A GEOMETRICAL VIEW OF BLIND EQUALIZATION

Z. Gu and W. A. Sethares

Department of Electrical and Computer Engineering,
University of Wisconsin, Madison, WI 53706-1691 USA.

ABSTRACT

This paper presents a geometrical analysis of the equalization problem. The input signal forms a $(m+L)$-Dimensional hypercube that is mapped via a convolution matrix to a m-D parallelotope, where L is the order of the channel and m is the order of the transversal equalizer. The properties of this mapping are discussed, and a criterion for equalization called the *minimum width criterion* is proposed. Virtually all of the standard equalizer schemes can be viewed as special cases of this minimum width criterion, including the L_∞, L_1, MSE, LS, Sato, Godard, and kurtosis methods. It is possible to build "new" equalization algorithms by combining the basic distance elements uncovered by this geometric anaylsis.

1. INTRODUCTION

Equalization is used to reduce or cancel the ISI (InterSymbol Interference) caused by data communication channels. Such channels are often modeled as FIR (Finite Impulse Response) filter in which the term with the main (or reference) tap represents the desired signal, and the summation of all other terms represents the ISI.

There are many different approaches to equalization (surveys can be found in [1, 2]), which are derived as optimization problems with various criteria. This paper interprets these criteria geometrically, and shows that most of the equalization approaches can be derived from a single geometrical criterion which we call the minimum width criterion. Based on the geometric analysis, it is also possible to derive new approaches that combine the strengths of the various criteria.

2. PROBLEM DESCRIPTION

For simplicity and clarity, the input is assumed to be ± 1, though the present analysis extends to any M-level PAM or QAM signals. The equalizer is assumed to be a transversal (linear FIR) filter, and the channel is modelled as the L_{th} order system

$$y_k = a_o u_k + \sum_{i=1}^{L} a_i u_{k-i} \qquad (1)$$

A linear equalizer concatenated with the channel is chosen to be an approximation to the inverse of the channel so that the impulse response of the pair is close to a δ function. The transversal equalizer, which is assumed to be of fixed length m, can be interpreted as a $(m-1)$-Dimensional hyperplane passing through the origin in a m-D signal space. If this hyperplane can separate the two groups of points which represent the different input values $u_{k-d} = \pm 1$ for at least one value of d (where d is some delay), then the equalizer can "open the eye," or remove the ISI sufficiently to reconstruct the original signal. Suppose the m equalizer taps are $q_i, i = 1, ..., m$. Then the recovered signal $\hat{u}_{k-d}$ is equal to $\sum_{i=1}^{m} q_i y_{k+i-1}$. Thus, for an L_{th} order channel and its equalizer with length m, each recovered signal is a function of $m + L$ sequencial input signals. All these possible $(m + L)$-tuples compose a binary $(m + L)$-cube which we call the "input cube." The next two sections discuss properties of the input cube and how it is mapped through the channel.

3. THE BINARY N-CUBE (BNC) PROPERTIES

In order to recover the input signal, it is important to understand the structure of the input first.

BnC1: A binary n-cube has 2^n vertices and each vertex can be represented by an n-tuple $(u_1, u_2, ..., u_n)$, where $u_i \in \{-1, 1\}, i = 1, 2, ..., n$. More generally, it has a total of $2^{n-k}\binom{n}{k}$ k-faces, where a k-face means a k-dimensional boundary surface of the n-cube, $k = 0, 1, ...n$. All k-faces (except 0-faces) are either orthogonal or parallel each other.

BnC2: A k-face can be represented as an n-tuple with k x's and $n-k$ u's, where "x" represents a "don't care" bit. Thus, the number of x bits represents the dimension of the face. The "don't care" bits x contain information about the orientation of the k-face, while the constant bits u contain the position information of the k-face. Accordingly, there are a total of $\binom{n}{n-k} = \binom{n}{k}$ different orientations of k-faces in an n-cube. Along each orientation, there are 2^{n-k} different positions that the k-faces can assume.

Comment:

Although the above properties are aimed at signals with binary inputs, they can be easily extended to any M-level n-cube. Observe that the convex cover, (which plays an important role in the geometrical analysis of polytopes), of an M-level n-cube is also a binary n-cube assuming that the input is bounded.

4. THE DIMENSIONALLY DEGENERATE MAPPING (DDM) PROPERTIES

This section describes how the input cube is mapped into a m-D signal space through the channel (1). If we view

0-7803-0946-4/93 $3.00

the m equations of the channel (1) as a function, then its domain is the multi-dimensional input $(m+L)$-cube, and its range is the m-D signal space, where L is the order of the channel and m is the length of the equalizer. This mapping is dimensionally degenerate if $L > 0$ and it can be written in matrix form as

$$\mathbf{y}_k = A\mathbf{u}_k \tag{2}$$

where $\mathbf{y}_k$ is the m-D vector composed of the received signal sequence from y_k to y_{k+m-1}, $\mathbf{u}_k$ is the $(m+L)$-vector composed of the input sequence from u_{k-L} to u_{k+m-1}, and A is the convolution matrix. The DDM properties are summarized as follows.

DDM1: All sub-dimensional faces of the input $(m+L)$-cube with dimension $k \leq m$ are mapped into the m-D value space without any dimensional degeneration. Orthogonality, however, is destroyed. Thus the original orthogonal cubical k-faces in the input space become k-D parallelotopes in the m-D output space. All exterior subfaces of the k-faces are mapped to exterior faces of the parallelotope.

DDM2: Since a m-D space can only accommodate a m-D object, if $k > m$, the k-face of the input cube is mapped onto a m-D parallelotope. The convex cover of the image of the k-face is composed of the images of all its parallel pairs of $(m-1)$-faces which have different orientations and maximum distance to the parallelotope's geometrical center. Thus, the parallelotope has a total of $2\begin{pmatrix} k \\ m-1 \end{pmatrix}$ $(m-1)$-faces. The effect of this dimensionally degenerate mapping is that only some of the exterior faces are mapped to exterior faces of the parallelotope.

DDM3: The $(m-1)$-faces of the image of the whole $(m+L)$-D input cube with maximum distance to the origin of the m-D signal space are orthogonal to the axis along which the maximum distance occurs. This distance is the magnitude of the minimum and maximum values of the received signal y_k in (1). Thus, the supporting $(m-1)$-D hyper planes of these faces form a m-cube with a width of $2|y|_{max}$, where $|y|_{max}$ is the maximum magnitude of the received signal. In another words, the whole image of the input $(m+L)$-cube, which is a m-D parallelotope, is exactly held in a m-cube with width $2|y|_{max}$ in the m-D signal space.

Fig. 1 shows an example of the 2-D image of a 4-D input cube mapped through a second order nonminimum phase channel.

5. THE $(M-1)$-D SEPARATING PLANE AND THE MINIMUM WIDTH CRITERION

The properties of the input cube and its image after mapping through the L_{th} order channel (1) can be used to describe geometrically the properties of desirable equalizers. Consider the images of the two input $(m+L-1)$-faces defined by $u_{k-d} = \pm 1$. By DDM2, the effect of the dimensionally degenerate mapping is to map the two $(m+L-1)$-faces of the input cube defined by $u_{k-d} = +1$ and $u_{k-d} = -1$ onto pairs of m-D parallelotopes with $\begin{pmatrix} m+L-1 \\ m-2 \end{pmatrix}$ different widths.

The goal of a transversal equalizer is to find a $(m-1)$-D hyper plane $\mathbf{q}^T Y_k = 0$ which passes through the origin and separates these two m-D parallelotopes, where Y_k represents the m variables in the m-D signal space. The coefficients of the equalizer $\mathbf{q}$ must be scaled so that the recovered signal $\hat{u}_{k-d} = \mathbf{q}^T\mathbf{y}_k$, where $\mathbf{y}_k$ is the recieved signal regressor vector. Clearly, the process of finding this separating plane is dependent on the boundary $(m-1)$-faces of the two parallelotopes. We thus may expect that the best equalization will occur when the separating hyper plane is parallel to the sides of the parallelotope having minimum width. Intuitively, the direction in which the image of $u_{k+d} = 1$ (or -1) looks the "thinnest," will be the best direction for the separating hyper plane. We thus call this the ***minimum width criterion.***

It is computationally expensive to find the exterior $(m-1)$-face pair of the image of $u_{k-d} = 1$ (or -1) with the minimum width directly since it involves searching $\begin{pmatrix} m+L-1 \\ m-2 \end{pmatrix}$ pairs of different $(m-1)$-D faces. An easier way is to consider the distance between each of the $(m-1)$-faces and the separating hyper plane. Since only one of the $(m-1)$-face pairs maps to an exterior $(m-1)$-face of the image of the whole input cube, this suggests two different methods to obtain the optimal separating plane.

5.1. THE MIN-MAX (L_∞) APPROACH

Consider the exterior $(m-1)$-faces of the image of $u_{k-d} = 1$ (or -1). The minimum width criterion implies that the maximum distance from all signal points to the separating plane should be minimized. This gives the min-max distance criterion, which is also known as the L_∞ criterion. The distance from a point $\mathbf{y}_k$ to a plane $\mathbf{q}$ passing through the origin is equal to $|\mathbf{q}^T\mathbf{y}_k|/\| q \|_2$. Thus, the min-max distance is in the form of $min(max|\mathbf{q}^T\mathbf{y}_k|/\| q \|_2)$ which, by normalizing q to unity, can be rewritten as

$$\begin{aligned} &min\ (max \quad |\mathbf{q}^T\mathbf{y}_k|) \\ &\qquad s.t. \quad \| q \|_2 = 1 \end{aligned} \tag{3}$$

Note the criterion (3) does not require knowledge of the channel, nor of the input in order to determine the separating hyperplane. Thus, it is an ideal criterion for blind equalization. An alternative way to set the constraint (note if there were no constraint, the solution would be trivial) is to let one of the equalizer's taps be fixed at 1, which gives the modified min-max distance criterion

$$\begin{aligned} &min\ (max \quad |\mathbf{q}^T\mathbf{y}_k|) \\ &\qquad s.t. \quad q_j = 1 \end{aligned} \tag{4}$$

where j is indexed from 1 to m.

This L_∞ criterion which comes directly from a geometrical consideration of the problem is identical to the criterion in [3] which was derived from a L_1 criterion using the theory of [4].

Since $\mathbf{y}_k = A\mathbf{u}_k$, and $max|\sum u_i| = max\sum|u_i|$ for all possible $\mathbf{u}_k$, from the min-max distance criterion (3, 4), we have

$$min(max \quad |\mathbf{q}^T A\mathbf{u}_k|) \Rightarrow$$

$$\begin{aligned} & min \quad \|\mathbf{q}^T A\|_1 \\ = \; & min \quad \sum_{i=-L}^{m} |s_i| \end{aligned} \tag{5}$$

where $\{s_i\}$ is the impulse response of the convolution of the channel $\{a_i\}$ and the equalizer $\{q_i\}$. This turns the min-max distance criterion into a L_1 minimization problem. As in the above analysis, a constraint must be added to avoid a trivial solution. Fixing an equalizer tap at unity gives the same solution as in (4), while setting the constraint to $s_o = 1$, where s_o is the center tap of the system, gives the peak distortion criterion as in [5].

Comments:

1. Both the L_1 and L_∞ criteria are equivalent. While solving (5) requires knowledge of the channel, solving (3, 4) can be done blindly, i.e. without the knowledge of the channel or the input. Both have to be solved by LP (Linear Programming) whose solution is obtained by selecting and solving a set of m independent equations from a large set of equations.
2. In order to avoid LP, an iterative approximation L_p scheme was presented in [3] which could be applied to recieved signals only. When $p = 2$, the L_2 algorithm can equalize minimum (or maximum) phase channels only [6]. The L_2 solution is symmetric when the reference tap is set at the center. This implies that the L_2 algorithm is bound to fail to open the eye on strongly nonminimum phase channels.
3. The maximum kurtosis criterion presented in [7] results in a L_4 algorithm under certain conditions. Similar arguments to these for the L_2 algorithm show that the algorithm cannot equalize nonminimum phase systems unless the channel is well behaved or a pre-whitening filter is applied to the recieved signal.

5.2. THE MAX-MIN APPROACH

The previous section shows that the minimum width criterion leads to a min-max distance criterion by considering the outside exterior $(m-1)$-face of the image of $u_{k-d} = 1$ only. On the other hand, considering the $(m-1)$-face which is not an exterior face of the image of the input cube, (and if the eye is open at the choice of delay d), then the minimum width criterion also implies that the minimum distance from all the points to the separating plane should be maximized. This imples that the separating plane is parallel to one of the inside $(m-1)$-faces. Similar arguments lead to the max-min distance criterion:

$$max \; (min|\mathbf{q}^T \mathbf{y}_k| / ||q||_2)$$

which is equivalent to

$$\begin{aligned} max \; (min \quad & |\mathbf{q}^T \mathbf{y}_k|) \\ s.t. \quad & ||\mathbf{q}||_2 = 1 \end{aligned} \tag{6}$$

This can also be rewritten in terms of the inputs as

$$\begin{aligned} max \; (min \quad & |\mathbf{q}^T A \mathbf{u}_k|) \\ s.t. \quad & ||\mathbf{q}||_2 = 1 \end{aligned} \tag{7}$$

The constraint on the equalizer taps q can also be chosen to be a single fixed tap constraint as in (4). Note that (6) is related to the recieved signals only, so it can also be used for blind equalization. Since this distance (6) is closest to the separating plane, if it is maximized, its solution will often have a better bit error rate than those of min-max distance solutions. Such examples are not difficult to find.

Comments:

1. The max-min distance criterion and the min-max distance criterion can be considered a dual pair, and their solutions are not, in general, the same. The reason for this is the discreteness of the FIR model. However, as the equalizer length m increases, the two solutions converge, that is, the two sides of the $(m-1)$-faces approach the pair of parallel exterior $(m-1)$-faces with minimum width.
2. The maximum kurtosis criterion in [7] can also be geometrically interpreted as a max-min criterion by removing the restriction on the average power constraint. In this case, if the kurtosis of the input is less than zero (and the input is real), the criterion can be rewritten as

$$\begin{aligned} max \; K(z_i) \quad & = \quad 2E^2\{|z_i|^2\} + |E\{z_i\}^2|^2 - E\{|z_i|^4\} \\ & = \quad 2E^2\{|z_i|\}^2 - E[|z_i|^2 - E\{|z_i|^2\}]^2 \\ s.t. \quad & ||q||_2 = 1 \end{aligned}$$

where $|z_i| = |\mathbf{q}^T \mathbf{y}_k|$ is the distance from a signal point to the separating plane under the constraint. In the second equation, the first term is a measure of the summation of the distance from the geometric centers of the images of $u_{k-d} = \pm 1$ to the separating plane. The second term is a measure of the differences of the distance from the two sides of the image to the geometric center, where all distances are averaged through the expectation operator. Thus, the kurtosis criterion results in a measure of distance from the two inside $(m-1)$-faces of the image of $u_{k-d} = \pm 1$, which is then maximized.

5.3. THE MEAN SQUARED ERROR (MSE) APPROACH

An alternative way to view the minimum width criterion is to consider the separating plane in the input domain. Let $u_{k-d} = 0$ be the separating plane in the $(m+L)$-D input space contained inside the input cube. By the DDM properties and the linearity, this bounded separating plane is mapped onto a parallelotope centered at the origin that lies between the images of $u_{k-d} = \pm 1$. Thus, if the two images of $u_{k-d} = \pm 1$ are separable, then the best equalizer can be obtained by letting the image of the separating plane $u_{k-d} = 0$ in the input cube degenerate to a $(m-1)$-D plane along its minimum width direction. This results in the MSE approach to blind equalization.

The images of $u_{k-d} = 1, -1$ and 0 are shifted versions of each other. Applying a plane fitting scheme to all the points of the image of $u_{k-d} = 1$ (or -1) only, results in a plane which has the same normal direction as the separating plane that passes through the geometric center of the image of $u_{k-d} = 1$ (or -1). One representation of an arbitrary plane is

$$c_1 x_1 + c_2 x_2 + \ldots + c_n x_n = 1 \tag{8}$$

Thus, the separating plane can be obtained by solving the system of equations

$$\mathbf{q}^T\mathbf{y}_k = 1 \tag{9}$$

for all recieved signal regressors $\mathbf{y}_k$ corresponding to $u_{k-d} = 1$ (or -1).

The system of equations (9) can be solved iteratively or in block form by minimizing the cost function $\sum(\mathbf{q}^T\mathbf{y}_k - 1)^2$ for all y_k. Thus the minimum width criterion leads to the MSE solution for the separating plane which, under reasonable conditions on the input signal, is unique up to multiplication by an unknown gain. This gain is easy to deal with as mentioned in [6].

Observe that in (9), only half of the y_k are used. Modifying (9) to

$$|\mathbf{q}^T\mathbf{y}_k| = 1 \tag{10}$$

allows the use of all the recieved data, and the cost function becomes $\sum(|\mathbf{q}^T\mathbf{y}_k| - 1)^2$, which is exactly the cost function of Sato's algorithm [8] for binary inputs.

Comments:

1. Sato's algorithm performs a gradient descent using the cost function (10). This may also be solved in block (Least Squares) form, and this LS solution will be the same (up to an unknown gain) as the well known LS solution of $A^T\mathbf{q} = 1_d$, which requires that the channel be known, where 1_d is the zero vector except for a 1 at the d_{th} position, and A is the convolution matrix.
2. The plane fitting of the MSE approach leads to a separating plane that is in general not parallel to any of the $(m-1)$-faces of the m-D parallelotope. This results in the performance difference between the MSE solution and the min-max and max-min solutions. It is not easy to chose among the various criteria since the system performance is dependent on details of the channel, the equalizer length and the distributions of the inputs and disturbances.
3. Godard's algorithm [9] is a variant of (10) in which the $|\mathbf{q}^T\mathbf{y}_k|$ term is raised to the p power. This does not fundamentally change the geometric meaning of the cost criterion.

6. CONCLUSION

All the main existing linear transversal equalization schemes can be viewed as special cases of the minimum width criterion of this geometric analysis. Since we have succesfully interpreted the kurtosis criterion (which is a 4_{th} order cumulant method) we expect that other high order cumulant algorithms can also be interpreted in the same framework, though this investigation is underway. The minimization criteria have used four basic distance measures: the distance from the outside exterior of the images $u_{k-d} = \pm 1$ to the separating plane, the distance from the inside exterior of the images $u_{k-d} = \pm 1$ to the separating plane, the distance from the geometric center of the image to the separating plane, and the width of the image of the separated parallelotope. These four distances are not independent. Using these distance measures as building blocks may lead to numerous "new" criteria and algorithms for equalization. Our goal is to explore this space of algorithms with an eye toward designing optimal algorithms for particular environments. The geometric analysis gives a way to view the universe of possible equalization algorithms from a single vantage point.

7. REFERENCES

[1] S. U. H. Qureshi, "Adaptive equalization", *Proc. IEEE*, Vol. 73, No.9, pp 1349-87, 1985.

[2] J. G. Proakis, "Adaptive equalization techniques for acoustic telemetry channels," *IEEE J. of Oceanic Engr.*, Vol. 16, No.1, Jan. 1991, pp 21-31.

[3] W. A. Sethares, R. A. Kennedy and Z. Gu, "An approach to blind equalization of non-minimum phase systems," *ICASSP*, Toronto, Canada, May 1991.

[4] C. A. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*, Academic Press, New York, NY, 1975.

[5] R. W. Lucky, "Techniques for adaptive equalizations of digital communication systems," *Bell Syst. Tech. J.*, Vol. 45, pp. 255-286, Feb. 1966.

[6] S. Verdu, B. D. O. Anderson and R. A. Kennedy, "Blind Equalization Without Gain Identification," *IEEE Trans. on Commun.*, Jan. 1993.

[7] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. on Info. Theory*, Vol. IT-36, No.2, pp.312-321, March 1990.

[8] Y. Sato, "A method of self-recovering equalization for multilevel amplitude-modulation systems," *IEEE Trans. Commun.*, Vol. COM-23, pp.593-619, June 1975.

[9] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, Vol. COM-28, pp 1867-1875, Nov. 1980.

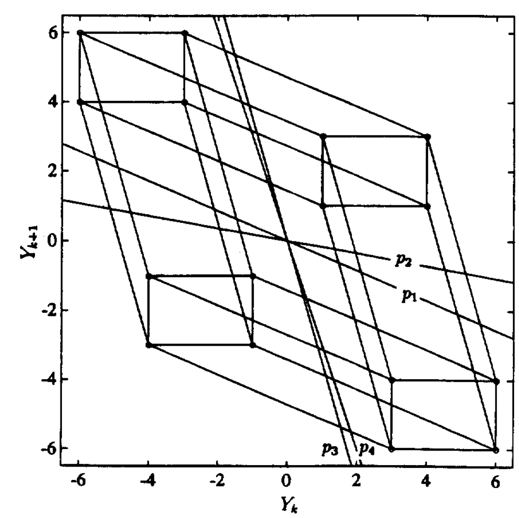

Fig. 1 The image of the 4-D input cube mapped through the second order nonminimum phase channel $y_k = u_k - 3.5u_{k-1} + 1.5u_{k-2}$. p_1 and p_2 are the L_∞ and MSE solutions respectively for $\hat{u}_{k-1}$ while p_3 and p_4 are the L_∞ and MSE solutions for $\hat{u}_{k-2}$.