H^∞ EQUALIZATION OF COMMUNICATION CHANNELS

Alper T. Erdogan, Babak Hassibi and Thomas Kailath

Information Systems Laboratory, Electrical Engineering Department,
Stanford University,
CA 94305

ABSTRACT

As an alternative to existing techniques and algorithms, we investigate the merit of the H-infinity approach to the equalization of communication channels. We first look at causal H-infinity equalization problem and then look at the improvement due to finite delay. By introducing the risk sensitive property, we compare the average performance of the central H-infinity equalizer with the MMSE equalizer in equalizing minimum phase channels.

1. INTRODUCTION

Equalization is a well studied problem in the area of communications. It can be considered as a special case of an estimation problem with the data model generally described by a linear model of the type shown in Figure 1. The discrete data sequence $\{b_i\}$ passes through the linear time-invariant channel $H(z)$, which causes inter-symbol interference (ISI). The observation sequence $\{y_i\}$ is then formed by the addition of an unknown measurement disturbance $\{v_i\}$ with the output of the communication channel $H(z)$. The purpose is to design an equalizer $K(z)$ which estimates b_{i-d} from the observations $\{y_i\}$, where $d \geq 0$ represents a possible delay.
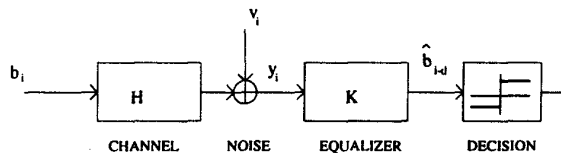

Figure 1: Linear Data Model

Various structures and methods have been proposed to recover transmitted data from their filtered and noise corrupted versions, and each method has its own advantages and disadvantages in terms of performance and complexity [5]. All these current techniques make some assumption about the underlying statistics and structure of the model. In many applications, however, the true information about the model is not available, and the algorithms use some estimate of the model parameters. For example, in mobile communications, the channel (and other statistical) parameters are often estimated from the observations through the use of certain

training sequences and therefore always contain errors. Time variation of the parameters and the errors due to tracking is another important issue. Therefore the question is whether small variations from the true model, and small disturbances, can cause large degradation in the performance of the algorithms. This brings us to the issue of "robustness".

In this paper, we address the robustness question by approaching the equalization problem from the H^∞ estimation,[4, 8, 7] point of view. The richness of robust H^∞ theory, and especially its stochastic interpretation of risk sensitive estimation, has been the basic motivation for our approach. Moreover, the availability of fast algorithms [4] is another major driving force for looking into H^∞ estimation as an equalization alternative. Finally, and perhaps most importantly, 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.

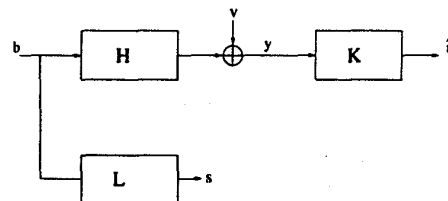## 2. H^∞ AND RISK SENSITIVE ESTIMATION

Figure 2: Setup for linear estimation

2.1. H^∞ Estimation

The basic setup for a general linear estimation problem is illustrated in Figure 2. In this setup we assume that $H(z)$ and $L(z)$ are causal linear time-invariant filters that map the input sequence $\{b_i\}$ to their respective outputs. The driving input $\{b_i\}$ and the additive disturbance sequence $\{v_i\}$ are assumed to be unknown. The estimation problem is to design a causal linear time-invariant estimator $K(z)$ that estimates s_i, the unobservable output of $L(z)$, using the observations $\{y_j, j \leq i\}$. We will denote such estimates by $\hat{s}_{i|i}$ and the resulting estimation errors by $\tilde{s}_{i|i} = s_i - \hat{s}_{i|i}$. Moreover, let $T_K(z)$ denote the transfer matrix that maps the unknown disturbances $\{b_i\}$ and $\{v_i\}$ to the estimation errors $\{\tilde{s}_{i|i}\}$.

Thus, $T_K(z) : \begin{bmatrix} b \\ v \end{bmatrix} \rightarrow \tilde{s}$. We may therefore write

$$T_K(z) = \begin{bmatrix} L(z) - K(z)H(z) & -K(z) \end{bmatrix} \qquad (1)$$

The choice of $K(z)$, and thereby the estimates $\hat{s}_{i|i}$, depends on our choice of performance criterion. In H^∞ estimation $K(z)$ is chosen to minimize the maximum error gain of $T_K(z)$, also known as the H^∞ norm of T_K defined as

$$\|T_K(z)\|_\infty^2 = \sup_{b,v \in l^2, b,v \neq 0} \frac{\sum_{i=-\infty}^\infty |\hat{s}_{i|i}|^2}{q^{-1} \sum_{i=\infty}^\infty |b_i|^2 + r^{-1} \sum_{i=-\infty}^\infty |v_i|^2}. \qquad (2)$$

Problem 1 (Optimal H^∞ Filtering Problem) *Find a causal estimator $K(z)$ that satisfies $\inf_{K(z)} \|T_K(z)\|_\infty^2$. Moreover, find the min-max energy gain γ_{opt}^2.*

There are very few cases where a closed-form solution to the optimal H^∞ filtering problem can be found, and in general one relaxes the minimization and settles for a suboptimal solution.

Problem 2 (Suboptimal H^∞ Filtering Problem) *Given $\gamma > 0$, find, if possible, a causal estimator $K(z)$ that guarantees*

$$\|T_K(z)\|_\infty^2 \leq \gamma^2 \qquad (3)$$

This clearly requires checking whether $\gamma > \gamma_{opt}$.

It will now be useful to give some flavor of the solution to Problem 2. (See [4] for more details). We introduce the following so-called *Popov function,*

$$\Sigma(z) = \begin{bmatrix} r + qH(z)H^*(z^{-*}) & -qH(z)L^*(z^{-*}) \\ -qL(z)H^*(z^{-*}) & -\gamma^2 + qL(z)L^*(z^{-*}) \end{bmatrix},$$

which can be regarded as a certain indefinite generalization of the spectral density function, $r + qH(z)H^*(z^{-*})$. Then a causal estimator, $K(z)$, that achieves $\|T_K(z)\|_\infty < \gamma$ exists if, and only if, the Popov function admits a canonical factorization of the form

$$\Sigma(z) = \begin{bmatrix} L_{11}(z) & L_{12}(z) \\ L_{21}(z) & L_{22}(z) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} L_{11}^*(z^{-*}) & L_{21}^*(z^{-*}) \\ L_{12}^*(z^{-*}) & L_{22}^*(z^{-*}) \end{bmatrix} \qquad (4)$$

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

$$K(z) = (L_{22}(z)Q(z) - L_{21}(z)) (L_{11}(z) - L_{12}(z)Q(z))^{-1}, \qquad (5)$$

where $Q(z)$ is any causal and strictly contractive operator, *i.e.,* $Q(z)$ is causal and is such that $|Q(e^{j\omega})|^2 < 1$, for all $\omega \in [0, 2\pi]$.

An important choice results from taking $Q = 0$, so that

$$K_{cen}(z) = -L_{21}(z)L_{11}^{-1}(z) \qquad (6)$$

which is the so-called "central" filter.

2.2. Risk Sensitive Estimation

Although the aforementioned H^∞ estimation formulation is a deterministic one, it has a nice stochastic interpretation which we now describe.

Given the basic model in Figure 2, in risk sensitive filtering, [6, 9], we assume that the disturbances $\{b_i\}$ and $\{v_i\}$ are stationary independent Gaussian random processes with variances q and r, respectively. The risk sensitive filtering problem is to find a causal $K(z)$ that minimizes

$$\mu(\theta) = \frac{2}{\theta} \log \left(E \exp \left(\frac{\theta}{2} \sum_{i=-\infty}^\infty |s_i - \hat{s}_i|^2 \right) \right), \qquad (7)$$

where $\theta > 0$ is known as the risk-sensitivity parameter.

The cost function 7 shows that as we increase the value of θ we put more penalty on large values of error as compared to the MMSE estimator, which minimizes $E(\sum_{i=-\infty}^\infty |s_i - \hat{s}_i|^2)$. However, the θ parameter can not be made arbitrarily large. In [2] it was shown that for any $\theta \leq \frac{1}{\gamma_{opt}^2}$, the causal $K(z)$ that minimizes the risk-sensitive cost function is given by the central H^∞ filter corresponding to $\gamma = \sqrt{\frac{1}{\theta}}$ and with energy weights q and r equal to the variances of $\{b_i\}$ and $\{v_i\}$, respectively. In the equalization application, only the large values of error, which are greater than the threshold for detection, are important. Errors below the threshold do not play a role. Therefore one may expect that the risk sensitive criterion is a good choice for equalization since it penalizes high errors more severely than the MMSE criterion. Nonetheless, as illustrated in future sections, it has certain drawbacks.

3. H^∞ EQUALIZERS

3.1. The Causal Case

The equalization problem of Section 1 is a special case of the linear estimation setup with $L(z) = z^{-d}$. When $d = 0$, the equalizer is constrained to be causal. In [3], the factorization (4) was explicitly obtained and thereby a characterization for all H^∞ equalizers was derived. The main results can be summarized as follows.

(i) If the channel $H(z)$ is *non-minimum phase*: We have

$$\gamma_{opt}^2 = q, \qquad (8)$$

which is the same energy gain obtained from $K(z) = 0$, *i.e.,* not equalizing at all! Therefore there is no hope for causally equalizing a non-minimum phase channel.

(ii) If the channel $H(z)$ is *minimum phase*: we have

$$\gamma_{opt}^2 = \frac{rq|L(e^{jw})|^2}{r + q|H(e^{jw})|^2} = \gamma_{opt,smoothing}^2, \qquad (9)$$

where $\gamma_{opt,smoothing}^2$ is the minimax energy gain for the "optimal" smoothing filter which turns out to be the celebrated Wiener smoother. This implies that one can perform as well as the non-causal (smoothing) solution!

In the minimum phase case, in [3], it is shown that the central equalizer is given by

$$K_{central}(z) = -L_{21}(z)L_{11}^{-1}(z) = \frac{h_0(1 - \gamma_{opt}^2)}{h_0 H(z) - (1 - \gamma_{opt}^2)R_\Delta \Delta(z)}. \qquad (10)$$

where the monic and minimum phase transfer function $\Delta(z)$ and the scalar R_Δ are found from the standard spectral factorization

$$R_\Delta \Delta(z)\Delta^*(z^{-*}) = \frac{\gamma^2}{1-\gamma_{opt}^2}H(z)H^*(z^{-*}) - r \geq 0. \quad (11)$$

We should also remark that another H^∞ optimal equalizer is $K(z) = (1-\gamma_{opt}^2)H^{-1}(z)$, which is simply a scaled version of the zero-forcing equalizer. Thus, an appropriately scaled zero-forcing equalizer is H^∞-optimal.

Comparison of the error spectra in Figure 3 illustrates that the smoother outperfroms all other equalizers at every frequency. The causal MMSE (H^2) filter has the best average performance among the causal equalizers, however the peak value of its spectrum is greater than the others. The risk sensitive and the scaled zero forcing equalizers have peak spectra equal to the smoother, but the risk sensitive one has better average performance.
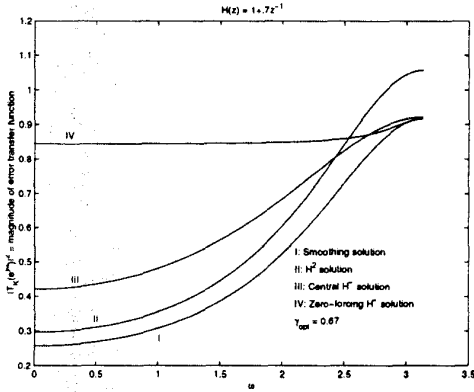

Figure 3: Comparison of the Error Spectra for the channel $H(z) = 1 + .7z^{-1}$

3.2. Finite Delay Case

It is clear from the previous section that we need non-causal equalizer structures to equalize non-minimum phase channels, which is equivalent to have finite delay, i.e. $d > 0$.

In order to illustrate the effect of delay, it will be instructive to look at the special case of equalizing the single zero channel $H(z) = 1 + az^{-1}, a \in R$.

We know from Section 3.1 that when $d = 0$,

$$\gamma_{opt}^2 = \begin{cases} \sup_{\omega \in [0,2\pi]} \frac{r}{r+|H(e^{j\omega})|^2} = \frac{1}{1+(1-|a|)^2} & \text{for } |a| < 1 \\ 1 & \text{for } |a| \geq 1 \end{cases}$$

Delay=1: After some algebraic manipulation, the Popov function for this case simplifies to

$$\Sigma(z) = \begin{bmatrix} 2 + a^2 + az^{-1} + az & -(z+a) \\ -(z^{-1}+a) & (1-\gamma^2) \end{bmatrix} \quad (12)$$

To attempt to factorize $\Sigma(z)$ as in Equation 4 consider the following factorization of the Popov function

$$\Sigma(z) = \begin{bmatrix} 1 & 0 \\ -\frac{1}{a} & 1 \end{bmatrix} \begin{bmatrix} 2 + a^2 + az^{-1} + az & \frac{2}{a} + z^{-1} \\ \frac{2}{a} + z & \frac{2}{a^2} - \gamma^2 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{a} \\ 0 & 1 \end{bmatrix}, \quad (13)$$

observe that the center matrix takes the form of the Popov function for the $d = 0$ case. Indeed by making the following substitutions, $\hat{H}(z) = -(a + \frac{a^2}{2}z^{-1})$, $\hat{L}(z) = 1 = z^0$, $\hat{r} = \frac{a^2}{2}$ and $\hat{q} = \frac{2}{a^2}$ we can rewrite the Popov function as

$$\begin{bmatrix} 1 & 0 \\ -\frac{1}{a} & 1 \end{bmatrix} \begin{bmatrix} \hat{r} + \hat{q}\hat{H}(z)\hat{H}^*(z^{-*}) & -\hat{q}\hat{H}(z)\hat{L}^*(z^{-*}) \\ -\hat{q}\hat{L}(z)\hat{H}^*(z^{-*}) & \hat{q}\hat{L}(z)\hat{L}^*(z^{-*}) - \gamma^2 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{a} \\ 0 & 1 \end{bmatrix} \quad (14)$$

It thus follows that we need to distinguish between the two cases where $\hat{H}(z)$ is minimum phase and where $\hat{H}(z)$ is non-minimum phase. When $\hat{H}(z)$ is minimum phase, which is the case for $|a| < 2$, we have

$$\gamma_{opt}^2 = \sup_{\omega \in [0,2\pi]} \frac{\hat{r}\hat{q}}{\hat{r} + \hat{q}|H'(e^{j\omega})|^2} \quad (15)$$

which after some simplification becomes

$$\gamma_{opt}^2 = \sup_{\omega \in [0,2\pi]} \frac{1}{1 + |H(e^{j\omega})|^2} = \gamma_{opt,smoothing}^2. \quad (16)$$

When $\hat{H}(z)$ is non-minimum phase, which is the case for $|a| \geq 2$,

$$\gamma_{opt}^2 = \hat{q} = \frac{2}{a^2}. \quad (17)$$

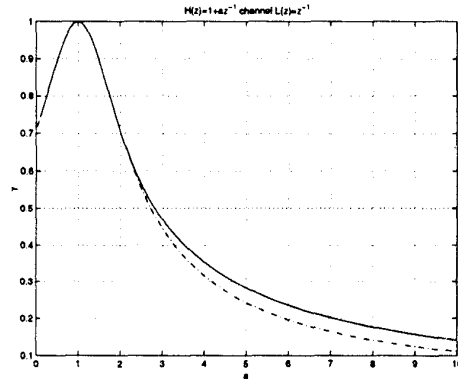Figure 4 shows the optimal value of γ as a function of a.

Figure 4: Optimal γ plot for single delay case. Dashed line refers to smoothing value

Delay $\geq$ 2: It can be similarly shown [1] that, for $d \geq 2$, $\gamma_{opt}^2 = \gamma_{opt,smoothing}^2$, i.e., a delay of two units is sufficient to obtain the same H^∞ performance as the smoother in equalizing a single-zero channel.

Unfortunately, there is no known explicit factorization for arbitrary $d > 0$ and for general non-minimum phase channels. However, in [3] it has been shown that in order to get an improvement over $\gamma_{opt} = 1$, the delay d should be chosen greater than the number of non-minimum phase zeros of the channel, i.e., the number of zeros outside of the unit circle.

4. COMPARISON OF CENTRAL H^∞ AND H^2 EQUALIZERS IN EQUALIZING MINIMUM PHASE CHANNELS

In this section of the paper, we will compare the central H^∞ and the H^2 equalizers in terms of average BERs. Figure 5 shows that the H^2 and central H^∞ equalizers have the same average BER performances for $H(z) = 1 + 0.5z^{-1}$.

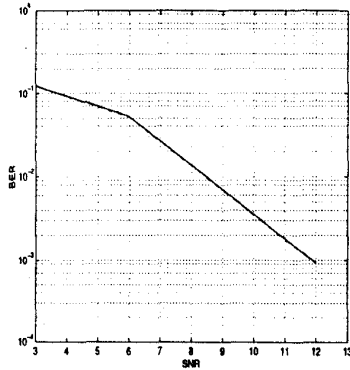
Figure 5: BER vs SNR curve for the channel $H(z) = 1 + 0.5z^{-1}$: Solid line-Risk Sensitive Equalizer, Dashed Line-MMSE Equalizer

However as illustrated in Figure 6, the H^2 equalizer has better average BER performance than the central H^∞ equalizer for the channel $H(z) = 1 + 0.95z^{-1}$.
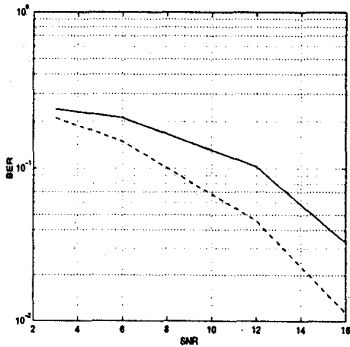

Figure 6: BER vs SNR curve for the channel $H(z) = 1+0.95z^{-1}$: Solid line-Risk Sensitive Equalizer, Dashed Line-MMSE Equalizer

In general, for the various channels that we have studied, when $H(z)$ and the signal statistics are known exactly, the risk sensitive and MMSE equalizers have either similar BERs, or the MMSE equalizer outperforms the risk sensitive one. It thus may appear that, there is no gain in using central H^∞ equalizers, compared to MMSE ones, in the ideal setup. However, in the face of model

uncertianity and lack of statistical knowledge, it is expected that the H^∞ equalizer will have acceptable performance H^2 equalizer [1].

5. CONCLUSION

In this paper, we introduced the H^∞ criterion as an alternative method for the equalization of communication channels, which concentrates on the worst case performance whereas the previous algorithms concentrate on the average performance. We showed that causal H^∞ filter has same worst-case performance as the non-causal smoothing filter in equalizing minimum phase channels. For non-minimum phase channels we need a number of delays at least equal to the number of non-minimum phase zeros of the channel. Therefore, study of H^∞ estimation provides us with a rigorous basis for the importance of the concepts of minimum phase channels and delay in the equalization problem.

We looked at the central H^∞, or risk sensitive, filter as a choice which has good average performance besides its optimal worst case performance . We showed that it has good average properties due to its stochastic interpretation although it does not appear to be better than the MMSE equalizer in terms of BER in the ideal case.

Formulation of equalizers for more general case and the performance under modeling errors is the area that we are currently pursuing.

6. REFERENCES

[1] A. T. ERDOGAN, B. HASSIBI, AND T. KAILATH, *On H^∞ Equalization of Communication Channels*, submitted to IEEE Tr. SP., (1997).

[2] K. GLOVER AND J. DOYLE, *State-space formulae for all stabilizing controllers that satisfy an H^∞ norm bound and relations to risk sensitivity*, System and Control Letters, 11 (1988), pp. 167–172.

[3] B. HASSIBI, A. T. ERDOGAN, AND T. KAILATH, *Equalization with an H^∞ Criterion*, submitted to IEEE Tr.Inf.Theory, (1997).

[4] B. HASSIBI, A. SAYED, AND T. KAILATH, *Indefinite Quadratic Estimation and Control: A Unified Approach to H^2 and H^∞ Theories*, SIAM Studies in Applied Mathematics, New York, 1998.

[5] J.G.PROAKIS, *DIGITAL COMMUNICATIONS*, McGRAW-HILL, 1995.

[6] J.L.SPEYER, J. DEYST, AND D. JACOBSON, *Optimization of stochastic linear systems with additive measurement and process noise using exponential performance criteria*, IEEE Transactions on Automatic Control, 19 (1974), pp. 358–366.

[7] P. KHARGONEKAR AND K. NAGPAL, *Filtering and smoothing in an H^∞ setting* , IEEE Transactions on Automatic Control, 36 (1991), pp. 151–166.

[8] U. SHAKED AND Y. THEODOR, *H^∞ optimal estimation: A tutorial*, In Proceedings of the IEEE Conference on Decision and Control, (1992), pp. 2278–2286.

[9] P. WHITTLE, *Risk Sensitive Optimal Control*, John & Wiley Sons, New York, (1990).