MIMO Decision Feedback Equalization from an H^∞ Perspective

Alper T. Erdogan[*], Babak Hassibi[†], Thomas Kailath[‡]

Abstract

We approach the Multiple Input Multiple Output (MIMO) decision feedback equalization problem in digital communications from an H^∞ estimation point of view. Using the standard (and simplifying) assumption that all previous decisions are correct, we obtain an explicit parameterization of all H^∞ optimal decision feedback equalizers. In particular we show that, under the above assumption, minimum mean square error decision feedback equalizers are H^∞ optimal. The H^∞ approach also suggests a method for dealing with errors in previous decisions.

KEYWORDS: H-Infinity Estimation, Decision Feedback Equalization, Risk Sensitive Estimation. **EDICS:** 3-COMM, 2- ESTM. 3-CEQU

CONTACT:

Alper Tunga Erdogan

Address: Electrical and Electronics Engineering, Koc University, Rumeli Feneri Yolu, Sariyer 80910 Istanbul Turkey

Email: alperdogan@ku.edu.tr

Phone: +90 216 3881490

[*]Koc University

[†]California Institute of Technology

[‡]Stanford University

1 Introduction

The ultimate goal of digital communication is the reliable transmission of information at the highest possible data rates. One major obstacle in achieving this goal is the inter symbol interference (ISI) imposed by the communication channel. The inter symbol interference refers to the effect of neighboring symbols on the current symbol and unless it is handled properly it can lead to high Bit Error Rates (BER) in the recovery of the transmitted sequence at the receiver. Therefore, various methods have been developed to increase the communications systems' performance by reducing the effects of the ISI.

Linear equalization is one of the major attempts in this direction. However, linear equalization doesn't exploit the fact that the transmitted sequence has a "finite alphabet" structure. To take advantage of this property, the Decision Feedback Equalization (DFE) is proposed. Decision feedback equalizers use old decisions to improve the equalizer performance. This has been a research focus for more than two decades. The reference paper [1] provides a good summary and a historical overview of these research efforts. A more recent treatment of decision feedback equalization with minimum mean square error criterion (MMSE-DFE) is in [2].

Almost all the techniques proposed for equalization makes some assumptions about the underlying characteristics of the disturbance signals and the structure of the communication channel model. In many applications, however, true information about the channel is not available and algorithms have to use the estimates of the model parameters. For example, in mobile communications the channel parameters are often estimated via use of training sequences. The time variations in these parameters also necessitate the need for tracking them and the errors due to tracking is another point of concern. These concerns bring the question of robustness, that is, whether the small variations from the true model, and small disturbances, can cause large degradations in the performances of the algorithms using these parameters.

Recently the H^∞ criterion has been proposed [3] for the linear equalization with the belief that the resulting H^∞ equalizers will be more robust against the model uncertainties and the lack of statistical information of the exogenous signals. In [4] this approach has been further studied, yielding various new insights into the linear equalization problem such as role of non-minimum phase zeros and the delay in equalization. As outlined in [4], we can list the reasons for the use of H^∞ criterion to the equalization problem can be summarized as

- The Risk-Sensitive optimality of the central H^∞ equalizers, which provides an ensemble average optimality property similar to the average optimality of the MMSE equalizers.
- The worst case optimality, which reduces the maximum performance deviation from the average performance. This property provides a basis for the robust equalizer design framework.
- Existence of the fast algorithms for the implementation.

In this paper, we approach the multiuser decision feedback equalization problem from the H^∞ estimation point of view. In the first part of the paper, we introduce multiuser decision feedback equalization problem. Then we introduce an equivalent model and provide the MMSE-DFE solution for this model. Starting with Section 5, we look at the formulation of H^∞ equalizers under the assumption that the previous decisions input to the feedback filter are correct. Here among other results we show that MMSE equalizers are H^∞ optimal under this assumption. In the last part of the paper, we abandon the assumption about the correctness of previous decisions which complicates the decision feedback problem due to the extreme difficulty in the modelling of the decision errors. However, we will show that H^∞ criterion based approach can still provide a solution in this case.

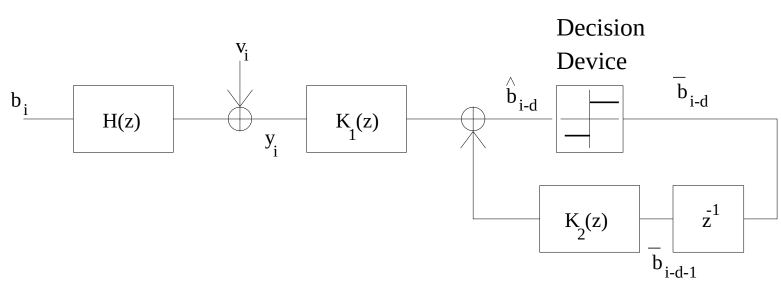

Figure 1: Decision Feedback Equalization

2 Decision Feedback Equalization Problem

The standard discrete time model for the decision feedback equalization problem is illustrated in Figure 1. In this figure, $\{b_i\}$ represents the discrete time finite-alphabet input-data sequence. If we assume the number of co-channel users to be M then $b_i \in \mathcal{C}^M$. The distortion effects of the communications medium are represented by a linear time invariant transfer function matrix $H(z)$, which reflects the effects of the transmit and receive filters and of the propagation environment (e.g., the multipath vector channel of an antenna array system in wireless communications system). The dimensions of $\{b_i\}$ and $\{y_i\}$ determine the dimensions of the channel $H(z)$. If there are N antennas or branches at the receiver, i.e., $y_i \in \mathcal{C}^N$, then H(z) is assumed to be a causal and stable $N \times M$ matrix function in z with Laurent series expansion

$$H(z) = h_0 + h_1 z^{-1} + ... \tag{1}$$

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 $H_{ij}(z)$ entry in the matrix refers to the effective channel between the user j and the antenna i. We also assume that the number of users is less than or equal to the number of antennas, i.e., $M \leq N$.

The sequence $\{v_i\}$ represents the noise disturbance (e.g., receiver antenna noise, co-channel interference, etc.) corrupting the observations. Modeling errors due to imperfect knowledge of the true channel can also be incorporated into the disturbance $\{v_i\}$. We shall, therefore, for the most part not make any statistical assumptions about the disturbance sequence $\{v_i\}$ and will simply consider it as an unknown sequence of elements in $\mathcal{C}^N$.

The frequency selective property of the $H(z)$ results in ISI for the observed signal and therefore, it is desirable to reduce the frequency selective property of the channel to reduce ISI. We also need to take the effects of noise into consideration. Referring to Figure 1, our aim in decision feedback equalization is to design causal filters $K_1(z)$ and $K_2(z)$ to estimate b_{i-d} , where $d \geq 0$ is the parameter indicating the delay in estimating the transmitted sequence. Here $K_1(z)$ is the feedforward filter that has the observations $\{y_i\}$ as its input, and $K_2(z)$ is the feedback filter that has the previous decisions $\{\bar{b}_{i-d-1}\}$ as its input. The estimate, denoted by $\hat{b}_{i-d}$, is the sum of the outputs of the $K_1(z)$ and $K_2(z)$, whereas the decisions $\bar{b}_{i-d-1}$ are obtained by passing b_{i-d} through a decision device. The design of the filters $K_1(z)$ and $K_2(z)$ depends upon the criterion chosen to define the closeness of $\hat{b}_{i-d}$ to b_{i-d}.

In almost all the decision feedback equalizer designs in the literature, the decisions input to the filter $K_2(z)$ are assumed to be correct; this simplifying assumption converts the original non-tractable nonlinear problem into a solvable linear one. Moreover, most of the research in the decision feedback area is focused on the mean square error criterion [5, 2, 6], mostly because it allows the derivation of explicit formulas for both the feedforward and feedback filters. As summarized in [1], in most of these derivations, the feedforward filter is assumed to be non-causal, i.e., a smoothing filter, and therefore in applications it should be approximated by a causal filter with a certain delay. Again, as shown in the same paper, the formulation reduces to solving a mean square error linear prediction problem.

In this article, we will use the H^∞ criterion as the the basis for the derivation of the filters. In doing so, we will consider the setup of Figure 1 constraining the feedforward filter to be causal.

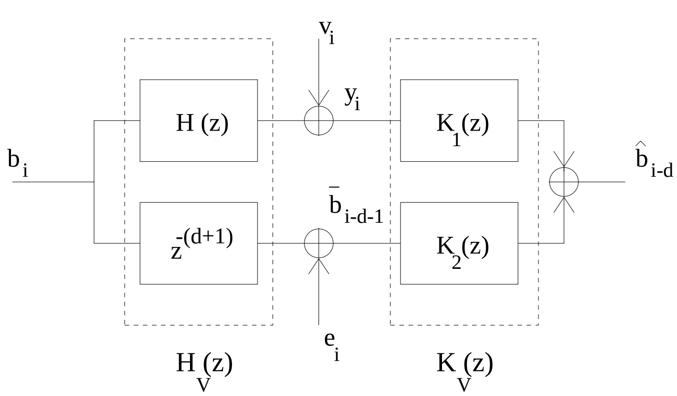

Figure 2: Equivalent model for Decision Feedback Equalization.

3 An Equivalent Model for Decision Feedback Equalization

We can remodel the decision feedback equalization problem in Figure 1 as shown in Figure 2 so that it takes the form of a general estimation problem with $L(z) = z^{-d}I$, as we are trying to estimate the d delayed input symbols. In this figure, $\bar{b}_{i-d-1}$ represents possibly incorrect previous decisions and $e_i = \bar{b}_{i-d-1} - b_{i-d-1}$ represents the corresponding errors in the decisions. Thus, in this model

$$H_V(z) = \begin{bmatrix} H(z) & z^{-(d+1)}I \end{bmatrix}^T \tag{2}$$

is the equivalent matrix channel. According to the same model

$$Y_V(z) = \begin{bmatrix} Y(z) & \bar{B}(z) \end{bmatrix}^T \text{ and } N_V(z) = \begin{bmatrix} V(z) & E(z) \end{bmatrix}^T \tag{3}$$

are the equivalent observation and noise vectors, respectively. We define the energy weighting matrix for the noise N_V as

$$R_{N_V} = \begin{bmatrix} R & 0 \\ 0 & \epsilon I \end{bmatrix} \tag{4}$$

where R represents the weight we assign to the additive noise $\{v_i\}$ and ϵ represents the weight assigned to the decision error sequence $\{e_i\}$. Then the decision feedback equalization problem is equivalent to finding

$$K_V(z) = \begin{bmatrix} K_1(z) & K_2(z) \end{bmatrix} \tag{5}$$

that minimizes a certain norm of the transfer function

$$T_{K_V}(z) = \begin{bmatrix} (z^{-d}I - K_V(z)H_V(z))Q^{1/2} & -K_V(z)R_{N_V}^{1/2} \end{bmatrix}. \tag{6}$$

In the rest of the paper, without loss of generality, we will assume $Q = I$ to simplify the expressions.

4 MMSE Decision Feedback Equalization

In this section, we formulate the MMSE decision feedback equalizers for the equivalent channel model of the previous section under correct decisions assumption (without this assumption the MMSE procedure isn't tractable). The solution to this problem is well known. However, we will repeat here the formulation as it provides a good basis of comparison with the H^∞ approach, and the treatment will be more general than most of the approaches in the literature where the feed-forward filter is assumed to be non-causal and to have access to infinite future observations. The following theorem gives the formulation for MMSE-DFE equalizers. It uses the general H^2 approach outlined in [7]:

Theorem 1 (H^2-optimal DFE) *The solution to the problem*

$$\min_{causal K_V} \|T_{K_V}(z)\|_2 \tag{7}$$

where $T_{K_V}(z)$ is given in Eq. (6), for the case $\epsilon = 0$, is given by

$$K_V = \{z^{-d}H^*(z^{-*})M^{-*}(z^{-*})\}_+ M^{-1}(z) \tag{8}$$

where $M(z)$ is found from the canonical factorization of the power spectrum matrix

$$S_{Y_V}(z) = R_{N_V} + H_V(z)H_V^*(z^{-*}) = M(z)M^*(z^{-*}) \tag{9}$$

with $M(z)$ causal and causally invertible and where $\{A(z)\}_+$ denotes the causal part of the transfer operator, $A(z)$. In addition the corresponding error spectrum is given by

$$Er(z) = \{z^{-d}H^*(z^{-*})M^{-*}(z^{-*})\}_- \{\{z^{-d}H^*(z^{-*})M^{-*}(z^{-*})\}_-\}^* \tag{10}$$

where $\{A(z)\}_-$ refers to strictly anti-causal part of the transfer function $A(z)$ and $\{A(z)\}^* = A^*(z^{-*})$.

5 H^∞ Decision Feedback Equalization

In this section, we look at the formulation of the decision feedback equalizers with respect to the H^∞ criterion. First, we will assume that the old decisions are correct, as we did in the MMSE formulation, and look at the derivation of $H^\infty - DFE$ equalizers. We then abandon this assumption in the last section and look at the solution provided by the H^∞ framework for this case.

5.1 Correct Decisions Case

We approach the derivation of H^∞ decision feedback equalizers under the correct decisions assumption by first concentrating on the case in which $d = 0$ and then by showing how we can generalize this approach. Under this assumption, e_i in Figure 2 is equal to zero for all i and therefore the corresponding weight is $\epsilon = 0$. The major result under the correct decisions assumption is that the MMSE-DFE equalizer turns out to be H^∞ optimal. This is a striking result since H^2 and H^∞ approaches generally, with the exception of some trivial cases, yield different results.

5.1.1 $H^\infty - DFE$ Equalizers for $d = 0$

Focusing on the $d = 0$ case gives us the flavor of the general formulation of the $H^\infty - DFE$ filters. We will state the result related to this case by the following theorem:

Theorem 2 **(H^∞ DFE for $d = 0$)** *For the setting described by Figure 2, under the correct decisions assumption, i.e., $e_i = 0$ and for $d = 0$, the solution to the problem*

$$\min_{causal K_V(z)} ||T_{K_V}(z)||_\infty \tag{11}$$

where $T_{K_V}(z)$ is given by Eq. (6), can be obtained for

$$\gamma^2 \geq \frac{1}{1 + \sigma_{min}(h_0^* R^{-1} h_0)} \tag{12}$$

and is given by

$$K_V(z) = (L_{22}(z)C(z) - L_{21}(z))\,(L_{11}(z) - L_{12}(z)C(z))^{-1} \tag{13}$$

where $C(z)$ is a causal and strictly contractive transfer function and

$$L(z) = \begin{bmatrix} I & h_1 + h_2 z^{-1} + \ldots & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}$$
$$\left[\begin{array}{cc|c} (h_0 h_0^* + R)^{1/2} & 0 & 0 \\ z^{-1} h_0^* (h_0 h_0^* + R)^{-1/2} & (\frac{\gamma^2 - 1}{\gamma^2} I - h_0^* R^{-1} h_0)^{-1/2} & z^{-1}(\frac{I - \gamma^2 h_0^* R^{-1} h_0}{1-\gamma^2}(\gamma^2 I - (I + h_0^* R^{-1} h_0)^{-1})^{1/2}) \\ \hline -h_0^* (h_0 h_0^* + R)^{-1/2} & 0 & (\gamma^2 I - (I + h_0^* R^{-1} h_0)^{-1})^{1/2} \end{array}\right]. \tag{14}$$

Proof: Proof of this theorem is given in Appendix A.

Remarks:

- It is interesting to compare the performance of the H^∞ decision feedback equalizer with the H^∞ linear equalizer by comparing the corresponding optimal H^∞ norms. As shown in [4], for a scalar channel with $R = r$, the linear equalizer has

$$\gamma^2_{opt,linear} = \frac{r}{r + \min_w |H(e^{j\omega})|^2} \tag{15}$$

for a *minimum phase* $H(z)$ and

$$\gamma^2_{opt,linear} = 1 \tag{16}$$

for a *non-minimum phase* $H(z)$. For the decision feedback equalizer, irrespective of the minimum phase property of the channel, Eq. (12) yields

$$\gamma^2_{opt,dfe} = \frac{r}{r + |h_0|^2}. \tag{17}$$

For non-minimum phase channels, obviously $\gamma^2_{opt,dfe} \leq \gamma^2_{opt,linear}$ since $|h_0|^2 \geq 0$ and therefore $\gamma^2_{opt,dfe} \leq 1$. For minimum phase channels, over the region $|z| \geq 1$, the minimum value of the $H(z)$ is achieved on the unit circle. This is due to the observation that $H^{-1}(z)$ has all its poles inside the unit circle and therefore, by the maximum modulus theorem, $H^{-1}(z)$ achieves its maximum on the unit circle for $|z| \geq 1$. Thus, $H(z)$ achieves its minimum on the unit circle for this region . Since $h_0 = H(\infty)$, we have $|h_0|^2 \geq min_\omega |H(e^{j\omega})|^2$, so that,

$$\gamma^2_{opt,dfe} \leq \gamma^2_{opt,linear} \tag{18}$$

i.e., the performance of the H^∞ decision feedback equalizer is better than the performance of the H^∞ linear equalizer with respect to the H^∞ criterion.

- Another important observation is obtained when we look at the central solution to the decision feedback equalization problem by using $L(z)$ given by Eq. (14):

$$\begin{aligned} K_{central}(z) &= -L_{21}(z)L_{11}^{-1}(z) \\ &= \begin{bmatrix} h_0^*(h_0h_0^* + R)^{-1} & -(h_1 + h_2z^{-1} + ...)h_0^*(h_0h_0^* + R)^{-1} \end{bmatrix} \end{aligned}$$

which turns out to be the MMSE decision feedback equalizer (MMSE-DFE) for the given setup and with the additional statistical assumptions . This is an important observation which we generalize in the next section.

5.1.2 H^∞ optimality of MMSE Decision Feedback Equalization

An important result from the previous section is that under the correct previous decisions assumption, and for $d = 0$, the MMSE decision feedback equalizer is H^∞ optimal. If we carry out the factorization for a general $d > 0$, we see that the MMSE solution still coincides with the corresponding central H^∞ solution. In this section, we shall prove this fact using a different route.

For the equivalent channel model described in the previous section, the MMSE decision feedback equalizer for any $d \geq 0$ can be found using

$$K_{MMSE}(z) = \{z^{-d} H_V^*(z^{-*}) M^{-*}(z^{-*})\}_+ M^{-1}(z) \tag{19}$$

where $M(z)$ is as defined in Eq. (9) and the error spectrum corresponding to the equalizer is given by

$$Er(z) = \{z^{-d} H_V^*(z^{-*}) M^*(z^{-*})\}_- \{\{z^{-d} H_V^*(z^{-*}) M^*(z^{-*})\}_-\}^* \tag{20}$$

where $\{.\}_-$ extracts the strictly non-causal part of its argument and $\{A(z)\}^* \triangleq A^*(z^{-*})$ for any function $A(z)$.

To obtain the spectral factorization of $S_{Y_V}(z)$, let us first write

$$\begin{aligned}
S_{Y_V}(z) &= R_{N_V} + H_V(z) H_V^*(z^{-*}) \\
&= \begin{bmatrix} R + H(z)H^*(z^{-*}) & H(z)z^{d+1} \\ z^{-d-1}H^*(z^{-*}) & I \end{bmatrix} \\
&= \underbrace{\begin{bmatrix} R^{1/2} & H(z)z^{d+1} \\ 0 & I \end{bmatrix}}_{N(z)} \underbrace{\begin{bmatrix} R^{1/2} & 0 \\ H^*(z^{-*})z^{-d-1} & I \end{bmatrix}}_{N^*(z^{-*})}.
\end{aligned}$$

Comparing with the spectral factorization in Eq. (9), we conclude that

$$M(z) = N(z)\Theta(z) \tag{21}$$

$$= \begin{bmatrix} R^{1/2} & H(z)z^{d+1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \theta_{11}(z) & \theta_{12}(z) \\ \theta_{21}(z) & \theta_{22}(z) \end{bmatrix} \tag{22}$$

$$= \begin{bmatrix} R^{1/2}\theta_{11}(z) + H(z)z^{d+1}\theta_{21}(z) & R^{1/2}\theta_{21}(z) + H(z)z^{d+1}\theta_{22}(z) \\ \theta_{21}(z) & \theta_{22}(z) \end{bmatrix} \tag{23}$$

where $\Theta(z)\Theta^*(z^{-*}) = I$ is chosen such that $M(z)$ is causal and causally invertible. The causality of the $M(z)$ constrains $\theta_{21}(z)$ and $\theta_{22}(z)$ to be causal.

We can write the error spectrum $Er_{MMSE}(z)$ as

$$\begin{aligned}
& \left\{z^{-d}\begin{bmatrix} H^*(z^{-*}) & z^{d+1}I \end{bmatrix} M^{-*}(z-*)\right\}_- \left\{\left\{z^{-d}I\begin{bmatrix} H^*(z^{-*}) & z^{d+1}I \end{bmatrix} M^{-*}(z-*)\right\}_-\right\}^* \\
= & \left\{\begin{bmatrix} z^{-d}H^*(z^{-*}) & zI \end{bmatrix} N^{-*}(z^{-*})\Theta(z)\right\}_- \left\{\left\{\begin{bmatrix} z^{-d}H^*(z^{-*}) & z \end{bmatrix} N^{-*}(z^{-*})\Theta(z)\right\}_-\right\}^* \\
= & \left\{\begin{bmatrix} 0 & zI \end{bmatrix}\begin{bmatrix} \theta_{11}(z) & \theta_{12}(z) \\ \theta_{21}(z) & \theta_{22}(z) \end{bmatrix}\right\}_- \left\{\left\{\begin{bmatrix} 0 & zI \end{bmatrix}\begin{bmatrix} \theta_{11}(z) & \theta_{12}(z) \\ \theta_{21}(z) & \theta_{22}(z) \end{bmatrix}\right\}_-\right\}^* \\
= & \begin{bmatrix} z\theta_{21,0} & z\theta_{22,0} \end{bmatrix}\begin{bmatrix} z^{-1}\theta^*_{21,0} \\ z^{-1}\theta^*_{22,0} \end{bmatrix} \\
= & \theta_{21,0}\theta^*_{21,0} + \theta_{22,0}\theta^*_{22,0}
\end{aligned}$$

where we used the fact that $\theta_{21}(z)$ and $\theta_{22}(z)$ are causal.

Eq. (24) shows that the resulting error spectrum is frequency independent. For the scalar case, the MMSE equalizers minimize the area under the error spectrum, whereas the H^∞ equalizers minimize the peak of the error spectrum. Thus, the frequency independence implies the flatness of the error spectrum, which in turn implies that the MMSE-DFE equalizer is H^∞ optimal. The reason is that any other DFE equalizer with a maximum value of the error spectrum less than that of the MMSE-DFE equalizer will have to have a smaller area under its spectrum than the MMSE-DFE case, which is a contradiction.

This property of MMSE-DFE can be extended for more general matrix channels by the use of operator techniques developed in [8] and outlined in Appendix B. In order to show that the H^∞ and the H^2 solutions coincide, we need to show that γ^2_{opt} is equal to the maximum singular value of the MMSE error spectrum which is a constant matrix. From Appendix B, we know that

$$\gamma_{opt} = ||\mathcal{E}_- + P^*_H P_H||_\infty \tag{24}$$

where the $\mathcal{E}_-$ is equal to zero, since the smoothing spectrum is equal to zero. Therefore,

$$\sigma_{opt} = ||P_H^* P_H||_\infty. \tag{25}$$

Here, using the results we obtained above,

$$\begin{aligned} P(z) &= M^{-1}(z)H_V(z)L^*(z) \\ &= \Theta^*(z)N^{-1}(z)H_V(z)z^{-d} \\ &= \Theta^*(z)\begin{bmatrix} 0 \\ z^{-1} \end{bmatrix} = z^{-1}\begin{bmatrix} \theta_{21}^*(z^{-*}) \\ \theta_{22}^*(z^{-*}) \end{bmatrix}. \end{aligned}$$

Therefore, since $\theta_{21}(z)$ and $\theta_{22}(z)$ are causal operators, we can write, $P_H = \begin{bmatrix} \theta_{21,0}^* & \theta_{22,0}^* \end{bmatrix}^T$. As a result,

$$\gamma_{opt} = ||P_H^* P_H||_\infty \tag{26}$$

$$= ||\theta_{21,0}\theta_{21,0}^* + \theta_{22,0}^*\theta_{22,0}^*||_\infty \tag{27}$$

$$= ||Er_{MMSE}||_\infty \tag{28}$$

which proves that the H^2 and the H^∞ solutions coincide for the more general matrix channel case.

This is a striking result, which sheds further light on the properties of the MMSE-DFE equalizer. Moreover, this is a rare case, where the solutions to the H^∞ and MMSE filtering problems coincide. In general, except for some trivial cases, solutions to both problems differ and a trade off exists between two criteria. In fact, one active research area is the design of mixed H^2/H^∞ filters. We should note that the equivalence relation shown should not be confused with the well-known limiting equivalence of the H^∞ estimator to the H^2 estimator, where the limit is on the γ parameter. The equivalence obtained is for finite γ levels and it's a property of the decision feedback structure.

5.1.3 Derivation of $H^\infty - DFE$ Equalizers for $d > 0$

We previously looked only at the case $d = 0$. For any $d > 0$ it is also possible to obtain explicit formulas for the $H^\infty - DFE$ equalizers since the factorization of the Popov function can be achieved easily but with increasing complexity of the expressions due to the following lemma

Lemma 1 *The Popov function*

$$\begin{bmatrix} R_{N_V} + H_V(z)H_V^*(z^{-*}) & -H_V(z)z^d \\ -z^{-d}H_V^*(z^{-*}) & (1-\gamma^2)I \end{bmatrix},$$

under the correct decisions assumption, i.e., $\epsilon = 0$, *is always unimodular.*

Proof: For any delay $d \geq 0$, we can factor the Popov function as

$$\begin{bmatrix} R_{N_V} + H_V(z)H_V^*(z^{-*}) & -H_V(z)z^d \\ -z^{-d}H_V^*(z^{-*}) & (1-\gamma^2)I \end{bmatrix}$$
$$= \underbrace{\begin{bmatrix} I & -\frac{H_V(z)z^d}{1-\gamma^2} \\ 0 & I \end{bmatrix}}_{F(z)} \begin{bmatrix} R_{N_V} - \frac{\gamma^2}{1-\gamma^2}H_V(z)H_V^*(z^{-*}) & 0 \\ 0 & (1-\gamma^2)I \end{bmatrix} \underbrace{\begin{bmatrix} I & 0 \\ -\frac{H_V^*(z^{-*})z^{-d}}{1-\gamma^2} & I \end{bmatrix}}_{F^*(z^{-*})}.$$

The factors $F(z)$ and $F^*(z^{-*})$ are clearly unimodular matrices, since they are triangular matrices with constant diagonals. The center matrix is also unimodular since

$$R_{N_V} - \frac{\gamma^2}{1-\gamma^2}H_V(z)H_V^*(z^{-*}) = \begin{bmatrix} I & H(z)z^{d+1} \\ 0 & I \end{bmatrix} \begin{bmatrix} R & 0 \\ 0 & \frac{-\gamma^2}{1-\gamma^2}I \end{bmatrix} \begin{bmatrix} I & 0 \\ z^{-(d+1)}H^*(z^{-*}) & I \end{bmatrix}$$

is unimodular, since as shown above, it is a product of 3 unimodular matrices. As a result, the Popov function is unimodular. ∎

Therefore we can systematically factor the Popov function using unimodular lower-upper and upper-lower factors. In doing so, one can follow an approach similar to the approach for the $d = 0$ case. We begin by first factoring $H(z)$ as in Appendix A , but for a general $d \geq 0$:

$$\begin{aligned} H(z) &= h_0 + h_1 z^{-1} + ...h_d z^{-d} + z^{-(d+1)}(h_{d+1} + h_{d+2}z^{-1} + ...) && (29) \\ &= h_0 + h_1 + ... + h_d z^{-d} + z^{-(d+1)}H_c(z), && (30) \end{aligned}$$

which leads to

$$H_v(z) = \begin{bmatrix} H(z) \\ z^{-(d+1)}I \end{bmatrix} = \begin{bmatrix} I & H_c(z) \\ 0 & I \end{bmatrix} \begin{bmatrix} h_0 \\ Iz^{-(d+1)} \end{bmatrix} = F_1(z)H_{ev}(z). \quad (31)$$

Therefore, we can reduce the factorization of the original Popov function $\Sigma(z)$ to factorization of the "equivalent" Popov function

$$\Sigma_{ev}(z) = \begin{bmatrix} R_{N_V} + H_{ev}(z)H_{ev}^*(z^{-*}) & -z^d H_{ev}(z) \\ -H_{ev}^*(z^{-*})z^{-d} & (1-\gamma^2)I \end{bmatrix} \tag{32}$$

In Appendix C, in Theorem 6, we show that for $d > 0$, under the correct decisions assumption

$$\gamma^2_{opt,delay} = \sigma_{max}(\begin{bmatrix} I_M & 0...0 \end{bmatrix} (I + H_d^* r^{-1} H_d) \begin{bmatrix} I_M & 0...0 \end{bmatrix}^T)$$

where

$$H_d = \begin{bmatrix} h_0 & 0 & ... & ... & ... & 0 \\ h_1 & h_0 & ... & .. & .... & 0 \\ \vdots & ... & \ddots & ... & ... & 0 \\ h_{d-2} & ... & h_1 & h_0 & 0 & 0 \\ h_{d-1} & ... & h_2 & h_1 & h_0 & 0 \\ h_d & ... & h_3 & h_2 & h_1 & h_0 \end{bmatrix}. \tag{33}$$

This means that we can directly calculate the optimal value of γ for the H^∞-DFE problem for any delay using the channel impulse response coefficients.

5.2 Error in Previous Decisions

In the previous sections, to simplify the analysis and the derivations of the filters, we assumed that the decisions used by the feedback filter were always correct. This assumption can hold in systems that use precoding techniques to implement the feedback part in the transmitter section, as in the well-known Tomlinson-Harashima precoding procedure [9]. However, such procedures require a priori knowledge of the channel and the statistics of the exogenous input signals for the design of the transmitter. In applications requiring adaptive communication capabilities, such as wireless communications systems, this is not a reasonable assumption. Since the channel is estimated at the receiver, adapting the transmitter with this information is not feasible in time-variant environments. Therefore, in such situations the feedback filter should be implemented at the receiver, which inevitably leads to incorrect decisions input to the feedback filter. If filters are designed under the correct previous decisions assumption, the existence of decision errors leads to a degradation in the performance.

In this section, we shall not assume that the decision errors are zero, but that they form some non-zero sequence $\{e_i\}$. Since $\{e_i\}$ is a complicated function of the feed-forward and feedback filters, as well as other parameters in the system, it is almost impossible to give an explicit statistical description of the errors and therefore design filters with respect to the statistical criterion such as MMSE criterion.

However, as far as the H^∞ criterion is concerned, $\{e_i\}$ is a nonzero sequence with small power given by ϵ, and therefore the H^∞ approach can provide a solution which safeguards against the worst-case decision errors.

When $\epsilon \neq 0$, the corresponding Popov function is no longer unimodular and therefore the J-Spectral factorization based approach is not as easy as the correct decisions case. However, we can still obtain numerical solutions by solving Riccati equations (or recursions) and we can implement equalizers with state space models as shown in [10]. In the design of equalizers we need to choose the parameter γ and the ϵ parameter in R_{N_V}. The γ should clearly be greater than γ_{opt}. Although there is no explicit expression for γ_{opt}, one can use the upper bound :

$$\gamma_{opt}^2 \leq \sigma_{max}(\left[\begin{array}{cc} I_M & 0..0 \end{array} \right] ((I + r^{-1} H_d^* H_d)^{-1} + \epsilon Z) \left[\begin{array}{cc} I_M & 0...0 \end{array} \right]^T), \tag{34}$$

where

$$Z = (rI + H_d^* H_d)^{-1} H_d^* h_d h_d^* H_d (rI + H_d^* H_d)^{-1}, \tag{35}$$

which is derived in Appendix C.

The choice of the ϵ parameter is critical, since it represents the power of the decision errors, which is not known beforehand. Figure 3 illustrates the variation of the BER of the equalizer as a function of the ϵ parameter for the example $H(z) = 0.56 - 0.06z^{-1} + 1.07z^{-2} + 1.6z^{-3} - 0.13z^{-4}$, delay $d = 2$ and $SNR = 18dB$ for the central H^∞ equalizer. Initially as we increase the value of ϵ from 0, the BER decreases. This is due to the fact that the equalizer is taking into account the existence of the decision error, so the performance improves. However, after a certain point, this trend reverses and the BER begins to increase because ϵ begins to overestimate the decision error power.

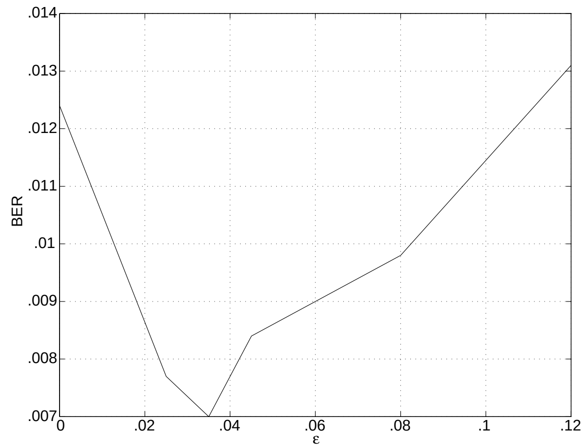

Figure 3: BER vs. ϵ

For binary signaling (with levels +1 and -1), the only possible values of the errors are 2 and -2. Therefore, the error power would be 4 times the BER. Since the BER itself is also dependent on the ϵ parameter chosen, it is hard to obtain an explicit expression for the optimal value of ϵ. But given that the error power at $\epsilon = 0$ is 4 times the $BER_{MMSE-DFE}$(BER of MMSE-DFE filter assuming correct decisions), ϵ_{opt} should be less than this value since its BER is lower than $BER_{MMSE-DFE}$. Figure 3 suggests 3 times $BER_{MMSE-DFE}$ for the optimal value of the epsilon, which is a reasonable choice for practical applications. Here the $BER_{MMSE-DFE}$ can be approximated , for example, using the upper bound formulas suggested in [11]. Similar reasoning can be followed in the computation of the optimal value of ϵ for constellations other than binary.

6 Conclusion

We studied the problem of decision feedback equalization from the H^∞ estimation point of view. If we make the assumption that the previous decisions are correct, the H^∞ formulation of the decision feedback equalizers can be obtained simply through factorization of the Popov function. As an important result, under this assumption the MMSE and H^∞ solutions coincide, which is an interesting result both from an equalization and estimation theory point of view. Once we remove this assumption, it is hard to formulate DFE Filters with respect to the

MMSE criterion, however, the H^{∞} framework still provides a solution due to its deterministic worst-case setup.

Appendix A

We formulate the H^{∞}-DFE equalizers for the zero delay case, which provides a proof for the Theorem 2. We follow the J-Spectral factorization based approach, and therefore, we begin by writing the Popov function for the equivalent DFE model of Figure 2 of Section 3:

$$\Sigma(z) = \begin{bmatrix} R_{N_V} + H_V(z)H_V^*(z^{-*}) & -H_V(z)z^d \\ -z^{-d}H_V^*(z^{-*}) & (1-\gamma^2)I \end{bmatrix}$$

where, for simplicity, we assumed binary antipodal signaling for which $b_i \in \{-1,1\}^M$ and therefore $Q = I$. The results can be easily generalized to more complex signal constellations.

We follow the factorization procedure outlined below:

1. We will iteratively extract factors from the Popov function $\Sigma(z)$.

2. By assuming a positive range for γ, we write $\Sigma(z) = P(z)JP^*(z^{-*})$ where $P(z)$ is a causal and causally invertible matrix and J is the mixed inertia matrix we previously defined. Here we note that the inertia condition for the factorization is always satisfied since the smoothing error spectrum is identical to zero which further implies $\gamma_{smoothing} = 0$. Therefore, as long as $\gamma > \gamma_{smoothing} = 0$ the inertia condition is satisfied.

3. If there exists a J-unitary matrix Θ such that $L(z) = P(z)\Theta$ has strictly causal block $L_{12}(z)$ and, causal and causally invertible block $L_{11}(z)$ then the assumed range for γ is greater than γ_{opt}.

Before we begin extracting factors from the Popov function, we first define the following factorization of H(z):

$$\begin{aligned} H(z) &= h_0 + z^{-1}(h_1 + h_2 z^{-1} + ...) &(36)\\ &= h_0 + z^{-1}H_c(z). &(37)\end{aligned}$$

So that we can write

$$H_v(z) = \begin{bmatrix} H(z) \\ z^{-1}I \end{bmatrix} = \begin{bmatrix} I & H_c(z) \\ 0 & I \end{bmatrix} \begin{bmatrix} h_0 \\ z^{-1}I \end{bmatrix} = F(z)H_{ev}(z). \qquad (38)$$

Note that, since $R_{N_V} = \begin{bmatrix} R & 0 \\ 0 & 0 \end{bmatrix} = F(z)R_{N_V}F^*(z^{-*})$, we obtain the following equality for the $\Sigma(z)$:

$$\Sigma(z) \quad = \quad \underbrace{\begin{bmatrix} F(z) & 0 \\ 0 & I \end{bmatrix}}_{F_1(z)} \underbrace{\begin{bmatrix} R_{N_V} + H_{ev}(z)H_{ev}^*(z^{-*}) & -H_{ev}(z) \\ -H_{ev}^*(z^{-*}) & I - \gamma^2 \end{bmatrix}}_{\Sigma_{ev}(z)} \underbrace{\begin{bmatrix} F^*(z^{-*}) & 0 \\ 0 & I \end{bmatrix}}_{F_1^*(z^{-*})}.$$

Due to the structure of $F_1(z)$, which is causal and causally invertible and diagonal, we can obtain the desired J-spectral factorization of $\Sigma(z)$ by first finding the J-spectral factorization of the equivalent Popov function $\Sigma_{ev}(z)$ as $L_{ev}(z)JL_{ev}^*(z^{-*})$ and then we can write $L(z) = F_1(z)L_{ev}(z)$.

Therefore, we continue by further factorization of $\Sigma_{ev}(z)$:

$$\Sigma_{ev}(z) = \underbrace{\begin{bmatrix} I & -\frac{H_{ev}(z)}{1-\gamma^2} \\ 0 & I \end{bmatrix}}_{F_2(z)} \quad \underbrace{\begin{bmatrix} R_{N_V} - \frac{\gamma^2}{1-\gamma^2}H_{ev}(z)H_{ev}^*(z^{-*}) & 0 \\ 0 & (1-\gamma^2)I \end{bmatrix}}_{\Sigma_1(z)} \underbrace{\begin{bmatrix} I & 0 \\ -\frac{H_{ev}^*(z^{-*})}{1-\gamma^2} & 1 \end{bmatrix}}_{F_2^*(z^{-*})}.$$

We can further factorize $\Sigma_1(z)$ as

$$\Sigma_1(z) = F_3(z)\Sigma_2(z)F_3^*(z^{-*})$$

where

$$F_3(z) = \begin{bmatrix} R^{1/2} & 0 & 0 \\ -\frac{\gamma^2}{1-\gamma^2}h_{m0}^*(I - \frac{\gamma^2}{1-\gamma^2}h_{m0}h_{m0}^*)^{-1}z^{-1} & I & 0 \\ 0 & 0 & I \end{bmatrix}$$

and

$$\Sigma_2(z) = \begin{bmatrix} I - \frac{\gamma^2}{1-\gamma^2}h_{m0}h_{m0}^* & 0 & 0 \\ 0 & (-\frac{1-\gamma^2}{\gamma^2} + h_{m0}^*h_{m0})^{-1} & 0 \\ 0 & 0 & (1-\gamma^2)I \end{bmatrix},$$

where $h_{m0} = R^{-1/2}h_0$.

Now if we write the singular value decomposition for h_{m0} as

$$h_{m0} = U \begin{bmatrix} S \\ 0_{(N-M)\cdot M} \end{bmatrix} V^*$$

where $S = diag(\sigma_1(h_{m0}), \sigma_2(h_{m0}), ..., \sigma_1(h_{m0}))$, with $\sigma_1(h_{m0}) \geq \sigma_2(h_{m0}) \geq ... \geq \sigma_M(h_{m0})$. Here, without loss of generality, we will assume distinct singular values. We can rewrite $\Sigma_2(z)$ as

$$\Sigma_2(z) = \begin{bmatrix} U(I - \frac{\gamma^2}{1-\gamma^2} \begin{bmatrix} S^2 & 0 \\ 0 & 0 \end{bmatrix})U^* & 0 & 0 \\ 0 & V(-\frac{1-\gamma^2}{\gamma^2} I - S^2)^{-1} V^* & 0 \\ 0 & 0 & (1-\gamma^2)I \end{bmatrix}.$$

We can now write $\Sigma_1(z)$ as

$$\Sigma_1(z) = F_4(z)\Sigma_3 F_4^*(z^{-*})$$

where

$$F_4(z) = \begin{bmatrix} R^{1/2}U & 0 & 0 \\ -\frac{\gamma^2}{1-\gamma^2} h_{m0}^*(I - \frac{\gamma^2}{1-\gamma^2} h_{m0}h_{m0}^*)^{-1} U z^{-1} & V & 0 \\ 0 & 0 & I \end{bmatrix}$$

and

$$\Sigma_3 = \begin{bmatrix} I - \frac{\gamma^2}{1-\gamma^2} \begin{bmatrix} S^2 & 0 \\ 0 & 0 \end{bmatrix} & 0 & 0 \\ 0 & (-\frac{1-\gamma^2}{\gamma^2} I - S^2)^{-1} & 0 \\ 0 & 0 & (1-\gamma^2)I \end{bmatrix}.$$

Here which entries of the diagonal matrix Σ_3 are negative or positive depends on the value of γ. Therefore, we continue further factorization of $\Sigma_{ev}(z)$ by assuming different ranges for γ:

1. We assume $\frac{1}{1+\sigma_i^2(h_{m0})} < \gamma^2 < \frac{1}{1+\sigma_{i+1}^2(h_{m0})}$ and $i < M$. Under this assumption,

$$I - \frac{\gamma^2}{1-\gamma^2} S^2 = \begin{bmatrix} -A & 0 \\ 0 & B \end{bmatrix}$$

where

$$A = -I + \frac{\gamma^2}{1-\gamma^2} diag(\sigma_1^2(h_{m0}), \sigma_2^2(h_{m0}), ..., \sigma_i^2(h_{m0})),$$

$$B = I - \frac{\gamma^2}{1-\gamma^2} diag(\sigma_{i+1}^2(h_{m0}), \sigma_{i+2}^2(h_{m0}), ..., \sigma_m^2(h_{m0})).$$

where, due to the assumed range for γ, both A and B are positive matrices. After some algebra we can write

$$\Sigma_1(z) = F_5(z) \begin{bmatrix} I_{N+M} & 0 \\ 0 & -I_M \end{bmatrix} F_5^*(z^{-*})$$

where

$$F_5(z) = \begin{bmatrix} 0 & 0 & \sqrt{1-\gamma^2} \\ (R^{1/2}U_{i+1:M}B^{1/2})^T & -(C^2 h_{m0}^* U_{i+1:M} B^{-1/2} z^{-1})^T & 0 \\ (R^{1/2}U_{M+1:N})^T & -(C^2 h_{m0}^* U_{M+1:N} z^{-1})^T & 0 \\ 0 & (V_{1:i} C A^{-1/2})^T & 0 \\ 0 & (V_{i+1:M} C B^{-1/2})^T & 0 \\ (R^{1/2}U_{1:i}A^{1/2})^T & (C h_{m0}^* U_{1:i} A^{-1/2} z^{-1})^T & 0 \end{bmatrix}^T$$

and $C = \sqrt{\frac{\gamma^2}{1-\gamma^2}}$. In above expressions, we used the partitions

$$U = \begin{bmatrix} U_{1:i} & U_{i+1:M} & U_{M+1:N} \end{bmatrix} \text{ and } V = \begin{bmatrix} V_{1:i} & V_{i+1:M} \end{bmatrix}.$$

Therefore, we can write $\Sigma_{ev}(z) = P_{ev}(z) J P_{ev}^*(z^{-*})$ where $J = \begin{bmatrix} I_{N+M} & 0 \\ 0 & -I_M \end{bmatrix}$ and

$$\begin{aligned} P_{ev}(z) &= F_2(z)F_5(z) \\ &= \begin{bmatrix} -\frac{h_0^T}{\sqrt{1-\gamma^2}} & -\frac{z^{-1}}{\sqrt{1-\gamma^2}} I & \sqrt{1-\gamma^2} I \\ B^{1/2}U_{i+1:M}^T (R^{1/2})^T & -(C^2 h_{m0} U_{i+1:M} B^{-1/2})^T z^{-1} & 0 \\ (R^{1/2}U_{M+1:N})^T & -(C^2 h_{m0}^* U_{M+1:N})^T z^{-1} & 0 \\ 0 & (V_{1:i} C A^{-1/2})^T & 0 \\ 0 & (V_{i+1:M} C B^{-1/2})^T & 0 \\ (R^{1/2}U_{1:i}A^{1/2})^T & (C h_{m0}^* U_{1:i} A^{-1/2})^T z^{-1} & 0 \end{bmatrix}^T. \end{aligned}$$

We note that $P_{ev}(z)$ contains the constant term $V_{i+1:M} C B^{-1/2}$ at the upper right $(N + M) \times M$ corner. We cannot remove this term by multiplying $P(z)$ from the right by a J-unitary matrix since the first two block entries of $P_{ev}(z)$ at the corresponding row are strictly causal and the other constant term $V_{1:i} C A^{-1/2}$ in the same row is orthogonal to $V_{i+1:M} C B^{-1/2}$. Therefore, we cannot convert $P_{ev}(z)$ to a matrix with a strictly causal

(1, 2) block by multiplication from the right. This implies that γ should be greater than the range assumed at the beginning.

2. In the previous part, we have seen that γ should be greater than or equal to $\frac{1}{1+\sigma_M^2(h_{m0})}$. We will pursue the factorization under this condition.
For this case, $P_{ev}(z)$ can be written as

$$P_{ev}(z) = \begin{bmatrix} -\frac{h_0}{\sqrt{1-\gamma^2}} & R^{1/2}U_{M+1:N} & 0 & R^{1/2}U_{1:M}A^{1/2} \\ -\frac{z^{-1}I}{\sqrt{1-\gamma^2}} & -\frac{\gamma^2}{1-\gamma^2}h_{m0}^*U_{M+1:N}z^{-1} & V\sqrt{\frac{\gamma^2}{1-\gamma^2}}A^{-1/2} & -\frac{\gamma^2}{1-\gamma^2}h_{m0}^*U_{1:M}A^{-1/2}z^{-1} \\ \sqrt{1-\gamma^2}I & 0 & 0 & 0 \end{bmatrix}.$$

Since $P_{ev}(z)$ doesn't have a top-right $N + M \times M$ strictly causal entry we need to find a J-unitary matrix to multiply $P_{ev}(z)$ from the right so that the resulting $L_{ev}(z)$ has strictly causal at that position. For that purpose, we first find

$$P_{ev}(\infty) = \begin{bmatrix} -\frac{h_0}{\sqrt{1-\gamma^2}} & R^{1/2}U_{M+1:N} & 0 & R^{1/2}U_{1:M}A^{1/2} \\ 0 & 0 & V\sqrt{\frac{\gamma^2}{1-\gamma^2}}A^{-1/2} & 0 \\ \sqrt{1-\gamma^2}I & 0 & 0 & 0 \end{bmatrix}.$$

Thus we want to find a J-unitary matrix Θ such that

$$P_{ev}(\infty)\Theta = \begin{bmatrix} W & 0 \\ V & Y \end{bmatrix}.$$

Since Θ is J-unitary

$$\begin{bmatrix} W & 0 \\ V & Y \end{bmatrix} J \begin{bmatrix} W^* & V^* \\ 0 & Y^* \end{bmatrix} = P_{ev}(\infty)JP_{ev}^*(\infty). \tag{39}$$

From this equality, we obtain

$$\begin{aligned} WW^* &= \begin{bmatrix} \frac{h_0h_0^*}{1-\gamma^2} + R^{1/2}(U_{M+1:N}U_{M+1:N}^* - U_{1:M}AU_{1:M}^*)R^{*/2} & 0 \\ 0 & \frac{\gamma^2}{1-\gamma^2}VA^{-1}V^* \end{bmatrix} \\ &= \begin{bmatrix} R + h_0h_0^* & 0 \\ 0 & (\frac{\gamma^2-1}{\gamma^2}I - h_{m0}^*h_{m0})^{-1} \end{bmatrix}, \end{aligned}$$

which is positive, and therefore,

$$W = \begin{bmatrix} (R + h_0 h_0^*)^{1/2} & 0 \\ 0 & (\frac{\gamma^2 - 1}{\gamma^2} I - h_{m0}^* h_{m0})^{-1/2} \end{bmatrix}.$$

Again using Equation 39,

$$VW^* = \begin{bmatrix} -h_0^* & 0 \end{bmatrix}.$$

Therefore $V = \begin{bmatrix} -h_0^*(R + h_0 h_0^*)^{-1/2} & 0 \end{bmatrix}$. Finally, since $VV^* - YY^* = (1 - \gamma^2)I$,

$$\begin{aligned} YY^* &= I\gamma^2 - I + h_0^*(R + h_0 h_0^*)^{-1} h_0 \\ &= -(I + h_0^* R^{-1} h_0)^{-1} + \gamma^2 I, \end{aligned}$$

which implies $Y = (\gamma^2 I - (I + h_0^* R^{-1} h_0)^{-1})^{1/2}$.

Using the above expressions we have for W, V and Y, we can obtain the J-unitary matrix Θ as

$$\begin{aligned} \Theta &= P_{ev}(\infty)^{-1} \begin{bmatrix} W & 0 \\ V & Y \end{bmatrix} \\ &= \begin{bmatrix} \frac{-h_0^*(h_0 h_0^* + R)^{-1/2}}{\sqrt{1-\gamma^2}} & 0 & \frac{(\gamma^2 I - (I + h_0^* R^{-1} h_0)^{-1})^{1/2}}{\sqrt{1-\gamma^2}} \\ U_{M+1:N}^* R^{-1/2}(h_0 h_0^* + R)^{1/2} & 0 & 0 \\ 0 & I & 0 \\ \Theta_{41} & 0 & \Theta_{43} \end{bmatrix} \end{aligned}$$

where

$$\Theta_{41} = A^{-1/2} U_{1:M}^* R^{-1/2}((h_0 h_0^* + R)^{1/2} - \frac{h_0 h_0^*}{1 - \gamma^2}(h_0 h_0^* + R)^{-1/2})$$

and

$$\Theta_{43} = A^{-1/2} U_{1:M}^* \frac{h_{m0}}{1 - \gamma^2}(\gamma^2 I - (I + h_0^* R^{-1} h_0)^{-1})^{1/2}.$$

If we apply this J-Unitary transformation to $P_{ev}(z)$ we obtain

$$L_{ev}(z) \quad = \quad P_{ev}(z)\Theta$$

$$= \begin{bmatrix} (h_0h_0^* + R)^{1/2} & 0 & 0 \\ z^{-1}h_0^*(h_0h_0^* + R)^{-1/2} & (\frac{\gamma^2-1}{\gamma^2}I - h_0^*R^{-1}h_0)^{-1/2} & L_{ev,23}(z) \\ -h_0^*(h_0h_0^* + R)^{-1/2} & 0 & L_{ev,33}(z) \end{bmatrix}$$

where

$$L_{ev,23}(z) = -z^{-1}\frac{(1-\gamma^2)I + \gamma^2 h_0^*R^{-1}h_0}{(1-\gamma^2)^2}(\gamma^2 I - (I + h_0^*R^{-1}h_0)^{-1})^{1/2}$$

and

$$L_{ev,33}(z) = (\gamma^2 I - (I + h_0^*R^{-1}h_0)^{-1})^{1/2}.$$

Therefore the resulting $L_{ev}(z)$ matrix has strictly causal L_{12} block. Besides, the L_{11} block is unimodular and causal and therefore causally invertible. Therefore, the desired form of factorization is achieved for

$$\gamma^2 \geq (1 + \sigma_{min}(h_{m0}^*h_{m0}))^{-1} = (1 + \sigma_{min}(h_0^*R^{-1}h_0))^{-1}$$

thus,

$$\gamma_{opt}^2 = (1 + \sigma_{min}(h_0^*R^{-1}h_0))^{-1}$$

Note that we obtain the transfer function $L(z)$ in Theorem 2 by

$$L(z) = F_1(z)L_{ev}(z).$$

Appendix B

We summarize the operator based approach in the formulation of the H^∞ problem based on the notation and the techniques introduced in [8]. This approach provides a simplified and alternative solution to the problems in certain cases.

Notation

Let the input-output rule for a linear time invariant system be given by the convolution expression

$$y_i = \sum_{j=-\infty}^{\infty} T_{i-j}u_j, \tag{40}$$

where $u = \{u_i\} \in l^{2,m}$, i.e., space of square-summable sequence of vectors with dimension m, and $y = \{y_i\} \in l^{2,p}$. If we define the mapping as $\mathcal{T}$, then the z-transform of the operator $\mathcal{T}$ can be written as

$$T(z) = \sum_{j=-\infty}^{\infty} T_i z^{-i}, \tag{41}$$

which is uniformly convergent and analytic on an annulus containing the unit circle, since $\mathcal{T}$ maps $l^{2,m}$ to $l^{2,p}$. Therefore the Fourier Transform $T(e^{j\omega})$ is well defined for all $\omega \in [0, 2\pi)$.

We partition the sequences u and y into their past, $u_- \triangleq \{u_i, i < 0\}$ and $y_- \triangleq \{y_i, i < 0\}$, and present and future, $\{u_+ \triangleq u_i, i \geq 0\}$ and $\{y_+ \triangleq y_i, i \geq 0\}$ components. This corresponds to the partitioning of $l^{2,m}$ and $l^{2,p}$ into orthogonal subspaces $l_-^{2,m}$ and $l_+^{2,m}$, and $l_-^{2,p}$ and $l_+^{2,p}$ respectively. Under this orthogonal partitioning of the input and output spaces, we can partition the operator $\mathcal{T}$ as

$$\mathcal{T} = \left[\begin{array}{c|c} \mathcal{T}_- & \mathcal{T}_A \\ \hline \mathcal{T}_H & \mathcal{T}_+ \end{array} \right]. \tag{42}$$

Here the operators of interest are:

- $\mathcal{T}$: *Laurent* operator, maps $l^{2,m}$ to $l^{2,p}$. Its H^∞ norm is defined as

$$\|\mathcal{T}\|_\infty \triangleq \sup_{u \neq 0 \in l^{2,m}} \frac{\|\mathcal{T}u\|_2}{\|u\|_2}. \tag{43}$$

- $\mathcal{T}_-$: *Toeplitz* operator, maps $l_-^{2,m}$ to $l_-^{2,p}$, i.e., past inputs to past outputs. Its H^∞ norm is defined as

$$\|\mathcal{T}_-\|_\infty \triangleq \sup_{u \neq 0 \in l_-^{2,m}} \frac{\|\mathcal{T}_- u\|_2}{\|u\|_2}. \tag{44}$$

- $\mathcal{T}_H$: *Hankel* operator, maps $l_-^{2,m}$ to $l_+^{2,p}$, i.e., past inputs to present and future outputs. Its H^∞ norm is defined as

$$\|\mathcal{T}_H\|_\infty \triangleq \sup_{u \neq 0 \in l_-^{2,m}} \frac{\|\mathcal{T}_H u\|_2}{\|u\|_2}. \tag{45}$$

We can also provide frequency domain characterization of the H^∞ norms of the Laurent and Toeplitz operators:

$$\|\mathcal{T}\|_\infty = \|\mathcal{T}_-\|_\infty = \sup_{\omega \in [0, 2\pi)} \sigma_{max} \left[T(e^{j\omega} \right]. \tag{46}$$

The Two-Block Problem

In this section, we will concentrate on the two block operator

$$\mathcal{T}_{\mathcal{K}} = \left[\begin{array}{cc} \mathcal{L} - \mathcal{K}\mathcal{H} & -\mathcal{K} \end{array} \right], \tag{47}$$

where $\mathcal{L}$ and $\mathcal{H}$ are causal Laurent operators, and $\mathcal{K}$ is a Laurent operator which is not necessarily causal. We are interested in $\mathcal{T}_{\mathcal{K}}$ since it is the error transfer operator that maps the input disturbances to the output estimation error in the general linear estimation setup.

We give the following theorems for the two-block problem[8]:

Theorem 3 Smoothing Problem *Consider the causal Laurent operators, $\mathcal{L}$ and $\mathcal{H}$, we would like to solve*

$$\gamma_s \triangleq \inf_{\mathcal{K}} \| \left[\begin{array}{cc} \mathcal{L} - \mathcal{K}\mathcal{H} & -\mathcal{K} \end{array} \right] \|_\infty. \tag{48}$$

Then we have

$$\gamma_s = \|\mathcal{L}(I + \mathcal{H}^*\mathcal{H})^{-1}\mathcal{L}^*\|_\infty = \sup_{\omega\in[0,2\pi)} \sigma_{max} \left[L(e^{j\omega}) \left(I + H^*(e^{j\omega})H(e^{j\omega}) \right)^{-1} L^*(e^{j\omega}) \right]. \tag{49}$$

Proof: Note that we may write

$$\mathcal{T}_{\mathcal{K}}\mathcal{T}_{\mathcal{K}}^* = (\mathcal{L} - \mathcal{K}\mathcal{H})(\mathcal{L} - \mathcal{K}\mathcal{H})^* + \mathcal{K}\mathcal{K}^*, \tag{50}$$

so that after a completion of squares,

$$\begin{aligned} \mathcal{T}_{\mathcal{K}}\mathcal{T}_{\mathcal{K}}^* &= \left(\mathcal{K} - \mathcal{L}\mathcal{H}^*(I + \mathcal{H}\mathcal{H}^*)^{-1} \right) (I + \mathcal{H}\mathcal{H}^*) \left(\mathcal{K} - \mathcal{L}\mathcal{H}^*(I + \mathcal{H}\mathcal{H}^*)^{-1} \right)^* \\ &\quad + \mathcal{L}(I + \mathcal{H}^*\mathcal{H})^{-1}\mathcal{L}^*. \end{aligned}$$

Therefore $\mathcal{T}_{\mathcal{K}}\mathcal{T}_{\mathcal{K}}^*$ is minimized for $\mathcal{K} = \mathcal{L}\mathcal{H}^*(I+\mathcal{H}\mathcal{H}^*)^{-1}$ which leads to $\mathcal{T}_{\mathcal{K}}\mathcal{T}_{\mathcal{K}}^* = \mathcal{L}(I+\mathcal{H}^*\mathcal{H})^{-1}\mathcal{L}^*$ and hence the desired result.

Theorem 4 Causal H^∞ Problem *Consider the causal Laurent operators, $\mathcal{L}$ and $\mathcal{H}$ and suppose we would like to solve*

$$\gamma_c \triangleq \inf_{causal\ \mathcal{K}} \| \left[\begin{array}{cc} \mathcal{L} - \mathcal{K}\mathcal{H} & -\mathcal{K}\| \end{array} \right] \|_\infty. \tag{51}$$

Then we have

$$\gamma_c = \|\mathcal{L}_-(I + \mathcal{H}_-^*\mathcal{H}_-)^{-1}\mathcal{L}_-^*\|_\infty \tag{52}$$

Proof: We refer to reference [8] for the proof.

It is an interesting fact that both γ_c and γ_s have similar structure; the only difference is that the Laurent operators in γ_s expression is replaced by Toeplitz operators in γ_c expression. Although $\mathcal{L}_-(I + \mathcal{H}_-^*\mathcal{H}_-)^{-1}\mathcal{L}_-^*$ is in terms of Toeplitz operators only, it is not necessarily Toeplitz and therefore a simple frequency domain formula for γ_c cannot generally be given. However, an alternative characterization of $\mathcal{L}_-(I + \mathcal{H}_-^*\mathcal{H}_-)^{-1}\mathcal{L}_-^*$ is given in the following theorem:

Theorem 5 Mixed Toeplitz-Plus-Hankel Operator *Consider the causal Laurent operators $\mathcal{L}$ and $\mathcal{H}$. Then we have*

$$\gamma_c^2 = \|\mathcal{E}_- + \mathcal{P}_{\mathcal{H}}^*\mathcal{P}_{\mathcal{H}}\|_\infty \tag{53}$$

where we have defined

$$\mathcal{E} = \mathcal{L}(I + \mathcal{H}^*\mathcal{H})^{-1}\mathcal{L}^*, \tag{54}$$

and

$$P = \Delta^{-1}\mathcal{H}\mathcal{L}^* \ , \ \Delta\Delta^* = I + \mathcal{H}\mathcal{H}^* \tag{55}$$

with Δ causal and causally invertible.

Proof : We refer to reference [8] for the proof.

We note that $\gamma_s^2 = \|\mathcal{E}_-\|_\infty$. Therefore, due to Eq. (53), $\gamma_c \geq \gamma_s$ and the increase depends on the Hankel operator $\mathcal{P}_{\mathcal{H}}$.

Appendix C

An explicit expression can be obtained for the $\gamma_{opt,dfe}$ value for the H^∞-DFE filter under the correct decisions assumption, i.e., $\epsilon = 0$. When $\epsilon > 0$, we can obtain an upper bound for $\gamma_{opt,dfe}$ as suggested by the following theorem:

Theorem 6 *Consider the $M \times N$ causal transfer matrix $H(z) = h_0 + h_1 z^{-1} + \ldots$ and suppose we are interested in the following problem*

$$\min_{causal K_V(.)} \| \left[\begin{array}{cc} (z^{-d}I - K_V(z)H_V(z))Q^{1/2} & -K_V(z)R_V^{1/2} \end{array} \right] \|_\infty = \gamma_{opt,dfe}^2$$

where

$$H_V(z) = \begin{bmatrix} H(z) \\ z^{-(d+1)} I \end{bmatrix} \qquad \text{and} \qquad R_V = \begin{bmatrix} rI & 0 \\ 0 & \epsilon I \end{bmatrix}.$$

1. *If $\epsilon = 0$, then*

$$\gamma^2_{opt,dfe} = \sigma_{max}(\begin{bmatrix} I_M & 0...0 \end{bmatrix} (I + H_d^* r^{-1} H_d) \begin{bmatrix} I_M & 0...0 \end{bmatrix}^T) \tag{56}$$

where H_d as defined in Equation 33.

2. *If $\epsilon > 0$, then we can show that*

$$\gamma^2_{opt,dfe} \leq \sigma_{max}(\begin{bmatrix} I_M & 0..0 \end{bmatrix} ((I + r^{-1} H_d^* H_d)^{-1} + \epsilon Z) \begin{bmatrix} I_M & 0...0 \end{bmatrix}^T), \tag{57}$$

where

$$Z = (rI + H_d^* H_d)^{-1} H_d^* h_d h_d^* H_d (rI + H_d^* H_d)^{-1}. \tag{58}$$

Proof: We follow the operator theory based methods described in [8]: We begin by defining the Toeplitz operator for the equivalent channel as

$$\mathcal{H}_V = \begin{bmatrix} \mathcal{H}_C^T & \mathcal{H}_D^T \end{bmatrix}^T \tag{59}$$

where $\mathcal{H}_C$ is the Toeplitz operator for the channel which is the semi-infinite matrix

$$\mathcal{H}_C = \begin{bmatrix} ... & . & . & . & . \\ ... & h_0 & 0 & 0 & 0 \\ ... & h_1 & h_0 & 0 & 0 \\ ... & h_2 & h_1 & h_0 & 0 \\ ... & h_3 & h_2 & h_1 & h_0 \end{bmatrix} \tag{60}$$

and $\mathcal{H}_D$ is the Toeplitz operator corresponding to the delay component of the equivalent channel which is another semi-infinite matrix given by

$$\mathcal{H}_D = \begin{bmatrix} I & 0_{\infty \times (d+1) \cdot M} \end{bmatrix}. \tag{61}$$

Furthermore, we also define the Toeplitz operator $\mathcal{L}$ for the delay operator $L(z)$ as

$$\mathcal{L} = \begin{bmatrix} I & 0_{\infty \times d \cdot M} \end{bmatrix}. \tag{62}$$

As shown in [8], in terms of the operators we defined above, the γ_{opt} is given by

$$\begin{aligned}\gamma_{opt}^2 &= \sigma_{max}(\mathcal{L}(I+\mathcal{H}_V^* R_{N_V}^{-1}\mathcal{H}_V)^{-1}\mathcal{L}^*) & (63)\\ &= \sigma_{max}(\mathcal{L}(I+r^{-1}\mathcal{H}_C^*\mathcal{H}_C+\epsilon^{-1}\mathcal{H}_D^*\mathcal{H}_D)^{-1}\mathcal{L}^*). & (64)\end{aligned}$$

We first note that

$$\mathcal{H}_D^*\mathcal{H}_D = \begin{bmatrix} I & 0_{\infty\times(d+1)\cdot M} \\ 0_{(d+1)\cdot M\times\infty} & 0_{(d+1)\cdot M\times(d+1)\cdot M}\end{bmatrix}. \quad (65)$$

We can partition $\mathcal{H}_C$ in a similar way as

$$\mathcal{H}_C = \begin{bmatrix} H_C & 0_{\infty\times(d+1)*M} \\ h_d & H_d \end{bmatrix} \quad (66)$$

where H_d as defined in Equation 33. We can write

$$\begin{aligned}(I+\mathcal{H}_C^*\mathcal{H}_C r^{-1}+\mathcal{H}_D^*\mathcal{H}_D\epsilon^{-1}) &= r^{-1}\begin{bmatrix} r(1+\epsilon^{-1})I+\mathcal{H}_C^*\mathcal{H}_C+h_d^*h_d & h_d^*H_d \\ H_d^*h_d & rI_{(d+1)*M}+H_d^*H_d\end{bmatrix}\\ &= \mathcal{F}\mathcal{M}\mathcal{F}^* \quad (67)\end{aligned}$$

where

$$\mathcal{F} = \begin{bmatrix} I & h_d^*H_d(rI_{d\cdot N}+H_d^*H_d)^{-1} \\ 0 & I_{(d+1)\cdot M}\end{bmatrix}, \quad (68)$$

and

$$\mathcal{M} = \begin{bmatrix} (1+\epsilon^{-1})I+r^{-1}\mathcal{H}_C^*\mathcal{H}_C+r^{-1}h_d^*(I_{d\cdot M}+r^{-1}H_dH_d^*)^{-1}h_d & 0 \\ 0 & r^{-1}H_d^*H_d+I_{d\cdot M}\end{bmatrix}. \quad (69)$$

Therefore,

$$\mathcal{L}(I+r^{-1}\mathcal{H}_C^*\mathcal{H}_C+\epsilon^{-1}\mathcal{H}_D^*\mathcal{H}_D)^{-1}\mathcal{L}^* = \mathcal{L}\mathcal{F}^{-*}\mathcal{M}^{-1}(\mathcal{L}\mathcal{F}^{-*})^*. \quad (70)$$

Here, we write $\mathcal{L}\mathcal{F}^{-*}$ as

$$\mathcal{L}\mathcal{F}^{-*} = \begin{bmatrix} I & 0 \\ -\begin{bmatrix} I_M & 0...0\end{bmatrix}(rI+H_d^*H_d)^{-1}H_d^*h_d & \begin{bmatrix} I_M & 0..0\end{bmatrix}\end{bmatrix}. \quad (71)$$

1. If $\epsilon \to 0$, then Equation 70 converges to

$$\mathcal{L}\mathcal{F}^{-*}\begin{bmatrix} 0 & 0 \\ 0 & (r^{-1}H_d^*H_d + I_{d*N})^{-1} \end{bmatrix}(\mathcal{L}\mathcal{F}^{-*})*, \tag{72}$$

therefore,

$$\gamma_{opt,dfe} = \sigma_{max}(\begin{bmatrix} I_M & 0...0 \end{bmatrix}(I + r^{-1}H_d^*H_d)^{-1}\begin{bmatrix} I_M & 0...0 \end{bmatrix}^T). \tag{73}$$

2. For $\epsilon > 0$, in expression 67, if we replace $\mathcal{M}$ with

$$\mathcal{M}_a = \begin{bmatrix} \epsilon^{-1}I & 0 \\ 0 & r^{-1}H_d^*H_d + I_{d*N} \end{bmatrix}, \tag{74}$$

since $M_a \leq M$ (and therefore $M_a^{-1} \geq M$), we obtain an upper bound for the γ_{opt}. If we look at the product $\mathcal{L}\mathcal{F}^{-*}\mathcal{M}_a^{-1}(\mathcal{L}\mathcal{F}^{-*})^*$, it is equal to

$$\underbrace{\begin{bmatrix} \epsilon I & -\epsilon\mathcal{X} \\ -\epsilon\mathcal{X}^* & \begin{bmatrix} I_M & 0..0 \end{bmatrix}(I + r^{-1}H_d^*H_d)^{-1}\begin{bmatrix} I_M & 0..0 \end{bmatrix}^T + \epsilon\mathcal{X}^*\mathcal{X} \end{bmatrix}}_{\mathcal{Y}} \tag{75}$$

where $\mathcal{X} = h_d^*H_d(rI + H_d^*H_d)^{-1}\begin{bmatrix} I_M & 0..0 \end{bmatrix}^T$.

Since we are interested in the eigenvalues of $\mathcal{Y}$, we look at $\lambda I - \mathcal{Y}$, which is congruent to

$$\begin{bmatrix} (\lambda-\epsilon)I & 0 \\ 0 & \lambda I - \begin{bmatrix} I_M & 0..0 \end{bmatrix}(I + r^{-1}H_d^*H_d)^{-1}\begin{bmatrix} I_M & 0..0 \end{bmatrix}^T - \epsilon\mathcal{X}^*\mathcal{X} - \frac{\epsilon^2}{\lambda-\epsilon}X^*X \end{bmatrix}. \tag{76}$$

It can be shown that the maximum eigenvalue is a concave function of ϵ; therefore the linear approximation obtained by ignoring the higher order terms provides an upper bound. Under the linear approximation Eq. (77) takes the form

$$\begin{bmatrix} (\lambda-\epsilon)I & 0 \\ 0 & \lambda I - \begin{bmatrix} I_M & 0..0 \end{bmatrix}(I + r^{-1}H_d^*H_d)^{-1}\begin{bmatrix} I_M & 0..0 \end{bmatrix}^T - \epsilon\mathcal{X}^*\mathcal{X} \end{bmatrix}. \tag{77}$$

Therefore, upperbound for the maximum eigenvalue of $\mathcal{Y}$ is given by

$$\sigma_{max}(\begin{bmatrix} I_M & 0..0 \end{bmatrix}(I + r^{-1}H_d^*H_d)^{-1}\begin{bmatrix} I_M & 0..0 \end{bmatrix}^T + \epsilon\mathcal{X}^*\mathcal{X}), \tag{78}$$

which yields

$$\gamma^2_{opt,dfe} \leq \sigma_{max}(\left[\begin{array}{cc} I_M & 0..0 \end{array} \right] ((I + r^{-1} H_d^* H_d)^{-1} + \epsilon Z) \left[\begin{array}{cc} I_M & 0..0 \end{array} \right]^T),$$

where

$$Z = (rI + H_d^* H_d)^{-1} H_d^* h_d h_d^* H_d (rI + H_d^* H_d)^{-1}. \tag{79}$$

References

[1] C. Belfore and J. Park, "Decision feedback equalization," *IEEE Proc.*, vol. 67, pp. 1143–1156, 1979.

[2] J.M.Cioffi, G.P.Dudevoir, M.V.Eyuboglu, and G.D.Forney, "Mmse decision-feedback equalizers and coding- parts i and ii," *IEEE Transactions on Communications*, vol. 43(10), pp. 2582–2604, October 1995.

[3] S.-C. Peng, "An equalizer design for nonminimum phase channel via two-block H^∞ optimization technique," *Signal Processing*, vol. 51, pp. 1–13, 1996.

[4] A. T. Erdogan, B. Hassibi, and T. Kailath, " Linear H^∞ Equalization of Communication Channels," *IEEE Trans. on SP.*, November 2000.

[5] J. Salz, "Optimum mean square decision feedback equalization," *Bell Syst. Tech. J.*, vol. 52, p. 1341, 1973.

[6] D. Falconer and G. Foschini, "Theory of mean square error QAM systems employing decision feedback equalization," *Bell Syst, Tech, J.*, vol. 52, pp. 1821–1849, 1973.

[7] B. Hassibi, "Indefinite metric spaces in estimation control and adaptive filtering," *Ph.D. Thesis*, August 1996.

[8] B. Hassibi, "On optimal solutions to two-block H^∞ problems," *Proceedings of American Control Conference*, 1998.

[9] H. Harashima and H. Miyakawa, "Matched transmission technique for channels with inter symbol interference," *IEEE Tran. Comm.*, vol. 29, pp. 774–780, 1972.

[10] B. Hassibi, A. 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.

[11] S. Altekar and N. Beaulieu, "Upper bounds to the Error Probability of Decision Feedback Equalization," *IEEE Trans. Inf. Theory*, vol. IT-39, pp. 145–156, 1993.