BLIND EQUALIZATION OF CONSTANT MODULUS SIGNALS VIA RESTRICTED CONVEX OPTIMIZATION

Boris Maričić, Zhi-Quan Luo[*] *and Timothy N. Davidson*

Department of Electrical and Computer Engineering, McMaster University
Hamilton, Ontario, Canada, L8S 4K1.

ABSTRACT

In this paper, we formulate the blind equalization of *Constant Modulus* (CM) signals as a convex optimization problem. This is done by performing an algebraic transformation on the direct formulation of the equalization problem and then restricting the set of design variables to a subset of the original feasible set. In particular, we express the blind equalization problem as a linear objective function subject to some linear and semidefiniteness constraints. Such *Semidefinite Programs* (SDPs) can be efficiently solved using interior point methods. Simulations indicate that our method performs better than the standard methods, whilst requiring significantly fewer data samples.

1. INTRODUCTION

Conventional equalization and carrier recovery algorithms generally require an initial training period during which a known data sequence is transmitted and synchronized at the receiver. In the case of highly non-stationary communications environments (e.g., digital mobile communications), it may be preferable to equalize the communication channel in an unsupervised manner. The resulting operation is referred to as *blind equalization*. Many digital communications schemes involve the transmission of *Constant Modulus* (CM) signals, hence several schemes for blind equalization of CM signals have been developed. Typically, they are based on gradient descent minimization of a specially designed cost function [1–3]. However, these algorithms can experience undesirable local convergence problems which may result in insufficient removal of channel distortion [4, 5]. Here we formulate the problem of blind equalization of CM signals as a convex optimization problem which has a unique global minimum. We compare our method with the standard blind adaptive equalization methods in [1–3]. (These standard methods are not globally convergent for arbitrary initialization [4, 5]). Simulation results indicate that our method performs better than the standard methods even when we use some a priori knowledge of the channel impulse response to aid initialization of the standard methods. In addition, our method requires significantly fewer samples, which might be useful in applications where convergence times requiring thousands of input samples are undesirable.

2. PROBLEM STATEMENT

The receiver structure we study is shown in Fig.1, where $a(k)$ represents the CM signal and $v(k)$ is additive white Gaussian noise.

[*]Email: luozq@mcmaster.ca

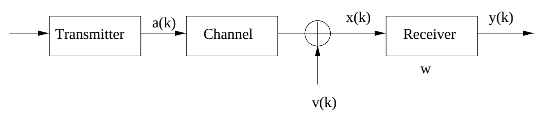

Figure 1: The block diagram of the system

The output of the equalizer $y(k)$ can be expressed as

$$y(k) = \mathbf{w}^H \mathbf{x}_k, \tag{1}$$

where

$$\mathbf{x}_k = [x(k), x(k-1), \ldots, x(k-n+1)]^T, \tag{2}$$

$\mathbf{w} \in \mathbb{C}^n$ is a weight vector and n is the length of the equalizer. If perfect equalization is achieved, the sequence $y(k)$ is also of the CM type. With that in mind, a natural optimization problem for the receiver to solve is [1]

$$\text{minimize} \quad \textstyle\sum_k (|y(k)|^2 - 1)^2 \qquad k = 1, \ldots, N, \tag{3}$$

where N is the length of the sequence $y(k)$, and we have assumed that the magnitude of the CM signal is equal to one. Using (1) and (3), the objective function can be written as:

$$\text{minimize} \quad \textstyle\sum_k (|\mathbf{w}^H \mathbf{x}_k|^2 - 1)^2 \qquad k = 1, \ldots, N, \tag{4}$$

where the weight vector $\mathbf{w}$ is the design variable. Since

$$|\mathbf{w}^H \mathbf{x}_k|^2 = \mathbf{w}^H \mathbf{x}_k (\mathbf{w}^H \mathbf{x}_k)^H = \mathbf{w}^H \mathbf{x}_k \mathbf{x}_k^H \mathbf{w}, \tag{5}$$

we have the following optimization problem:

$$\text{minimize} \quad \textstyle\sum_k (\mathbf{w}^H \mathbf{X}_k \mathbf{w} - 1)^2 \qquad k = 1, \ldots, N, \tag{6}$$

where $\mathbf{X}_k = \mathbf{x}_k \mathbf{x}_k^H$. In this paper we show that if we restrict the design variable $\mathbf{w}$ to a subset of the original feasible set, the blind equalization of CM signals can be expressed as a convex optimization problem. In particular, the problem can be formulated as a linear objective function, subject to linear and semidefiniteness constraints; that is, as a semidefinite program (SDP). Such SDPs can be efficiently solved using interior point methods [6].

3. SDP FORMULATION

In this section we outline the framework of our SDP method for blind equalization of CM signals for the special case of a BPSK signal transmitted through a real-valued channel. (The framework extends directly to complex-valued CM constellations and channels.) From (6) we see that our optimization problem can be written as,

$$\text{minimize} \quad f(\mathbf{w}), \tag{7}$$

where $f(\mathbf{w}) = \sum_k (\mathbf{w}^T \mathbf{X}_k \mathbf{w} - 1)^2$ is a 4-th order polynomial in $\mathbf{w}$. This optimization problem can be re-cast as:

$$\begin{aligned} &\text{maximize} \quad \tau \\ &\text{subject to} \quad f(\mathbf{w}) - \tau \geq 0 \qquad \text{for all } \mathbf{w}. \end{aligned} \tag{8}$$

We can think of τ as a horizontal hyperplane that lies beneath $f(\mathbf{w})$ for every value of $\mathbf{w}$. Instead of minimizing $f(\mathbf{w})$, we lift-up (maximize) the hyperplane while requiring it to always lie below $f(\mathbf{w})$. At the optimal solution, $\tau_{\max}$ equals to the optimal (minimum) value of $f(\mathbf{w})$. In order to express (6) as a convex problem we define two sets, $\mathscr{C}$ and $\mathscr{D}$:

$$\mathscr{C} = \{f \mid f(\mathbf{w}) \text{ is a 4-th order polynomial of } \mathbf{w} \text{ and } f(\mathbf{w}) \geq 0 \ \forall \mathbf{w}\};$$

$$\mathscr{D} = \{f \mid f(\mathbf{w}) = \sum_i g_i(\mathbf{w})^2,\ g_i(\mathbf{w}) \text{ is a quadratic polyn of } \mathbf{w}\}.$$

Obviously, $\mathscr{D} \subseteq \mathscr{C}$. In terms of $\mathscr{C}$ and $\mathscr{D}$ we can express our optimization problem as follows:

$$\begin{aligned} &\text{maximize} \quad \tau \\ &\text{subject to} \quad f(\mathbf{w}) - \tau \in \mathscr{C}. \end{aligned} \tag{9}$$

However, we consider the following optimization problem:

$$\begin{aligned} &\text{maximize} \quad \tau \\ &\text{subject to} \quad f(\mathbf{w}) - \tau \in \mathscr{D}. \end{aligned} \tag{10}$$

It is important to note that the optimization problem (10) is not equivalent to the one in (9), since we restrict the design variable $\mathbf{w}$ to a subset of the original feasible set. Eq. (10) is said to be a convex restriction of (9), and it provides a lower bound on the optimal value of τ in (9). However, in the noise free case, the optimal equalizer in (9) does lie in $\mathscr{D}$. This motivates us to consider (10). Moreover, the formulation in (10) is a convex optimization problem. To see this, we first show that:

$$f \in \mathscr{D} \Leftrightarrow f = \bar{\mathbf{w}}^T \mathbf{G} \bar{\mathbf{w}}, \qquad \text{for some} \qquad \mathbf{G} \succeq 0, \tag{11}$$

where

$$\bar{\mathbf{w}} = [w_1^2, \ldots, w_n^2, \mathbf{w}_2^T, w_1, \ldots, w_n, 1]^T \tag{12}$$

and $\mathbf{w}_2$ contains all products $w_i w_j,\ 1 \leq i < j \leq n$ in a specified order. The set $\mathscr{D}$ contains all 4-th order polynomial functions that can be written as:

$$f(\mathbf{w}) = \sum_i g_i^2(\mathbf{w}). \tag{13}$$

Since $g_i(\mathbf{w})$ is a quadratic polynomial function in $\mathbf{w}$, it can be written as

$$\begin{aligned} g_i(\mathbf{w}) = {} & q_{i,1} w_1^2 + \ldots + q_{i,n} w_n^2 \\ & + q_{i,n+1} w_1 w_2 + \ldots + q_{i,z} w_{n-1} w_n \\ & + q_{i,z+1} w_1 + \ldots + q_{i,z+n} w_n + q_{i,z+n+1}, \end{aligned} \tag{14}$$

where $z = n + \binom{n}{2}$. If we define a vector of coefficients $\mathbf{q}_i = [q_{i,1}, \ldots, q_{i,z+n+1}]^T$ then it follows that

$$g_i(\mathbf{w}) = \bar{\mathbf{w}}^T \mathbf{q}_i, \quad \text{and} \quad g_i^2(\mathbf{w}) = \bar{\mathbf{w}}^T \mathbf{q}_i \mathbf{q}_i^T \bar{\mathbf{w}},$$

and hence

$$f(\mathbf{w}) = \bar{\mathbf{w}}^T \sum_i (\mathbf{q}_i \mathbf{q}_i^T) \bar{\mathbf{w}} = \bar{\mathbf{w}}^T \mathbf{G} \bar{\mathbf{w}}. \tag{15}$$

Moreover, the reverse argument from (15) to (13) is trivial. Hence, we have proved (11). Using this fact and substituting $f'(\mathbf{w}) = f(\mathbf{w}) - \tau$ we can rewrite (10) as:

$$\begin{aligned} &\text{maximize} \quad \tau \\ &\text{subject to} \quad f'(\mathbf{w}) = \bar{\mathbf{w}}^T \mathbf{G} \bar{\mathbf{w}} \quad \text{for some } \mathbf{G} \succeq 0. \end{aligned} \tag{16}$$

Since $f'(\mathbf{w}) = \sum_k (\mathbf{w}^T \mathbf{X}_k \mathbf{w} - 1)^2 - \tau$ it can be written in the general form of an arbitrary 4-th order polynomial:

$$\begin{aligned} f'(\mathbf{w}) = {} & \sum_i p_i^{(4)} w_i^4 + \sum_i \sum_{j \neq i} r_{i,j}^{(4)} w_i^3 w_j + \sum_i \sum_{j > i} s_{i,j}^{(4)} w_i^2 w_j^2 \\ & + \sum_i \sum_{j \neq i} \sum_{i \neq k > j} t_{i,j}^{(4)} w_i^2 w_j w_k \\ & + \sum_i \sum_{j > i} \sum_{k > j} \sum_{l > k} u_{i,j}^{(4)} w_i w_j w_k w_l \\ & + \sum_i p_i^{(3)} w_i^3 + \sum_i \sum_{j \neq i} r_{i,j}^{(3)} w_i^2 w_j \\ & + \sum_i \sum_{j > i} \sum_{k > j} s_{i,j}^{(3)} w_i w_j w_k \\ & + \sum_i p_i^{(2)} w_i^2 + \sum_i \sum_{j > i} r_{i,j}^{(2)} w_i w_j \\ & + \sum_i p_i^{(1)} w_i + p^{(0)}. \end{aligned} \tag{17}$$

Hence, the constraint $f'(\mathbf{w}) = \bar{\mathbf{w}}^T \mathbf{G} \bar{\mathbf{w}}$ in (16) is equivalent to the set of linear constraints on $\mathbf{G}$ implicitly specified in (17). For example, we have that:

$$\begin{aligned} p_i^{(4)} &= \mathbf{G}_{i,i} & i &= 1, \ldots, n, \\ p_i^{(2)} &= \mathbf{G}_{m,i} + \mathbf{G}_{i,m} + \mathbf{G}_{z+i,z+i} & i &= 1, \ldots, n, \end{aligned}$$

where m is equal to the number of elements in (12) and $z = n + \binom{n}{2}$. The semidefinite constraint $\mathbf{G} \succeq 0$ ensures that $f'(\mathbf{w})$ is within the set $\mathscr{D}$. Finally, the SDP formulation of the problem of blind equalization of CM signals can be written as:

$$\begin{aligned} &\text{maximize} \quad \tau \\ &\text{subject to} \quad \mathbf{A}\mathbf{vec}(\mathbf{G}) = \mathbf{c} \\ &\qquad\qquad\quad\ \mathbf{G} \succeq 0, \end{aligned} \tag{18}$$

where $\mathbf{c}$ contains all the coefficients from (17), $\mathbf{A}$ is a selection matrix and $\mathbf{vec}(\mathbf{G})$ is the column-stacking operator. The SDP can be efficiently solved for the optimal $\mathbf{G}$ using interior point methods [6]. (We have used the SeDuMi implementation [8].) We next show how to find an equalizer $\mathbf{w}$ corresponding to the optimal $\mathbf{G}$.

3.1. Post-processing

The solution of (18) provides the optimal value of $\mathbf{G}$, denoted $\mathbf{G}_{opt}$. However, we want to implement an equalizer $\mathbf{w}$. How do we find $\mathbf{w}_{opt}$ once we have $\mathbf{G}_{opt}$? First note that $\mathbf{w} = (w_1, \ldots, w_n)$

is contained in $\bar{\mathbf{w}}$ (12), so once we have $\bar{\mathbf{w}}$ it is quite straightforward to obtain $\mathbf{w}$. Second, if perfect equalization is achieved, there is an optimal $\bar{\mathbf{w}}$, such that

$$f_{opt}^{'}(\mathbf{w}_{opt}) = \bar{\mathbf{w}}_{opt}^T \mathbf{G}_{opt} \bar{\mathbf{w}}_{opt} = 0. \qquad (19)$$

From (19) we can see that $\bar{\mathbf{w}}_{opt}$ lies in the null-space of $\mathbf{G}_{opt}$ and from (12) we can see that it has a specific structure. With this in mind, we devised the following alternating projections algorithm to determine $\bar{\mathbf{w}}_{opt}$: Given an initial vector, alternatingly project it onto $Null(\mathbf{G}_{opt})$ and then back onto the subspace $\mathscr{S}$ which contains all the vectors with structure given in (12). These projections are described as 1 and 2 below:

1. To project a vector onto $Null(\mathbf{G}_{opt})$ we premultiply it by the projection matrix $\mathbf{P} = \mathbf{V}_t \mathbf{V}_t^T$, where $\mathbf{V}_t$ contains the eigenvectors corresponding to the (almost) zero-valued eigenvalues of $\mathbf{G}_{opt}$. Then we re-scale the resulting vector, such that the last component becomes equal to one. We need to do this in order to perform the next step.

2. In order to project a vector onto subspace $\mathscr{S}$ we use the fact that there is a one-to-one correspondence between the vector $\bar{\mathbf{w}}$ given by (12) and the matrix $\bar{\mathbf{W}} = \mathbf{a}\mathbf{a}^T$, where $\mathbf{a} = [w_1, w_2, \ldots, w_n, 1]$; that is

$$\bar{\mathbf{W}} = \begin{bmatrix} w_1^2 & w_1 w_2 & \cdots & w_1 w_n & w_1 \\ w_2 w_1 & w_2^2 & \cdots & w_2 w_n & w_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ w_1 & w_2 & \cdots & w_n & 1 \end{bmatrix}. \qquad (20)$$

   Each element in $\bar{\mathbf{W}}$ is equal to a corresponding element in $\bar{\mathbf{w}}$. We see that a vector $\mathbf{y} \in \mathscr{S}$ if the corresponding matrix $\mathbf{Y}$ is a rank-one matrix. So, given arbitrary vector $\mathbf{y}$ we create the corresponding matrix $\mathbf{Y}$ and then perform a rank-one approximation of it. The resulting rank-one matrix $\hat{\mathbf{Y}}$ corresponds to a vector $\hat{\mathbf{y}} \in \mathscr{S}$ which is the projection of $\mathbf{y}$ onto $\mathscr{S}$.

4. IMPLEMENTATION AND SIMULATION

We demonstrate the effectiveness of our algorithm through several simulation examples. In all our examples, the SDPs were solved using SeDuMi [8].

4.1. Good telephone channel

In the first example, we consider the following channel $\mathbf{h}$=[0.04;-0.05;0.07;-0.21;-0.5;0.72;0.36;0.21;0.03;0.07], which is a typical response of a good quality telephone channel [7]. We compare our method with the algorithms given in [1, 2], which are blind adaptive algorithms. From [2] we used an adaptive algorithm that minimizes the cost function: $J = E\{\big||y(k)| - 1\big|\}$. In our test scenarios, the length of equalizers was 11 and the step size parameter for the adaptive algorithms was chosen to be 5×10^{-3}, which has proven to give the best results. We allowed 2,000 samples for the adaptation of weight coefficients for the methods in [1,2], but used only 500 samples for our method. We define the intersymbol interference as follows: $\text{ISI} = \left\|\frac{\mathbf{t}}{\max_i |[\mathbf{t}]_i|} - \mathbf{e}\right\|_2^2$, where $\mathbf{t} = \mathbf{h} * \mathbf{w}$ is a combined response of the channel and the equalizer, $*$ denotes convolution and $\mathbf{e}$ is a vector with 1 in the position $\arg\max_i \big|[\mathbf{t}]_i\big|$ and zeros elsewhere. We can see from Figs. 2 and 3 that the performance of blind adaptive methods in [1,2] depends on the initialization of the equalizer parameters. The curve denoted as CMA corresponds to the case when we use partial knowledge of the channel impulse response for initialization; i.e., we initialize the equalizer with a single 'spike' time-aligned with the channel response's center of mass. However, if such a knowledge is not available and the spike doesn't coincide with the channel response's center of mass, the adaptive algorithms may degrade in performance, as is shown with the curve CMA1. In both cases, simulation results indicate that our method achieves better average intersymbol interference suppression and a lower bit-error rate. This improved performance is achieved whilst requiring fewer samples than the algorithms in [1, 2].

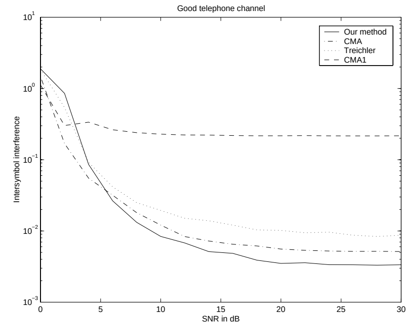

Figure 2: Intersymbol interference for Section 4.1

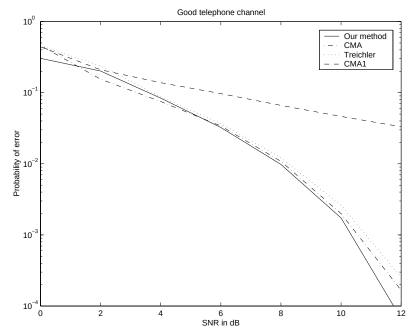

Figure 3: Probability of error for Section 4.1

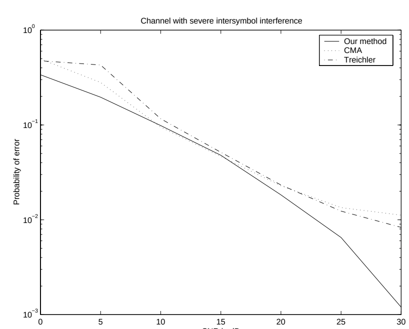

Figure 4: Probability of error for Section 4.2

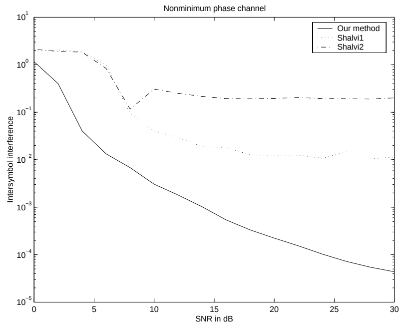

Figure 5: Intersymbol interference for Section 4.3

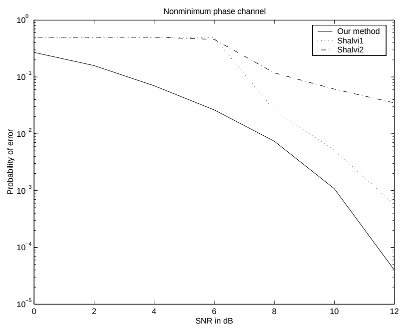

Figure 6: Probability of error for Section 4.3

4.2. Channel with severe interference

Here we consider the channel $\mathbf{h} = [0.407\ 0.815\ 0.407]$ given in [7]. This channel corresponds to a channel with severe ISI. This time we have used 12-tap equalizer for all methods and 1,000 samples for our method. Results shown in Fig.4 correspond to the case where the adaptive equalizers in [1] and [2] are initialized with a single 'spike' time-aligned with the channel response's center of mass. Again, we see that our method achieves a lower bit-error rate.

4.3. Nonminimum phase channel

In this example we compare our method with the method proposed in [3]. We consider the case where a communication source transmits a sequence of symbols through an unknown nonminimum phase channel $\mathbf{h}$ in which the unit sample response is given by:

$$[\mathbf{h}]_i = \begin{cases} 0, & i < 0 \\ -0.4, & i = 0 \\ 0.84 \times 0.4^{i-1}, & i > 0. \end{cases} \tag{21}$$

We have used a 9-tap equalizer for both methods. From Fig. 5 we can see that our method achieves better interference suppression which results in the improved bit-error rate performance shown in Fig.6. The curve Shalvi1 corresponds to the case where the adaptive equalizer in [3] is initialized with a single 'spike' in the middle of the equalizer, and Shalvi2 corresponds to the case where the spike is set at the beginning of the equalizer.

5. CONCLUSION

In this paper we have shown that blind equalization of constant modulus signals can be expressed as a convex optimization problem. A semidefinite programming formulation was made possible by performing an algebraic transformation on the direct formulation of the equalization problem and then restricting the design variables to a subset of the original feasible set. Simulation results indicate that our method has a better average performance than the methods proposed in [1–3], even in the case when we used *a priori* knowledge of the channel impulse response envelope to aid initialization for those standard methods. Furthermore, our method requires fewer samples, which might be useful for applications where convergence times requiring thousands of input samples are undesirable.

Our method incurs a higher computational cost than those in [1–3], but its improved performance in the preliminary simulations presented here motivates our current work on the development of specialized solvers which exploit the structure of the SDP in (18), analysis and refinements of the post-processing technique in Section 3.1, and the extension of the principles of this work to the complex-valued case. It also motivates generalizations to fractionally-spaced and multiple-sensor receivers, and comparisons with CM algorithms for those receivers, such as that in [9].

6. REFERENCES

[1] D. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. COM-28, no. 11, pp. 1867–1875, Nov. 1980.

[2] J. R. Treichler and M. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoustics, Speech, Signal Proc.*, vol. ASSP–31, pp. 459–472, Apr. 1983.

[3] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of non-minimum phase systems (channels)," *IEEE Trans. Info. Theory*, vol. IT-36, pp. 312–321, Mar. 1990.

[4] J. R. Treichler, V. Wolff, and C. R. Johnson, Jr., "Observed misconvergence in the constant modulus adaptive algorithm," *Proc. 25th Asilomar Conference on Signals, Systems and Computers*, pp. 663–667, Pacific Grove, CA, 1991.

[5] C. R. Johnson, Jr., P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind equalization using the constant modulus criterion: A review" *Proc. IEEE*, vol. 86, no. 10, pp. 1927–1950, Oct. 1998.

[6] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, pp. 49–95, 1996.

[7] J. G. Proakis, *Digital Communications*, McGrawHill, New York, fourth edition, 2000.

[8] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optim. Meth. Software*, vol. 11–12, pp. 625–653, 1999. Also: `http://www.unimaas.nl/~sturm/software/sedumi.html`.

[9] A.-J. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Signal Proc.*, vol. 44, May 1996.