmatrix}$ and $L_{11}(z)$ causal and causally invertible, and $L_{12}(z)$ strictly causal. If this is the case, then all possible H^∞ estimators of level γ are given by

$$K(z) = (L_{22}(z)C(z) - L_{21}(z))(L_{11}(z) - L_{12}(z)C(z))^{-1} \quad (8)$$

where $C(z)$ is any causal and strictly contractive operator, i.e., $C(z)$ is causal and is such that $\bar{\sigma}(C(e^{j\omega})) < 1$, for all $\omega \in [0, 2\pi]$. An important special choice is $C = 0$ which leads to the so-called central filter

$$K_{\text{cen}}(z) = -L_{21}(z)L_{11}^{-1}(z). \quad (9)$$

IV. H^∞ EQUALIZATION

The equalization problem is a special case of the general estimation problem of Section III corresponding to $L(z) = I_m$. Thus in the H^∞ approach to equalization, the goal is to minimize $\|[I_m - K(z)H(z) - K(z)]\|_\infty$.

We should also mention that it is possible to slightly generalize the estimation problem of Section III by introducing certain weighting matrices. In this case, we are interested in minimizing the H^∞ norm of the transfer operator from $\{Q^{-1/2}u_i, R^{-1/2}v_i\}$ to $\{Q^{-1/2}(u_i - \hat{u}_i)\}$, which is seen to be

$$\begin{aligned} \mathcal{T}_{K^{\text{weighted}}}(z) &= [I_m - Q^{-1/2}K(z)H(z)Q^{1/2} \quad -Q^{-1/2}K(z)R^{1/2}] \\ & \quad (10) \end{aligned}$$

for some weighting matrices $R > 0$ and $Q > 0$. Defining $K'(z) = Q^{-1/2}K(z)R^{1/2}$ and $H'(z) = R^{-1/2}H(z)Q^{1/2}$, we see that we can rewrite $\mathcal{T}_{K^{\text{weighted}}}(z) = [I_m - K'(z)H'(z) - K'(z)]$. Thus there is no loss of generality in assuming $Q = I_m$ and $R = I_p$, and we shall do so throughout.

The inequality $\gamma_s \leq \gamma_{\text{opt}}$ is quite clear, since noncausal equalizers have access to more information than causal ones and cannot therefore have worse performance. The second inequality follows from the fact that if we perform no equalization, i.e., $K(z) = 0$, then we have

$$\|T_{K_{\text{weighted}}}\|_\infty = \|[I_m \ 0]\|_\infty = 1. \quad (11)$$

Thus $\gamma = 1$ is the worst case energy gain obtained from doing no equalization. Therefore, how much the optimal values of γ_s and γ_{opt} can be reduced from unity shows how successful we are in the equalization problem.

A. Noncausal Case

For the case where $K(z)$ is not constrained to be causal we have the following theorem [17].

Theorem 2: (Linear Smoothing Equalization): The solution to the problem

$$\min_{K(z)} \|[(z^{-d}I - K(z)H(z)) \ -K(z)]\|_\infty \quad (12)$$

is given by $\gamma_{\text{opt,smoothing}}^2 = \max_{\omega \in [0, 2\pi]} \sigma_{\max}((I + H^*(e^{j\omega})H(e^{j\omega}))^{-1})$, where d represents the equalization delay.

Moreover, one noncausal solution is given by the noncausal minimum mean squared error (MMSE) equalizer

$$K_{\text{smoothing}}(z) = z^{-d}H^*(z^{-*})(I + H(z)H^*(z^{-*}))^{-1} \quad (13)$$

which has the property that for all $K(z)$

$$T_{K_{\text{smoothing}}}(e^{j\omega})T_{K_{\text{smoothing}}}^*(e^{j\omega}) \leq T_K(e^{j\omega})T_K^*(e^{j\omega}), \quad \omega \in [0, 2\pi]. \quad (14)$$

Remarks:

- Focusing on the scalar case is instructive. In this case, we have

$$\gamma_s^2 = \sup_{\omega \in [0, 2\pi]} \frac{1}{1 + |H(e^{j\omega})|^2}. \quad (15)$$

- In other words, γ_s is achieved at those frequencies for which $|H(e^{j\omega})|^2$ takes its smallest value.
- Consider now the general case where $H(z)$ is a $p \times m$ transfer matrix. If $m > p$ (so that there are more signals to estimate than to observe), we have

$$\gamma_s = 1 \quad (16)$$

since $H^*(z^{-*})H(z)$ will be rank deficient. More importantly, in the case of a square channel ($p = m$), if the channel has a unit circle zero, then $\gamma_s = 1$ (since $H^*(e^{j\omega})H(e^{j\omega})$ will be rank deficient at the frequency where the zero occurs).

- Recall from our earlier discussions that $\gamma_s = 1$ implies that we have no improvement over not equalizing at all ($K(z) = 0$). This is quite clear when we have a unit circle zero, corresponding to frequency ω_1 , say. In this case the output of the channel cannot contain sinusoids of frequency ω_1 , and hence if $\{u_i\}$ is precisely such a sinusoid, then we cannot estimate it since we cannot observe it at the output of $H(z)$.

V. CAUSAL H^∞ EQUALIZATION

Our main results are concerned with the case where the equalizer is restricted to being causal.

A. The Square Case

In this section, we shall assume that $p = m$, i.e., that the number of input and observations signals coincide. We shall presently show that, in this case, we can find an explicit expression for the factorization (7) of Theorem 1. This is what allows us to, contrary to general H^∞ estimation problems, obtain a closed-form solution to the causal H^∞ equalization problem.

1) *J-Spectral Factorization in the Square Case:* Note that in the equalization problem the transfer matrix whose J -spectral factorization we are seeking is given by

$$\begin{bmatrix} I + H(z)H^*(z^{-*}) & -H(z) \\ -H^*(z^{-*}) & -\gamma^2 I + I \end{bmatrix}. \quad (17)$$

[We have taken $L(z) = I_m$ in (17).] Since we know $\gamma \leq 1$ is achievable, let us consider (17) for a given $\gamma_s < \gamma < 1$. Once γ is strictly less than one, we may perform the following block upper diagonal lower factorization:

$$\begin{aligned} & \begin{bmatrix} I + H(z)H^*(z^{-*}) & -H(z) \\ -H^*(z^{-*}) & -\gamma^2 I + I \end{bmatrix} \\ &= \begin{bmatrix} I & -\frac{H(z)}{1-\gamma^2} \\ 0 & I \end{bmatrix} \\ & \times \begin{bmatrix} I + H(z)H^*(z^*) - \frac{H(z)H^*(z^*)}{1-\gamma^2} & 0 \\ \underbrace{I - \frac{\gamma^2}{1-\gamma^2}H(z)H^*(z^{-*})}_{0} & I - \gamma^2 I \end{bmatrix} \\ & \times \begin{bmatrix} I & 0 \\ -\frac{H^*(z^{-*})}{1-\gamma^2} & I \end{bmatrix}. \end{aligned} \quad (18)$$

Now since $\gamma > \gamma_s$, we have

$$\begin{aligned} \gamma^2 I &> [I + H^*(e^{jw})H(e^{jw})]^{-1} \\ &\Rightarrow \gamma^{-2} I < I + H^*(e^{jw})H(e^{jw}) \\ &\Rightarrow \frac{\gamma^2}{1-\gamma^2} H(e^{jw})H^*(e^{jw}) > I \end{aligned}$$

where in the fourth step we have used our assumption that $\gamma_s < 1$, so that $H(e^{jw})$ is full rank for all w . The last inequality implies that we can introduce the canonical spectral factorization

$$\Delta(z)R_\Delta\Delta^*(z^{-*}) = \frac{\gamma^2}{1-\gamma^2}H(z)H^*(z^{-*}) - I > 0 \quad (19)$$

with $\Delta(z)$ monic ($\Delta(\infty) = I$) and causal and causally invertible, and $R_\Delta > 0$. With this definition we can write the factorization (18) as

$$\begin{aligned} & \begin{bmatrix} I + H(z)H^*(z^{-*}) & -H(z) \\ -H^*(z^{-*}) & -\gamma^2 I + I \end{bmatrix} \\ &= \begin{bmatrix} -\frac{H(z)}{\sqrt{1-\gamma^2}} & \Delta(z)R_\Delta^{1/2} \\ \sqrt{1-\gamma^2}I & 0 \end{bmatrix} \\ & \times \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} -\frac{H^*(z^{-*})}{\sqrt{1-\gamma^2}} & \sqrt{1-\gamma^2}I \\ R_\Delta^{*/2}\Delta^*(z^{-*}) & 0 \end{bmatrix}. \end{aligned}$$

Note that the above factorization is a J -spectral factorization since the spectral factor

$$P(z) \triangleq \begin{bmatrix} -\frac{H(z)}{\sqrt{1-\gamma^2}} & \Delta(z)R_\Delta^{1/2} \\ \sqrt{1-\gamma^2}I & 0 \end{bmatrix}. \quad (20)$$

is causal, and causally invertible, since

$$P^{-1}(z) = \begin{bmatrix} 0 & \frac{1}{\sqrt{1-\gamma^2}}I \\ R_\Delta^{-1/2}\Delta^{-1}(z) & \frac{R_\Delta^{-1/2}\Delta^{-1}(z)H(z)}{1-\gamma^2} \end{bmatrix}. \quad (21)$$

is causal. However, it is not quite the desired factorization of Theorem 1 since the block (1, 2) entry $P_{12}(z)$ is not strictly causal. To transform $P(z)$ into a transfer matrix such that the (1, 2) entry is strictly causal, we must premultiply $P(z)$ by an appropriate constant¹ J -unitary matrix Θ . In fact, we need only concentrate on $P(\infty)$. Thus, we must have

$$\begin{aligned} P(\infty)\Theta &= \begin{bmatrix} -\frac{h_0}{\sqrt{1-\gamma^2}} & R_\Delta^{1/2} \\ \sqrt{(1-\gamma^2)}I & 0 \end{bmatrix} \begin{bmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{bmatrix} \\ &= \begin{bmatrix} A & 0 \\ B & C \end{bmatrix}. \end{aligned} \quad (22)$$

Since Θ is J -unitary, we have

$$\begin{aligned} P(\infty)JP^*(\infty) &= \begin{bmatrix} A & 0 \\ B & C \end{bmatrix} J \begin{bmatrix} A^* & B^* \\ 0 & C^* \end{bmatrix} \\ J &= \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}. \end{aligned} \quad (23)$$

Comparing the (1, 1) block entries in the above equality, we have

$$\frac{h_0h_0^*}{1-\gamma^2} - R_\Delta = AA^* \quad (24)$$

which implies that $A = ((h_0h_0^*)(1-\gamma^2) - R_\Delta)^{1/2}$ and that $P_{12}(z)$ can be made strictly causal only if

$$\frac{h_0h_0^*}{1-\gamma^2} - R_\Delta > 0. \quad (25)$$

Comparing the (2,1) block entries of (23) yields $-h_0^* = BA^*$, so that $B = -h_0^*A^{-*}$. Finally, comparing the (2,2) block entries yields

$$(1-\gamma^2)I = BB^* - CC^* \quad (26)$$

or

$$\begin{aligned} CC^* &= -(1-\gamma^2)I + h_0^* \left(\frac{h_0h_0^*}{1-\gamma^2} - R_\Delta \right)^{-1} h_0 \\ &= (1-\gamma^2)^2 [h_0^*R_\Delta^{-1}h_0 - (1-\gamma^2)I]^{-1} \end{aligned}$$

so that $C = (1-\gamma^2)[h_0^*R_\Delta^{-1}h_0 - (1-\gamma^2)I]^{-*/2}$. We should also remark that (25) guarantees $h_0^*R_\Delta^{-1}h_0 - (1-\gamma^2)I > 0$ so that C

¹The J -unitary matrix must be constant so as not to destroy the causal and causal invertibility property of $P(z)$.

is well defined. Indeed, two different block lower diagonal upper and block upper diagonal lower factorizations of the matrix

$$N = \begin{bmatrix} (1-\gamma^2) & h_0^* \\ h_0 & R_\Delta \end{bmatrix}$$

show that the matrices

$$\begin{bmatrix} (1-\gamma^2)I & 0 \\ 0 & R_\Delta - \frac{h_0h_0^*}{1-\gamma^2} \end{bmatrix} \quad \begin{bmatrix} (1-\gamma^2)I - h_0^*R_\Delta^{-1}h_0 & 0 \\ 0 & R_\Delta \end{bmatrix}$$

are congruent and therefore must have the same inertia. Since $(1-\gamma^2)I > 0$ and $R_\Delta - (h_0h_0^*)/(1-\gamma^2) < 0$, we conclude that N has m positive eigenvalues and m negative eigenvalues. Moreover, since $R_\Delta > 0$, we conclude that $(1-\gamma^2)I - h_0^*R_\Delta^{-1}h_0 < 0$, which is the desired result.

Having found the elements $\{A, B, C\}$, we can now find the desired J -unitary matrix

$$\Theta = \begin{bmatrix} -\frac{h_0}{\sqrt{1-\gamma^2}} & R_\Delta^{1/2} \\ \sqrt{(1-\gamma^2)}I & 0 \end{bmatrix}^{-1} \begin{bmatrix} A & 0 \\ B & C \end{bmatrix}$$

and thereby the desired canonical factor

$$\begin{bmatrix} L_{11}(z) & L_{12}(z) \\ L_{21}(z) & L_{22}(z) \end{bmatrix} = \begin{bmatrix} P_{11}(z) & P_{12}(z) \\ P_{21}(z) & P_{22}(z) \end{bmatrix} \Theta. \quad (27)$$

The results of this section are summarized in the following lemma.

Lemma 1— J -Spectral Factorization: Suppose the $m \times m$ causal matrix function $H(z) = h_0 + h_1z^{-1} + \dots$ has full rank on $|z| = 1$, and let γ be such that

$$1 > \gamma^2 > \sup_{w \in [0, 2\pi]} \bar{\sigma} [I + H^*(e^{jw})H(e^{jw})]^{-1}. \quad (28)$$

Then a factorization of the form

$$\begin{aligned} &\begin{bmatrix} I + H(z)H^*(z^{-*}) & -H(z) \\ H^*(z^{-*}) & -\gamma^2I + I \end{bmatrix} \\ &= \begin{bmatrix} L_{11}(z) & L_{12}(z) \\ L_{21}(z) & L_{22}(z) \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \\ &\quad \times \begin{bmatrix} L_{11}^*(z^{-*}) & L_{21}^*(z^{-*}) \\ L_{12}^*(z^{-*}) & L_{22}^*(z^{-*}) \end{bmatrix} \end{aligned} \quad (29)$$

with $\begin{bmatrix} L_{11}(z) & L_{12}(z) \\ L_{21}(z) & L_{22}(z) \end{bmatrix}$ causal and causally invertible, and $L_{12}(z)$ strictly causal, exists if, and only if

$$\frac{h_0h_0^*}{1-\gamma^2} - R_\Delta > 0 \quad (30)$$

where R_Δ is found from the canonical spectral factorization

$$\Delta(z)R_\Delta\Delta^*(z^{-*}) = \frac{\gamma^2}{1-\gamma^2}H(z)H^*(z^{-*}) - I > 0 \quad (31)$$

with $\Delta(z)$ monic and causal and causally invertible. If this is the case, then we have (32) as shown at the bottom of the page, where A and C are found from the factorizations

$$\begin{aligned} AA^* &= \frac{h_0 h_0^*}{1 - \gamma^2} - R_\Delta \\ CC^* &= (1 - \gamma^2)^2 [h_0^* R_\Delta^{-1} h_0 - (1 - \gamma^2)I]^{-1}. \end{aligned} \quad (33)$$

From the above lemma, and using Theorem 1, it follows that a level- γ H^∞ equalizer exists if, and only if, (30) is satisfied and

$$\begin{aligned} L_{11}^{-1}(z) \\ = (1 - \gamma^2)A^* [H(z)h_0^* - (1 - \gamma^2)\Delta(z)R_\Delta]^{-1} \end{aligned} \quad \text{is causal.} \quad (34)$$

Thus our problem is reduced to checking (30) and (34). This may appear as a difficult task; however, we shall show in the following sections that the satisfaction of these conditions is strictly dependent upon whether $H(z)$ is minimum phase or not.

2) *The Minimum Phase Case:* Let us first assume that the $m \times m$ causal matrix function $H(z)$ is minimum phase, i.e., that its inverse $H^{-1}(z)$ is causal. Then we have the following result.

Lemma 2: Consider the setting of Lemma 1 and suppose that $H(z)$ is minimum phase. Then for all γ satisfying (28) we have

$$\frac{h_0 h_0^*}{1 - \gamma^2} - R_\Delta > 0. \quad (35)$$

Proof: Note that from (31)

$$\Delta(e^{jw})R_\Delta \Delta^*(e^{jw}) = \frac{\gamma^2}{1 - \gamma^2} H(e^{jw})H^*(e^{jw}) - 1.$$

Thus

$$\begin{aligned} (1 - \gamma^2)\Delta(e^{jw})R_\Delta \Delta^*(e^{jw}) \\ = \gamma^2 H(e^{jw})H^*(e^{jw}) - (1 - \gamma^2) < H(e^{jw})H^*(e^{jw}) \\ \Rightarrow H^{-1}(e^{jw})\Delta(e^{jw})h_0 h_0^{-1} (1 - \gamma^2)R_\Delta h_0^{-*} h_0^* \Delta^* \\ \times (e^{jw})H^{-*}(e^{jw}) < I. \end{aligned}$$

Now the transfer matrix $M(z) \triangleq H^{-1}(z)\Delta(z)h_0$ is causal (since $H^{-1}(z)$ and $\Delta(z)$ are causal) and monic (since $M(\infty) = H^{-1}(\infty)\Delta(\infty)h_0 = h_0^{-1} \cdot I \cdot h_0 = I$), and we may write

$$M(z) = I + m_1 z^{-1} + m_2 z^{-2} + \dots$$

Therefore

$$\begin{aligned} (I + m_1 e^{-jw} + \dots)h_0^{-1}(1 - \gamma^2)R_\Delta h_0^{-*} \\ \times (I + m_1 e^{-jw} + \dots)^* < I \\ \Rightarrow \frac{1}{2\pi} \int_0^{2\pi} (I + m_1 e^{-jw} + \dots)h_0^{-1}(1 - \gamma^2)R_\Delta h_0^{-*} \\ \times (I + m_1 e^{-jw} + \dots)^* dw < \frac{1}{2\pi} \int_0^{2\pi} I dw \\ \Rightarrow h_0^{-1}(1 - \gamma^2)R_\Delta h_0^{-*} + m_1 h_0^{-1}(1 - \gamma^2) \\ \times R_\Delta h_0^{-*} m_1^* + \dots < I \Rightarrow (1 - \gamma^2)R_\Delta < h_0 h_0^*. \end{aligned}$$

The next result establishes the causal invertibility of $L_{11}(z)$.

Lemma 3: Consider the setting of Lemma 1 and suppose that $H(z)$ is minimum phase. Then for all γ satisfying (28) the transfer matrix

$$L_{11}^{-1}(z) = (1 - \gamma^2)A^* [H(z)h_0^* - (1 - \gamma^2)\Delta(z)R_\Delta]^{-1} \quad (36)$$

is causal.

Proof: Let us first write

$$\begin{aligned} L_{11}^{-1}(z) \\ = (1 - \gamma^2)A^* h_0^{-*} [I - (1 - \gamma^2)H^{-1}(z)\Delta(z)R_\Delta h_0^{-*}]^{-1} \\ \times H^{-1}(z) \end{aligned}$$

where we have used the fact that h_0 is invertible since $H(z)$ is minimum phase (if h_0 is singular, then $H(z)$ will have a zero at $z = \infty$). Now since $H^{-1}(z)$ is causal, $L_{11}^{-1}(z)$ will be causal if $[I - (1 - \gamma^2)H^{-1}(z)\Delta(z)R_\Delta h_0^{-*}]^{-1}$ is causal. We shall presently show this by appealing to Lemma 7 in the Appendix and showing that

$$A(z) \triangleq (1 - \gamma^2)H^{-1}(z)\Delta(z)R_\Delta h_0^{-*} \quad (37)$$

is a strict contraction. Indeed

$$\begin{aligned} A(e^{jw})A^*(e^{jw}) \\ = (1 - \gamma^2)H^{-1}(e^{jw})\Delta(e^{jw})R_\Delta \underbrace{h_0^{-*} h_0^{-1}}_{< \frac{1}{1 - \gamma^2} R_\Delta^{-1}} R_\Delta \Delta^* \\ \times (e^{jw})H^{-*}(e^{jw})(1 - \gamma^2) \\ < H^{-1}(e^{jw})\Delta(e^{jw})(1 - \gamma^2)R_\Delta \Delta^*(e^{jw})H^{-*}(e^{jw}) \\ < I \end{aligned}$$

where in the second step we used $h_0 h_0^* > (1 - \gamma^2)R_\Delta$ so that $h_0^{-*} h_0^{-1} < (1)/(1 - \gamma^2)R_\Delta^{-1}$ and where in the last step we used (36).

The above lemmas show that, in the minimum phase case, whenever a J -spectral factorization exists, then $L_{12}(z)$ can be made strictly causal and $L_{11}(z)$ can be made causally invertible. This essentially means that in the minimum phase case, we have

$$\gamma_{\text{opt}} = \gamma_s.$$

3) *The Nonminimum Phase Case:* We now assume that the $m \times m$ causal matrix function $H(z)$ is nonminimum phase, i.e., that its inverse $H^{-1}(z)$ is noncausal. Here we have the following result.

Lemma 4: Consider the setting of Lemma 1 and suppose that $H(z)$ is nonminimum phase, and that

$$\frac{h_0 h_0^*}{1 - \gamma^2} - R_\Delta > 0. \quad (38)$$

Then the transfer matrix

$$L_{11}^{-1}(z) = (1 - \gamma^2)A^* [H(z)h_0^* - (1 - \gamma^2)\Delta(z)R_\Delta]^{-1} \quad (39)$$

is noncausal.

$$\begin{bmatrix} L_{11}(z) & L_{12}(z) \\ L_{21}(z) & L_{22}(z) \end{bmatrix} = \begin{bmatrix} [H(z)h_0^* - (1 - \gamma^2)\Delta(z)R_\Delta] \frac{A^{-*}}{1 - \gamma^2} & \frac{-H(z) + \Delta(z)h_0}{1 - \gamma^2} C \\ -h_0^* A^{-*} & C \end{bmatrix} \quad (32)$$

Proof: First note that we may write

$$L_{11}^{-1}(z) = -A^* R_{\Delta}^{-*/2} \left[I - R_{\Delta}^{-1/2} \Delta^{-1}(z) H(z) h_0^* R_{\Delta}^{-*/2} \right]^{-1} \times R_{\Delta}^{-1/2} \Delta^{-1}(z).$$

Now when (38) holds, using an argument similar to the one presented in the proof of Lemma 3, one can show that the causal transfer matrix

$$C(z) \triangleq R_{\Delta}^{-1/2} \Delta^{-1}(z) H(z) h_0^* R_{\Delta}^{-*/2} \quad (40)$$

is a strict expansion, i.e., $C(e^{jw})C(e^{jw}) > I, \forall w$. Note, moreover, that since $H^{-1}(z)$ is noncausal and

$$H^{-1}(z) = h_0^* R_{\Delta}^{-*/2} C^{-1}(z) R_{\Delta}^{-1/2} \Delta^{-1}(z)$$

the transfer matrix $C^{-1}(z)$ is noncausal as well. (Otherwise, the noncausal $H^{-1}(z)$ would be the product of causal transfer matrices.) Lemma 8 now implies that, since $C(z)$ is a causal expansion and $C^{-1}(z)$ is noncausal, the transfer matrix $(I - C(z))^{-1}$ must be noncausal. This in turn implies that $L_{11}^{-1}(z)$ is noncausal, since

$$(I - C(z))^{-1} = -R_{\Delta}^{*/2} A^{-*} L_{11}^{-1}(z) \Delta(z) R_{\Delta}^{1/2}$$

cannot be the product of causal transfer matrices.

Recall that for a level- γ H^{∞} equalizer, with $\gamma < 1$, to exist we require both (30) and (34). What we have just shown is that, when $H(z)$ is nonminimum phase, even if (30) is satisfied (34) is not. This essentially means that in the nonminimum phase case, we have

$$\gamma_{\text{opt}} = 1 \quad (41)$$

since $\gamma = 1$ is always achievable by $K(z) = 0$.

4) *Main Result:* We are now in a position to state the main result of this paper.

Theorem 3 (H^{∞} Equalization in the Square Case): Consider the $m \times m$ causal transfer matrix $H(z) = h_0 + h_1 z^{-1} + h_2 z^{-2} + \dots$ and suppose that we want to solve the problem

$$\min_{\text{causal } K(\cdot)} \| [I - K(z)H(z) \quad -K(z)] \|_{\infty} \triangleq \gamma_{\text{opt}}.$$

1) If $H(z)$ is minimum phase, i.e., if $H^{-1}(z)$ is analytic on $|z| \geq 1$, then the minimax energy gain is given by

$$\gamma_{\text{opt}}^2 = \sup_{w \in [0, 2\pi]} \bar{\sigma} [I + H^*(e^{jw})H(e^{jw})]^{-1}. \quad (42)$$

Moreover, for any $\gamma > \gamma_{\text{opt}}$, all causal equalizers that guarantee $\| [I - K(z)H(z) \quad -K(z)] \|_{\infty}^2 \leq \gamma^2$ are given by

$$K(z) = (L_{22}(z)S(z) - L_{21}(z))(L_{11}(z) - L_{12}(z)S(z))^{-1} \quad (43)$$

where the $L_{ij}(z)$ are given as shown in (44) at the bottom of the page, with the monic and minimum phase transfer matrix $\Delta(z)$ and the matrix R_{Δ} found from the canonical spectral factorization

$$\Delta(z)R_{\Delta}\Delta^*(z^{-*}) = \frac{\gamma^2}{1-\gamma^2} H(z)H^*(z^{-*}) - I > 0 \quad (45)$$

where the matrices $\{A, C\}$ are found from the factorizations

$$\begin{aligned} AA^* &= \frac{h_0 h_0^*}{1-\gamma^2} - R_{\Delta} \\ CC^* &= (1-\gamma^2)^2 [h_0^* R_{\Delta}^{-1} h_0 - (1-\gamma^2)I]^{-1} \end{aligned} \quad (46)$$

and where $S(z)$ is any causal contraction, i.e., $S(z)$ is analytic on $|z| \geq 1$ and

$$S^*(z^{-*})S(z) \leq I, \quad \forall |z| = 1. \quad (47)$$

2) If $H(z)$ is nonminimum phase, i.e., if $H^{-1}(z)$ is not analytic on $|z| \geq 1$, then the minimax energy gain is given by

$$\gamma_{\text{opt}} = 1. \quad (48)$$

Remarks:

- i) Note that when $H(z)$ is minimum phase, $\gamma_{\text{opt}} = \gamma_s$! This implies that for minimum phase channels, causal equalizers perform as well as noncausal ones, and that (from an H^{∞} point of view) there is no gain in knowing future values of the observations signal $\{y_i\}$.
- ii) Recall from Theorem 2 that the central noncausal equalizer at least outperforms all other equalizers (causal or noncausal) at all frequencies. Theorem 3 therefore states that if $H(z)$ is minimum phase, then the central noncausal equalizer cannot outperform the best causal equalizer in terms of its worst case performance (which occurs at certain frequencies).
- iii) However, if $H(z)$ is nonminimum phase, then $\gamma_{\text{opt}} = 1$! Thus causal equalization of nonminimum phase channels is not possible, since $\gamma = 1$ is the same bound obtained by not equalizing at all ($K(z) = 0$).
- iv) Suppose now that $H(z)$ is rational. Then if all the zeros of $H(z)$ lie strictly inside the unit circle, causal equalizers perform as good as noncausal ones. As soon as a zero lies on the unit circle then equalization is not possible ($\gamma_s = \gamma_{\text{opt}} = 1$) since the channel cannot pass certain frequencies. However, as soon as a zero crosses the unit circle, noncausal equalizers can now perform whereas causal ones cannot. There is therefore a clear transition occurring at the unit circle and a strict dichotomy between minimum phase and nonminimum phase channels as far as causal equalization is concerned.
- v) Some understanding of the result of Theorem 3 for nonminimum phase channels can be obtained by considering the simple (perhaps the simplest) nonminimum phase channel $H(z) = z^{-1}$ (a pure delay). In this case it is quite clear that causal equalization is not possible, since at time i the equalizer has access only to noisy observations of $\{u_j\}_{j=-\infty}^{i-1}$ and has no knowledge of u_i in order to be able to estimate it.
- vi) A similar behavior to Theorem 3 can be observed had we studied the equalization problem from an H^2 point

$$\begin{bmatrix} L_{11}(z) & L_{12}(z) \\ L_{21}(z) & L_{22}(z) \end{bmatrix} = \begin{bmatrix} [H(z)h_0^* - (1-\gamma^2)\Delta(z)R_{\Delta}] \frac{A^{-*}}{1-\gamma^2} & \frac{-H(z)+\Delta(z)h_0}{1-\gamma^2} C \\ -h_0^* A^{-*} & C \end{bmatrix} \quad (44)$$

of view, though the result is not as pronounced. In the scalar case (which for simplicity we shall only consider), the H^2 norm of the H^2 -optimal causal equalizer is given by

$$d_2 = 1 - \frac{|h_0|^2}{R_e} \quad (49)$$

where R_e is found from the spectral factorization

$$1 + H(z)H^*(z^{-*}) = M^*(z^{-*})R_eM(z) \quad (50)$$

with $M(z)$ monic and minimum phase. It is now easy to show that for all channels $H_1(z)$ and $H_2(z)$ that have the same spectrum (i.e., $1 + H_1(z)H_1^*(z^{-*}) = 1 + H_2(z)H_2^*(z^{-*})$), so that noncausal equalizers have the same performance, $|h_0|^2$ is largest for the minimum phase channel. Thus d_2 is smallest for the minimum phase channel, which means that the corresponding equalizer has the best H^2 performance. (In fact, it can also be shown that d_2 increases as the number of non-minimum phase zeros increases.)

5) *Special Cases:* Theorem 3 gives a full parameterization of all possible level γ H^∞ equalizers in terms of a causal strict contraction $S(z)$. The most natural choice is $S(z) = 0$, which corresponds to the central equalizer

$$K_{\text{cen}}(z) = (1 - \gamma^2)h_0^* [H(z)h_0^* - (1 - \gamma)^2\Delta(z)R_\Delta]^{-1}. \quad (51)$$

As mentioned earlier, the central equalizer has various other desirable optimality properties, such as being risk-sensitive optimal [10] and being the maximum entropy solution [9], [11]. We shall not go into the details here and shall just mention that the risk-sensitive optimality property of minimizing

$$E \exp \left[\frac{1}{2\gamma^2} (u_i - \hat{u}_i)^*(u_i - \hat{u}_i) \right] \quad (52)$$

may be useful for digital communications since exponentially larger penalties are being applied to larger estimation errors, which is typically where detection errors (detecting a zero as a one, say) occur. This observation is studied in more detail in [12] with mixed results.

A less obvious, but nonetheless intriguing, choice is

$$S(z) = -(1 - \gamma^2)C^{-1}h_0^{-1}R_\Delta A^{-*} \quad (53)$$

a constant matrix. Indeed $S(z)$ is a strict contraction since we have the equation at the bottom of the page so that

$$SS^* = (1 - \gamma^2)C^{-1}h_0^{-1}R_\Delta A^{-*} A^{-1}R_\Delta h_0^{-*} C^{-*} (1 - \gamma^2) < I.$$

With this choice of $S(z)$, (43) shows that we obtain

$$K(z) = (1 - \gamma^2)H^{-1}(z) \quad (54)$$

i.e., one H^∞ optimal equalizer is simply a scaled version of the inverse of the channel. In other words, suitably scaling the zero-forcing equalizer is H^∞ -optimal.

Lemma 5 (Zero-Forcing Equalizer): For any γ such that

$$1 > \gamma^2 > \sup_{w \in [0, 2\pi]} \bar{\sigma}[I + H^*(e^{jw})h(e^{jw})]^{-1} \quad (55)$$

[(55) implies that $H(z)$ is minimum phase] the equalizer $K(z) = (1 - \gamma^2)H^{-1}(z)$ achieves

$$\|T_K(z)\|_\infty = \|[I - K(z)H(z) \quad -K(z)]\|_\infty < \gamma. \quad (56)$$

Proof: We have already presented a proof in the arguments leading to the statement of the lemma. A more direct proof is also possible. Note that with $K(z) = (1 - \gamma^2)H^{-1}(z)$, we have

$$T_K(z) = [\gamma^2 I \quad -(1 - \gamma^2)H^{-1}(z)].$$

Therefore

$$\begin{aligned} T_K(e^{jw})T_K^*(e^{jw}) &= \gamma^4 I + (1 - \gamma^2)^2 H^{-1}(e^{jw})H^{-*}(e^{jw}) \\ &< \gamma^4 I + (1 - \gamma^2)^2 \cdot \frac{\gamma^2}{1 - \gamma^2} I = \gamma^2 I \end{aligned}$$

where in the second step we have used $H^*(e^{jw})H(e^{jw}) > (1 - \gamma^2)/(\gamma^2)I$, so that $H^{-1}(e^{jw})H^{-*}(e^{jw}) < (\gamma^2)/(1 - \gamma^2)I$.

Note that if there were no additive disturbance present ($\{v_i\} = \{0\}$), the inverse of the channel would perfectly reconstruct the input signal. The scaling factor $1 - \gamma^2$, though, is crucial here—if $H(z)$ has zeros close to the unit circle the frequency response of $H(z)$ will be very large at certain frequencies, resulting in large amplification of the noise at those frequencies. However, in this case γ_{opt} will be close to one, and thus the factor $1 - \gamma^2$ prohibits such large amplifications.

Nonetheless, even though it is H^∞ -optimal, scaling and inverting the plant is not a very satisfactory solution since it has poor average performance and leads to flat error spectra.

$$\begin{aligned} CC^* &= (1 - \gamma^2)^2 [h_0^* R_\Delta^{-1} h_0 - (1 - \gamma^2)I]^{-1} \\ &= (1 - \gamma^2)^2 \left[h_0^{-1} R_\Delta h_0^{-*} - h_0^{-1} R_\Delta h_0^{-*} \left[h_0^{-1} R_\Delta h_0^{-*} - \frac{1}{1 - \gamma^2} I \right]^{-1} h_0^{-1} R_\Delta h_0^{-*} \right] \\ &= (1 - \gamma^2)^2 h_0^{-1} R_\Delta h_0^{-*} + (1 - \gamma^2)^2 h_0^{-1} R_\Delta A^{-*} A^{-1} R_\Delta h_0^{-*} \\ &< (1 - \gamma^2)^2 h_0^{-1} R_\Delta A^{-*} A^{-1} R_\Delta h_0^{-*} \end{aligned}$$

To illustrate this fact, and to compare the various equalizers considered so far, let us consider the problem of equalizing a scalar minimum phase channel $H(z)$. In Fig. 3, we have plotted the error spectra $|T_K(e^{j\omega})|^2$, corresponding to the noncausal H^2 (smoothing or noncausal central H^∞) optimal equalizer, the causal H^2 -optimal equalizer, the causal central H^∞ -optimal equalizer, and the causal H^∞ -optimal (scaled) zero-forcing equalizer, for a minimum phase channel. Reviewing this figure should allow the reader to gain some perspective and “feel” for the relative performances of these different equalizers.

As can be seen, the H^2 - and H^∞ -optimal noncausal equalizer outperforms all other equalizers at all frequencies. Its error spectrum thus serves as a lower bound for the error spectra of all other equalizers. The H^2 -optimal equalizer has the best average performance (under the appropriate stochastic assumptions), and hence has the smallest area under the spectrum curve, among all causal equalizers. The (central) H^∞ -optimal and the zero-forcing H^∞ -optimal equalizers have the best worst case performance, and hence their spectra has the smallest peak, over all causal equalizers. Since the channel is minimum phase, these peaks coincide with the peak of the noncausal H^∞ -optimal equalizer.

These figures also indicate the tradeoffs between H^2 and H^∞ equalizers. Although H^2 -optimal equalizers have the smallest area under the spectrum curve, they tend to have large peaks that correspond to nonrobust behavior at certain frequencies. On the other hand, although the H^∞ -optimal estimators are robust, and have the smallest possible peaks, they have larger areas under their spectrum curve and may have poor average performance. This is particularly true of the scaled zero-forcing equalizer, which has a very “flat” spectrum. However, the central H^∞ equalizer has a reasonably low error spectrum connected with its risk-sensitive optimality.

B. The Nonsquare Case

So far we have only considered the case of a square channel ($p = m$). When $p \neq m$, we have the following result.

Theorem 4 (Causal H^∞ Equalizer for Nonsquare Plant): Consider the $p \times m$ causal transfer matrix $H(z) = h_0 + h_1 z^{-1} + h_2 z^{-2} + \dots$, where $p \neq m$, and where we want to solve the problem

$$\min_{\text{causal } K(\cdot)} \|[I - K(z)H(z) \quad -K(z)]\|_\infty \triangleq \gamma_{\text{opt}}.$$

- i) Suppose $p > m$. Then if $H(z)$ has no zeros on or outside the unit circle

$$\gamma_{\text{opt}} < 1. \quad (57)$$

Otherwise $\gamma_{\text{opt}} = 1$.

- ii) Suppose $p < m$. Then

$$\gamma_{\text{opt}} = 1. \quad (58)$$

Proof: To prove part i), let us first assume that $H(z)$ has no zeros on $|z| \geq 1$. Since $p > m$, this means that $H(z)$ has

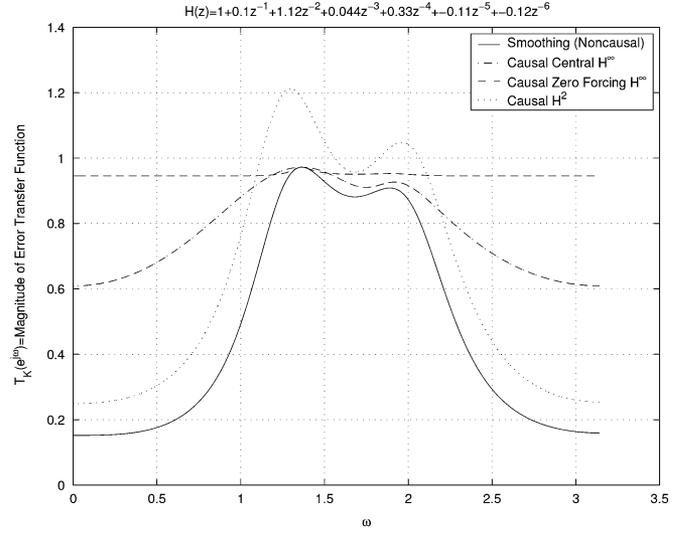


Fig. 3. Inversion of the noisy channel $H(z) = 1 + .1z^{-1} + 1.12z^{-2} + .044z^{-3} + .33z^{-4} - .11z^{-5} - .12z^{-6}$.

a causal left inverse, i.e., there exists an $m \times p$ transfer matrix $H^{-L}(z)$, analytic on $|z| \geq 1$, such that

$$H^{-L}(z)H(z) = I_m.$$

Since $H^{-L}(z)$ is analytic on $|z| \geq 1$, for any $\lambda > \|H^{-L}(z)\|_\infty^2$, we can write

$$H^{-L}(e^{j\omega})H^{-L*}(e^{j\omega}) < \lambda I, \quad \forall \omega.$$

Now let us choose our equalizer as $K(z) = 1/(1 + \lambda) \cdot H^{-L}(z)$. For this choice we have

$$T_K(z) = \begin{bmatrix} \frac{\lambda}{1+\lambda} \cdot I & -\frac{1}{1+\lambda} \cdot H^{-L}(z) \end{bmatrix}.$$

Thus

$$\begin{aligned} T_K(e^{j\omega})T_K^*(e^{j\omega}) &= \frac{\lambda^2}{(1+\lambda)^2} I + \frac{1}{(1+\lambda)^2} H^{-L}(e^{j\omega})H^{-L*}(e^{j\omega}) \\ &< \frac{\lambda^2}{(1+\lambda)^2} I + \frac{\lambda}{(1+\lambda)^2} I = \frac{\lambda}{1+\lambda} I < I \end{aligned}$$

so that $\|T_K(z)\|_\infty < 1$.

Let us now assume that $H(z)$ has zeros on $|z| \geq 1$. We shall assume that there exists a causal $K(z)$ such that $\|T_K(z)\|_\infty < 1$ and shall obtain a contradiction. Thus assume that there exists a $K(z)$, analytic on $|z| \geq 1$, such that

$$(I - K(e^{j\omega})H(e^{j\omega}))(I - K(e^{j\omega})H(e^{j\omega}))^* + K(e^{j\omega})K^*(e^{j\omega}) < I, \quad \forall \omega.$$

Since $T_K(z)$ is analytic on $|z| \geq 1$, the maximum-modulus principle implies that

$$(I - K(z)H(z))(I - K(z)H(z))^* + K(z)K^*(z) < I, \quad \forall |z| \geq 1. \quad (59)$$

But assume that $H(z)$ has a zero in $|z| \geq 1$, i.e., $H(w)$ has rank less than m for some w such that $|w| \geq 1$. Let X be a corresponding right null vector of $H(w)$, i.e., $H(w)X = 0$. Then we have $[I - K(w)H(w)]X = X$, so that unity is an eigenvalue of $I - K(w)H(w)$, and hence

$$\bar{\sigma}[I - K(w)H(w)] \geq 1$$

which clearly contradicts (59).

Finally, to prove ii), we note that when $p < m$, the number of signals to estimated is greater than the number of observations, and so there is no hope in estimating them. $K(z) = 0$ clearly gives $\gamma_{\text{opt}} = 1$.

Remark: Note that when $p > m$, $H(z)$ will generically have no zeros because it will generically have full rank for all z . To be more explicit, suppose that $p = 2$ and $m = 1$, so that

$$H(z) = [H_1(z) \ H_2(z)]^T. \quad (60)$$

Now $H(z)$ will have a zero outside the unit circle if, and only if, $H_1(z)$ and $H_2(z)$ share some nonminimum phase zero. But of course any two arbitrary rational functions will generically not have common zeros.

We should also mention that, when $H(z)$ has no zeros in $|z| \geq 1$, finding an explicit expression for γ_{opt} does not appear to be simple. The construction presented in the proof of Theorem 4 suggests the following upper bound:

$$\gamma_{\text{opt}}^2 \leq \inf_{H^{-L}(\cdot), H^{-L}(z)H(z)=I_m} \frac{\|H^{-L}(z)\|_\infty^2}{1 + \|H^{-L}(z)\|_\infty^2}. \quad (61)$$

Various other statements can also be made. Here is an example. Suppose that $H(z) = \begin{bmatrix} H_1(z) \\ H_2(z) \end{bmatrix}$, where both $H_1(z)$ and $H_2(z)$ are scalar and minimum phase. Then the best causal equalizer performs better than the best noncausal equalizers applied to $H_1(z)$ and $H_2(z)$ separately, but not as good as the best noncausal equalizer applied to $H_1(z)$ and $H_2(z)$ together. In other words

$$\begin{aligned} & \sup_{w \in [-\pi, \pi]} \frac{1}{1 + |H_1(e^{jw})|^2 + |H_2(e^{jw})|^2} \\ & \leq \gamma_{\text{opt}}^2 \leq \min_{i=1,2} \sup_{w \in [-\pi, \pi]} \frac{1}{1 + |H_i(e^{jw})|^2}. \end{aligned} \quad (62)$$

VI. SOME REMEDIES FOR NONMINIMUM PHASE CHANNELS

In Section V-A4, we noted that, in the square case ($p = m$), causal equalization of nonminimum phase channels is not possible (from the H^∞ point of view). Since nonminimum phase channels do occur in practice, it is important to have a means of circumventing this drawback. In this section, we shall look at three remedies for the nonminimum phase case: using multiple sensors, oversampling, or allowing for finite delay, as suggested by the framework of this paper.

A. Using Multiple Sensors

By adding multiple sensors, we increase the number of observation signals without changing the number of input signals.

Thus, we transform a square problem with $p = m$ to a nonsquare problem with $p > m$. This means that our original problem that had $\gamma_{\text{opt}} = 1$, when $H(z)$ is nonminimum phase, is transformed to one for which γ_{opt} is generically strictly less than one.

Take, for example, the scalar case and suppose that $H(z)$ is nonminimum phase. Then if we add a second sensor, corresponding to a scalar channel with transfer function $H'(z)$, our channel will become

$$[H(z) \ H'(z)]^T.$$

Thus, even though $H(z)$ is nonminimum phase, and $H'(z)$, too, may very well be nonminimum phase, it is still possible to causally equalize the aggregate channel as long as $H(z)$ and $H'(z)$ share no nonminimum phase zeros. This, of course, generically holds.

We should also mention that the property of noncommon zeros is similar to the requirement that arises in the blind equalization of vector channels using second-order statistics (see, e.g., [14]).

B. Oversampling

Oversampling is also currently used as a method for the blind equalization of nonminimum phase channels using second-order statistics (see, e.g., [15]), and can also be used for causal equalization. The idea is to sample the output of the communications channel at a rate higher than the rate of the input signal. This then leads us to a situation where we have more observation signals than input signals, which can be suitably transformed to an MIMO channel scenario with $p > m$ that was considered in Section VI-A. Indeed if we have a nonminimum phase scalar channel with transfer function $H(z)$, and oversample at a rate of r output samples per each input transmitted, the resulting MIMO channel becomes

$$[H(z) \ H(e^{j\frac{2\pi}{r}}z) \ \dots \ H(e^{j\frac{2(r-1)\pi}{r}}z)]^T. \quad (63)$$

Thus the condition for causal H^∞ equalization is that the transfer functions $\{H(e^{j(2k\pi)/r}z)\}_{k=0}^{r-1}$ share no common nonminimum phase zero.

C. Finite Delay

Another obvious remedy for the equalization of nonminimum phase channels is to allow for some finite delay, $d > 0$. In other words, we are interested in estimating u_i using the observations $\{\dots y_{i+d-1}, y_{i+d}\}$. Such a problem corresponds to the general estimation problem of Section III with $L(z) = z^{-d}$ (since we can also think of it as estimating u_{i-d} using the observations $\{\dots y_{i-1}, y_i\}$).

It is quite clear that allowing for infinite delay takes us back to the situation of a noncausal equalizer. The natural question to ask therefore is: what is the amount of finite delay $0 < d < \infty$ that we should incur so that:

- i) $\gamma_{\text{opt}} < 1$?
- ii) $\gamma_{\text{opt}} = \gamma_s$?

The answer to the first question would tell how much delay we need in order to be successful in causal equalization, and the answer to the second question would tell us how much delay we

need to in order obtain the same performance as the noncausal equalizer.

Unfortunately, we do not yet know an explicit factorization for the transfer matrix

$$\begin{bmatrix} I + H(z)H^*(z^{-*}) & -z^d H(z) \\ -z^{-d} H^*(z^{-*}) & (1 - \gamma^2)I \end{bmatrix} \quad (64)$$

for arbitrary $d > 0$ and general nonminimum phase channels. Therefore we do not yet have explicit expressions for γ_{opt} in this case and have no answer for question ii), raised above, of when $\gamma_{\text{opt}} = \gamma_s$. However, we do have the answer to question i), of when $\gamma_{\text{opt}} = 1$.

Lemma 6 (Delay Required for $\gamma_{\text{opt}} < 1$): Consider the $m \times m$ causal rational transfer matrix $H(z) = h_0 + h_1 z^{-1} + \dots$ and suppose that $H(z)$ has no unit circle zeros (i.e., $H(e^{jw})$ is full rank for all w). We are interested in the following problem:

$$\min_{\text{causal } K(z)} \|[z^{-d}I - K(z)H(z) \quad -K(z)]\|_{\infty} \triangleq \gamma_{\text{opt}}, \quad d \geq 0. \quad (65)$$

Let l denote the number of nonminimum phase (i.e., on or outside the unit circle) zeros of $H(z)$ (counting multiplicities). Then we have the following result.

i) If $d \geq l$, then

$$\gamma_{\text{opt}} < 1.$$

ii) If $d < l$, then

$$\gamma_{\text{opt}} = 1.$$

Proof: The proof to the above theorem is based on the operator theoretic techniques outlined in [16]. A proof for the more general $L(z)$ is given in [16]. However, we will provide the proof specifically for the delay case, i.e., $L(z) = z^{-d}$ following the same methodology.

We first define the Toeplitz operators \mathcal{L}_- for $L(z) = z^{-d}$, the delay operator, and \mathcal{H}_- for the channel $H(z)$ as

$$\mathcal{L}_- = [I \quad 0_{\infty \times dM}] \quad (66)$$

$$\mathcal{H}_- = \begin{bmatrix} \ddots & & & & \\ \ddots & H_0 & 0 & 0 & \\ \dots & H_1 & H_0 & 0 & \\ \dots & H_2 & H_1 & H_0 & \end{bmatrix} \quad (67)$$

respectively. We know that

$$\gamma_{\text{opt,delay}} = \mathcal{L}_-(I + \mathcal{H}_-^* \mathcal{H}_-)^{-1} \mathcal{L}_-^*. \quad (68)$$

Here $\gamma > \gamma_{\text{opt,delay}}$, if and only if

$$\begin{aligned} \gamma^2 - \mathcal{L}_-(I + \mathcal{H}_-^* \mathcal{H}_-)^{-1} \mathcal{L}_-^* &> 0 \\ \Leftrightarrow \begin{bmatrix} I + \mathcal{H}_-^* \mathcal{H}_- & \mathcal{L}_-^* \\ \mathcal{L}_- & \gamma^2 I \end{bmatrix} &> 0 \\ \Leftrightarrow I + \mathcal{H}_- \mathcal{H}_-^* - \gamma^{-2} \mathcal{L}_-^* \mathcal{L}_- &> 0 \\ \Leftrightarrow x^*(I + \mathcal{H}_- \mathcal{H}_-^* - \gamma^{-2} \mathcal{L}_-^* \mathcal{L}_-)x &> 0, \quad \text{for all } x \neq 0 \in l_-^2 \\ \Leftrightarrow \|x\|_2^2 + \|\mathcal{H}_- x\|_2^2 - \gamma^{-2} \|\mathcal{L}_- x\|_2^2 &> 0, \quad \text{for all } x \neq 0 \in l_-^2 \\ \Leftrightarrow \gamma^2 > \frac{\|\mathcal{L}_- x\|_2^2}{\|x\|_2^2 + \|\mathcal{H}_- x\|_2^2}, &\text{for all } x \neq 0 \in l_-^2. \end{aligned}$$

Therefore, we can write

$$\gamma_{\text{opt,delay}}^2 = \sup_{x \neq 0 \in l_-^2} \frac{\|\mathcal{L}_- x\|_2^2}{\|x\|_2^2 + \|\mathcal{H}_- x\|_2^2}. \quad (69)$$

Therefore we can write

$$\begin{aligned} \gamma_{\text{opt,delay}} &= 1 \\ \Leftrightarrow \frac{\|\mathcal{L}_- x\|_2^2}{\|x\|_2^2 + \|\mathcal{H}_- x\|_2^2} &= 1, \text{ for some } x \\ \Leftrightarrow \|\mathcal{L}_- x\|_2^2 - \|x\|_2^2 &= \|\mathcal{H}_- x\|_2^2 \\ \Leftrightarrow (\|x\|_2^2 - (x_0^2 + \dots + x_{-d+1}^2) - \|x\|_2^2) &= \|\mathcal{H}_- x\|_2^2 \\ \Leftrightarrow -(x_0^2 + \dots + x_{-d+1}^2) &= \|\mathcal{H}_- x\|_2^2. \end{aligned}$$

As a result, $\gamma_{\text{opt,delay}} = 1$ if, and only if, there exists an x which is in the null space of \mathcal{H}_- and which has $x_0 = x_{-1} = \dots = x_{-d+1} = 0$. Now suppose that $H(z)$ has l distinct nonminimum phase zeros at z_1, \dots, z_l . Since the nullspace vectors of \mathcal{H}_- has the form

$$v_{p_i} = [\dots \quad z_i^3 \quad z_i^2 \quad z_i \quad 1]^T \quad (70)$$

we can write this condition to be equivalent to finding $\alpha_1, \alpha_2, \dots, \alpha_l$ such that

$$\alpha_1 v_{p_1} + \dots + \alpha_l v_{p_l} = \begin{bmatrix} x' \\ 0_d \end{bmatrix} \quad (71)$$

where 0_d is the all-zero vector with dimension $d \times 1$. An equivalent condition can be written as finding $[\alpha_1, \dots, \alpha_l]$ such that

$$\underbrace{\begin{bmatrix} \vdots & \vdots & \vdots & \vdots \\ p_1^2 & p_2^2 & \dots & p_l^2 \\ p_1 & p_2 & \dots & p_l \\ 1 & 1 & \dots & 1 \end{bmatrix}}_P \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_l \end{bmatrix}}_{\alpha} = 0_d. \quad (72)$$

Since P is a matrix with dimensions $d \times l$, its nullspace is nonempty and therefore $\gamma_{\text{opt,delay}} = 1$ if and only if $d < l$.

The above lemma has an interesting interpretation. In order to get an improvement over not equalizing at all ($\gamma_{\text{opt}} = 1$), the delay d should be chosen greater than the number of nonminimum phase zeros of the channel. Thus the fundamental limitation in causal equalization is the number of zeros outside of the unit circle.

VII. EXAMPLE

In this section, we provide an example simulation, where $H(z)$ is a five-tap 5×4 channel with randomly generated (Gaussian) entries. All four users use four quaternary phase-shift keying constellation. We simulated average bit error rate (BER) performance for the average case with white disturbances and the worst case with the disturbances at the frequency where maximum error gain occurs. Both equalizers are implemented as recursive Kalman filters (H^2 or H^∞) [17] where equalization delay is set as $d = 3$. Fig. 4 shows BER

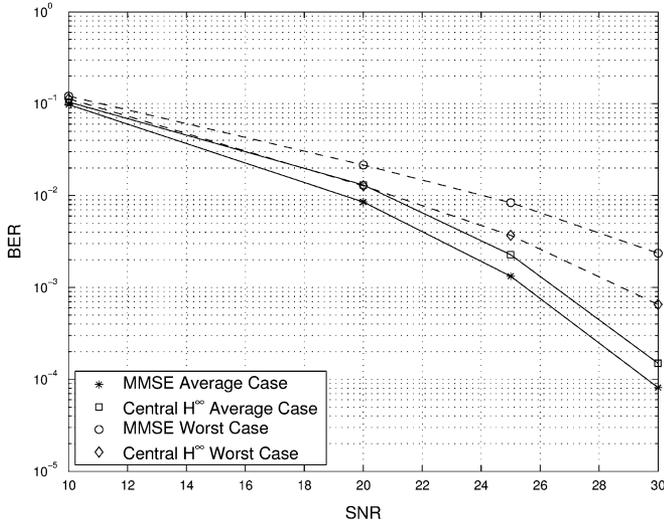


Fig. 4. Average and worst case BER performances of MMSE and the central H^∞ equalizers.

rate curves (averaged over all users). According to this figure, although the risk sensitive H^∞ equalizer has a slightly worse average performance than the MMSE equalizer, its worst case performance is much closer to its average performance in comparison to the MMSE equalizer.

VIII. CONCLUSION

In this paper, we studied the H^∞ criterion as an alternative method for the equalization of communication channels. Current equalization methods and algorithms are mainly concerned with average behavior (e.g., MSE or BER) under the conditions of perfect models and known statistical distributions. H^∞ theory, on the other hand, is concerned with the worst case behavior and the question of robustness with respect to model uncertainty and lack of statistical information.

We obtained a closed-form solution to the H^∞ equalization problem and parameterized all possible H^∞ -optimal equalizers for the square causal case. More importantly, not only did we obtain a solution to the problem, but we learned a lot about the equalization problem itself. In particular, we discovered an interesting dichotomy between minimum phase and nonminimum phase channels. In the minimum phase case causal equalizers perform as well as noncausal ones, whereas in the nonminimum phase case, causal equalizers cannot reduce the worst case estimation error bounds from their *a priori* values. Moreover, we showed that, for minimum phase channels, a scaled version of the zero-forcing equalizer is H^∞ -optimal. We also studied the use of adding more sensors, oversampling, and allowing for finite delay, as remedies for the nonminimum phase case. We also learned that the fundamental limit in causally equalizing a nonminimum phase channel is related to its number of nonminimum phase zeros.

The results presented in this paper also have ramifications in other areas. They essentially deal with the problem of “robustly” inverting a linear time-invariant system (as in filter banks) and so have implications to the important tracking problem of con-

trol theory. They also have various implications to more general control and estimation problems and, especially, to the questions of worst case controllability and estimability (essentially what systems are easy, or difficult, to control and estimate) [16].

Finally, we may mention some possible directions for future work. The first has to do with the extension of our explicit derivations for square case for more general case with non-square channel and arbitrary equalization delay. In connection with this, the question of how much delay is required to obtain the same H^∞ performance as the noncausal solution is to be addressed. Another has to do with incorporating “structural” modeling errors for the channel in the equalizer solution. Perhaps the most important direction relies on the following observation: As we saw for H^∞ -optimal zero-forcing equalization, having a good (indeed optimal) worst case performance alone is not sufficient for having acceptable performance in real applications. Since the H^∞ criterion provides a family of filters that achieve the same worst case performance, it is an important design problem to choose a filter among this family that has good performance with respect to other relevant criteria. This question was partially addressed (through the study of mixed $H^2 - H^\infty$ and risk sensitive criteria) for FIR equalizers in [18] and for filter banks in [19].

APPENDIX

Some important lemmas follow.

Lemma 7: Suppose the $n \times n$ transfer matrix $A(z)$ is causal and strictly contractive, i.e.,

- i) $A(z)$ is analytic on $|z| \geq 1$;
- ii) $A(z)A^*(z^{-*}) < I_n$ for $|z| = 1$.

Then the transfer matrix $(I_n - A(z))^{-1}$ is causal.

Proof: Since $A(z)$ is a contraction, this implies that $A(e^{jw})A^*(e^{jw}) < I_n, \forall w$. Moreover, since $A(z)$ is analytic on $|z| \geq 1$, due to the maximum modulus principle, it achieves its maximum on the boundary. Thus

$$A(z)A^*(z) < I_n, \quad \forall |z| \geq 1. \quad (\text{A.1})$$

This means that we can expand the inverse in $(I_n - A(z))^{-1}$ to write

$$(I_n - A(z))^{-1} = I_n + A(z) + A^2(z) + \dots \quad (\text{A.2})$$

where the infinite series converges absolutely for $|z| \geq 1$ since $A(z)$ is a strict contraction in this region. Thus, $(I_n - A(z))^{-1}$ is the sum of absolutely converging analytic functions on $|z| \geq 1$, and hence itself is an analytic function on $|z| \geq 1$. Thus, $(I_n - A(z))^{-1}$ is causal.

Lemma 8: Suppose that the $n \times n$ transfer matrix $C(z)$ is a causal strict expansion, i.e., that:

- i) $C(z)$ is analytic on $|z| \geq 1$;
- ii) $C(z)C^*(z^{-*}) > I_n$ for $|z| = 1$.

Then we have

$$C^{-1}(z) \text{ noncausal} \Rightarrow (I_n - C(z))^{-1} \text{ noncausal.}$$

Proof: We shall prove the lemma by contradiction, i.e., we shall assume that $(I_n - C(z))^{-1}$ is causal and prove that $C^{-1}(z)$ is causal. First note that $C(e^{jw})C^*(e^{jw}) > I_n$ implies

$$\begin{aligned} & (I_n - C(e^{jw}))^{-1}C(e^{jw})C^*(e^{jw})(I_n - C(e^{jw}))^{-*} \\ & > (I_n - C(e^{jw}))^{-1}(I_n - C(e^{jw}))^{-*} \\ & \Rightarrow (I_n - C(e^{jw}))^{-1}[I_n - (I_n - C(e^{jw}))] \\ & \quad \times [I_n - (I_n - C(e^{jw}))^*](I_n - C(e^{jw}))^{-*} \\ & > (I_n - C(e^{jw}))^{-1}(I_n - C(e^{jw}))^{-*} \\ & \Rightarrow \underbrace{[I_n - (I_n - C(e^{jw}))^{-1}]}_{\triangleq B(e^{jw})} \\ & \quad + [I_n - (I_n - C(e^{jw}))^{-1}]^* > I_n > 0. \end{aligned}$$

We thus have

$$\begin{cases} B(z) + B^*(z^{-*}) > 0, & \forall |z| = 1 \\ B(z) \text{ is causal} \end{cases}.$$

In other words, $B(z)$ is causal and positive real. Lemma 9 therefore implies that $B^{-1}(z)$ is causal. But since

$$\begin{aligned} C^{-1}(z) &= [I_n - (I_n - C(z))]^{-1} \\ &= I_n - [I_n - (I_n - C(z))^{-1}]^{-1} \\ &= I_n - B^{-1}(z) \end{aligned}$$

we conclude that $C^{-1}(z)$ is causal.

Lemma 9: Suppose that the $n \times n$ transfer matrix $B(z)$ is causal and positive real, i.e.,

- i) $B(z)$ is analytic on $|z| \geq 1$;
- ii) $B(z) + B^*(z^{-*}) > 0$ for $|z| = 1$.

Then $B^{-1}(z)$ is causal.

Proof: Since $B(z)$ is analytic on $|z| \geq 1$, then $B(e^{jw})$ is bounded for all $w \in [0, 2\pi]$ and, in particular

$$\sup_{w \in [0, 2\pi]} \bar{\sigma}[B(e^{jw})B^*(e^{jw})] = \|B(z)\|_\infty^2 < \infty.$$

Moreover, since $B(z)$ is positive real, there exists some $\alpha > 0$ such that

$$\begin{aligned} B(e^{jw}) + B^*(e^{jw}) &> \alpha \|B(z)\|_\infty^2 I \geq \alpha B(e^{jw})B^*(e^{jw}), \\ &\forall w \in [0, 2\pi]. \end{aligned}$$

Therefore the causal transfer matrix $I - \alpha B(z)$ is a strict contraction since

$$\begin{aligned} & (I - \alpha B(e^{jw}))(I - \alpha B(e^{jw}))^* \\ &= I - \alpha(B(e^{jw}) + B^*(e^{jw})) + \alpha^2 B(e^{jw})B^*(e^{jw}) \\ &= I - \alpha \left[\underbrace{B(e^{jw}) + B^*(e^{jw}) - \alpha B(e^{jw})B^*(e^{jw})}_{> 0} \right] < I. \end{aligned}$$

Therefore, using Lemma 7, $B^{-1}(z)$ is causal since $B(z) = (1/\alpha) \cdot [I - (I - \alpha B(z))]$.

REFERENCES

- [1] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.
- [2] P. P. Khargonekar and K. M. Nagpal, "Filtering and smoothing in an H^∞ setting," *IEEE Trans. Autom. Control*, vol. 36, pp. 151–166, 1991.
- [3] T. Basar, "Optimum performance levels for minimax filters, predictors and smoothers," *Syst. Contr. Lett.*, vol. 16, pp. 309–317, 1991.
- [4] M. J. Grimble, "Polynomial matrix solution of the H^∞ filtering problem and the relationship to Riccati equation state-space results," *IEEE Trans. Signal Process.*, vol. 41, no. 1, pp. 67–81, Jan. 1993.
- [5] M. Green and D. J. N. Limebeer, *Linear Robust Control*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [6] B. Hassibi, A. H. Sayed, and T. Kailath, "Linear estimation in Krein spaces—Parts I and II: Theory and applications," *IEEE Trans. Autom. Control*, vol. 41, no. 1, pp. 18–49, Jan. 1996.
- [7] S. C. Peng, "An equalizer design for nonminimum phase channel via two-block H^∞ optimization technique," *Signal Process.*, vol. 51, pp. 1–13, 1996.
- [8] W. Fan and V. Balakrishnan, "Robust steady-state filtering for systems with deterministic and stochastic uncertainties," *IEEE Trans. Signal Process.*, vol. 51, pp. 3227–3231, Oct. 2003.
- [9] D. Mustafa and K. Glover, *Minimum Entropy H^∞ Control*. New York: Springer-Verlag, 1990.
- [10] P. Whittle, *Risk-Sensitive Optimal Control*. New York: Wiley, 1990.
- [11] K. Glover and J. C. Doyle, "State-space formulae for all stabilizing controllers that satisfy an H^∞ -norm bound and relations to risk sensitivity," *Syst. Contr. Lett.*, vol. 11, pp. 167–172, 1988.
- [12] A. T. Erdogan, B. Hassibi, and T. Kailath, "On linear H^∞ equalization of communication channels," *IEEE Trans. Signal Process.*, vol. 48, pp. 3227–3231, 2000.
- [13] —, "MIMO decision feedback equalization from an H^∞ perspective," *IEEE Trans. Signal Process.*, vol. 52, pp. 734–745, 2004.
- [14] B. Khalaj, B. Hassibi, A. Paulraj, and T. Kailath, "Blind identification of FIR channels using antenna arrays," in *Proc. 27th Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Nov. 1993, pp. 721–726.
- [15] L. Tong, G. Xu, and T. Kailath, "A new approach to blind identification and equalization of multipath channels," in *Proc. 25th Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Nov. 1991.
- [16] B. Hassibi and T. Kailath, "On optimal solutions to two block H^∞ problems," in *Proc. Amer. Control Conf.*, vol. 3, 1998, pp. 24–26.
- [17] B. Hassibi, A. H. Sayed, and T. Kailath, *Indefinite Quadratic Estimation and Control: A Unified Approach to H^2 and H^∞ Theories*. New York: SIAM Studies in Applied Mathematics, 1998.
- [18] A. T. Erdogan, B. Hassibi, and T. Kailath, "FIR H^∞ equalization," *Signal Process.*, vol. 81, pp. 907–917, May 2001.
- [19] H. Vikalo, B. Hassibi, A. T. Erdogan, and T. Kailath, "On robust signal reconstruction in noisy filter banks," *Signal Process.*, vol. 85, pp. 1–14, May 2005.



Babak Hassibi was born in Tehran, Iran, in 1967. He received the B.S. degree from the University of Tehran in 1989 and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, in 1993 and 1996, respectively, all in electrical engineering.

From October 1996 to October 1998 he was a Research Associate with the Information Systems Laboratory, Stanford University, and from November 1998 to December 2000 he was a Member of the Technical Staff in the Mathematical Sciences Research Center, Bell Laboratories, Murray Hill,

NJ. Since January 2001, he has been an Assistant Professor of electrical engineering with the California Institute of Technology, Pasadena. He has also held short-term appointments with Ricoh California Research Center, the Indian Institute of Science, and Linköping University, Sweden. His research interests include wireless communications, robust estimation and control, adaptive signal processing, and linear algebra. He is the coauthor of *Indefinite Quadratic Estimation and Control: A Unified Approach to H^2 and H^∞ Theories* (New York: SIAM, 1999) and *Linear Estimation* (Englewood Cliffs, NJ: Prentice-Hall, 2000).

Prof. Hassibi received the 1999 O. Hugo Schuck best paper award from the American Automatic Control Council, the 2002 National Science Foundation Career Award, and the 2003 Okawa Foundation Research Grant for Information and Telecommunications.



Alper T. Erdogan was born in Ankara, Turkey, in 1971. He received the B.S. degree from the Middle East Technical University, Ankara, in 1993 and the M.S. and Ph.D. degrees from Stanford University, CA, in 1995 and 1999, respectively.

He was a Principal Research Engineer with Globespan-Virata Corporation (formerly Excess Bandwidth and Virata Corporations) from September 1999 to November 2001. He has been an Assistant Professor in the Electrical and Electronics Engineering Department, Koc University, Istanbul,

Turkey, since January 2002. His research interests include wireless and wireline communications, adaptive signal processing, optimization, system theory and control, and information theory.



Thomas Kailath received the B.E. (Telecom) degree from the College of Engineering, Pune, India, in 1956 and the S.M. and Sc.D., degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1959 and 1961, respectively.

From 1961 to 1962, he was with the Communications Research Division, Jet Propulsion Laboratories, Pasadena, CA, and taught part-time at the California Institute of Technology. Since then he has been with Stanford University, where he is currently Hitachi

America Professor of Engineering, Emeritus. He has also held shorter term appointments at several institutions around the world. His research has spanned a large number of disciplines, emphasizing information theory and communications in the 1960s, linear systems, estimation, and control in the 1970s, VLSI design and sensor array signal processing in the 1980s, and applications to semiconductor manufacturing and digital communications in the 1990s. Concurrently, he contributed to several fields of mathematics, especially stochastic processes, operator theory, and linear algebra. He has authored, edited, and coauthored several books. In the course of his research and teaching, he has mentored more than 100 doctoral and postdoctoral students and authored or coauthored more than 300 journal papers.

Prof. Kailath is a member of the National Academy of Engineering, the National Academy of Sciences, the American Academy of Arts and Sciences, the Third World Academy of Sciences, the Indian National Academy of Engineering, and the Royal Spanish Academy of Engineering. He has held Guggenheim, Churchill, and Humboldt fellowships, among others. He was President of the IEEE Information Theory Society in 1975 and received its Shannon Award in 2000. Among other awards he received are the Technical Achievement (1989) and Society (1991) awards of the IEEE Signal Processing Society, the 1995 IEEE Education Medal, and (with A. Sayed) the 1996 IEEE Donald G. Fink Prize Award. He has received outstanding paper prizes from the IEEE Information Theory Society, the IEEE Signal Processing Society, the European Signal Processing Society, and the IEEE Transactions on Semiconductor Manufacturing. He has received honorary degrees from Linköping University, Sweden; Strathclyde University, Scotland; the University of Carlos III, Madrid, Spain; and the University of Bordeaux, France.