Blind Channel Equalization and ϵ-Approximation Algorithms *

Qingyu Li[1], Er-Wei Bai[1*] and Yinyu Ye[2]

[1]Department of Electrical and Computer Engineering
[2]Department of Management Science
University of Iowa
Iowa City, IA 52242

Abstract

In this paper, we show that a blind equalizer can be obtained without using any statistical information on the input by formulating the blind channel equalization problem into a quadratic optimization with binary constraints. Then, efficient ϵ-approximation algorithms are presented and applied to find the equalizer coefficients. The proposed approach does not rely on any statistics, applies to non-minimum phase channels and converges to an ϵ neighborhood of a global optimum. Simulation results support our theoretical analyses.

SP EDICS: 3-CEQU.
Permission to publish this abstract separately is granted.

**Corresponding author. This work was supported in part by NSF grants DMI-9908077 and DMS-9703490.

1 Introduction

The problem of blind channel equalization is of considerable interest in data communication and other fields. Given a received output sequence, we want to recover the transmitted input sequence without knowing the channel dynamics. There exists a huge volume of research papers on this topic. Most early works require higher order statistics to identify the channel. These algorithms provide satisfactory results in some applications. Because of higher order statistics, however, a relatively long number of samples are needed and this may limit their applications in other cases. Recently, there has been an explosion in the research of blind channel equalization using only the second order statistics [1, 4, 5, 16, 17, 18]. The idea is to utilize multiple channel information by either over-sampling the output or using multiple antennas so that the lost phase information can be recovered. The key is multiple channels that share no common zeros.

There is much less work on blind channel equalization at symbol baud rates with no higher order statistics. In general, blind channel equalization at baud rates is not possible with no prior information on the input or the channel. Most early works in this direction use some optimization algorithms to adjust equalizer coefficients according to some cost functions [2, 6, 15]. The choice of the cost function is rather Ad Hoc and the resulting algorithms may or may not converge even to local minima. In fact, they some time diverge [9]. To overcome this problem, some global optimization algorithms have been proposed in the literature. For instance, the Viterbi algorithm which gives a maximum likelihood estimate. However, computational complexity is high. Recently, a direct examination method was also proposed in [8]. It was shown in the absence of noise that for a finite alphabet input set, the input signal can be identified by trying out all possible combinations. This is basically a semi-exhaustive approach and the complexity of computation is prohibitively high. In fact, exhaustive nature is a common property of all blind channel equalization algorithms if a global minimum is the objective. Moreover, most exhaustive methods are sensitive to the assumption on the channel order. Recently, an important result was obtained in [10]. It was shown that for a finite alphabet input set, if the received signal is fed into a stable filter and the output of the filter belongs to the the same alphabet with certain probabilistic distribution, then the output of the filter recovers the message signal upto a delay and a rotation constant. Along this line, some algorithms were developed in [11]. These algorithm are Ad Hoc and no complexity and convergence results are expected.

In this paper, we propose an approach of blind channel equalization at symbol baud rates. The proposed equalizer is actually a blind MMSE equalizer which does not use any statistical information of the input. Similar to those reported in the literature, the proposed approach requires no faster sampling. Unlike existing results, however, our approach has some unique features. It is based on a novel idea that formulates the blind channel equalization problem into an equivalent optimization problem with binary constraints. Then, efficient ϵ-approximation algorithms are employed to solve the problem. Our approach extends existing literature in several significant ways.

- The proposed blind channel equalization algorithms are guaranteed to converge to an ϵ neighborhood of a global minimum. The exact definition of "a neighborhood" will be defined later. Moreover, the accuracy of the ϵ solution has been improved significantly in this paper over the best available results in the literature by utilizing the special structure presented in the blind channel equalization setting.

- Besides convergences, the worst case computational complexity of our algorithms is polynomial. This property distinguishes the proposed approach from existing blind channel equalization algorithms at baud rates that have exponential computational complexity.

- The approach requires no faster sampling and no multiple antennas. It also applies to non-minimum phase channels with no high order statistics information. Moreover, the proposed approach does not depend on the exact order of the unknown channel and thus is robust.

Because of ϵ-convergence, a large number of simulation shows that our proposed approach recovers transmitted input signal very well at symbol baud rates.

The organization of the paper is as follows. Section 2 gives problem statement and some preliminary results. Blind channel equalization is discussed in this section. Section 2.1 is devoted to the case where the channel coefficients are real valued and the input is binary. For this simple case, we explain ideas and derive detailed formula. The case of complex channels and QAM inputs is discussed in Section 2.2. Extension to a more general alphabet is presented in Section 2.3. ϵ-approximation solutions as well as their convergence and complexity analyses are provided in Section 3. Algorithms and simulations are in Section 4. Some concluding remarks are given in Section 5.

Notation used in this paper is standard. In particular, A' denotes the transpose of the matrix A.

2 Problem formulation

A scalar communication channel discussed in this paper is described by

$$y(k) = \sum_{i=0}^{L} h_i s(k-i) + v(k) \tag{2.1}$$

where the FIR system

$$h(z) = h_0 + h_1 z^{-1} + ... + h_L z^{-L}$$

denotes the unknown channel with L being the order of the system which can be unknown. $y(\cdot)$ is the received output driven by the input sequence $s(\cdot)$ and $v(\cdot)$ represents the noise.

It is assumed in this paper that the input $s(\cdot)$ is possibly a complex random sequence belonging to a finite alphabet

$$\mathcal{A} = \{a(u), u \in S_A\},$$

which has at least two elements. Furthermore, the input $s(\cdot)$ is identically and independently distributed in $\mathcal{A}$ according to some probability distribution $P(\mathcal{A}, p)$ with $p(a(u)) = \text{Prob}\{s(k) = a(u)\}$. Moreover, it is assumed that $\min_u \ p(a(u)) > 0$.

To present the idea of our approach clearly, we assume in this section that the noise $v(k) = 0$. Now, if we knew $h(z)$ and also suppose $h(z)$ is stably invertible, then the channel equalization problem can be solved, i.e., suppose there is a stable $w(z) = \sum_{i=-\infty}^{\infty} w_i z^i$ such that

$$w(z)h(z) = 1,$$

then

$$\hat{s}(z) = w(z)y(z) = w(z)h(z)s(z) = s(z)$$

recovers the input $s(k)$ perfectly. The problems are: (1) we do not know if such a stable inverse $w(z)$ exists for an unknown $h(z)$. (2) If such a $w(z)$ exists, how to find it based only on the received output. To this end, let us recall the definition of a regular alphabet and the result of [10].

Definition 2.1 *A real alphabet $\mathcal{A} = \{a(u), u \in S_A\}$ is said to be regular if there exist real numbers $r_A > 0$ such that for all $u \in S_A$*

$$a(u) = a(0) + r_A u$$

and a complex alphabet $\mathcal{A}$ is said to be regular if there exist two real numbers $r_A^R, r_A^I > 0$ such that for all $(u_A, v_A) \in S_A$

$$a(u_A, v_A) = a(0,0) + r_A^R u_A + r_A^I v_A.$$

Theorem 2.1 *Consider a (real or complex valued) channel as in (2.1) and assume that the input $s(\cdot)$ is independently and identically distributed in a regular alphabet $\mathcal{A}$ according to some probability distribution $P(\mathcal{A}, p)$ with $\min_u \ p(a(u)) > 0$. Let $w(z) = \sum_{i=-\infty}^{\infty} w_i z^i$ be any nonzero l_1-stable filter and define*

$$\hat{s}(k) = \sum_{i=-\infty}^{\infty} w_i y(k-i).$$

Suppose $\hat{s}(\cdot)$ converges in distribution to some $P(\mathcal{A}, q)$ (q is not necessarily equal to p). Then, upto a possible delay and some rotation constants, we have

1. *$w(z)h(z) = rz^{-d}$.*

2. *$\hat{s}(k) = rs(k-d)$*

for some delay d and rotation constants r.

Now, suppose the unknown $h(z)$ does have a stable inverse $w(z)$. According to the above theorem, $w(z)$ and $s(\cdot)$ can be obtained by solving the following optimization problem

$$\min_{w_i, \bar{s}(k) \in \mathcal{A}} \sum_{k=1}^{\infty} | \sum_{i=-\infty}^{\infty} w_i y(k-i) - \bar{s}(k)|^2. \tag{2.2}$$

Clearly, the minimum is zero. Therefore, by solving (2.2), we obtain the inverse $w(z)$ of $h(z)$ and recover $s(k)$, upto a delay and a rotation constant. We remark that (2.2) may have more than one solution but all these solutions differ only by some delay and rotation constants. This fact is an immediate consequence of the uniqueness result (Theorem 3.1) of [3].

Now, the only remaining theoretical obstacle to our approach is how to check if $h(z)$ possesses a l_1-stable inverse. It is obvious to see that there is no l_1-stable inverse $w(z)$ if $h(z)$ has a zero on the unit circle. This can be seen easily. For instance, let $h(z) = 1 - rz^{-1}$

for some $|r| = 1$. Suppose there exists some $w(z) = \sum_{i=-\infty}^{\infty} w_i z^i$ such that

$$1 = w(z)h(z) = \sum_{i=-\infty}^{\infty} (w_i - rw_{i+1})z^i.$$

Then, we have $w_i = rw_{i+1}$, for $|i| > 0$. Since $1 = w(z)h(z)$, w_i can not be identically zero. On the other hand, $|w_i| = |rw_{i+1}| = |w_{i+1}|$ implies that the coefficient w_i does not decrease as $|i|$ increases and this shows that $w(z)$ can not be stable.

We now show that any $h(z)$, which can be minimum phase, maximum phase or mixed phase, always has a l_1-stable inverse as long as there is no zero on the unit circle.

Theorem 2.2 *Let*

$$h(z) = h_0 + h_1 z^{-1} + ... + h_L z^{-L} = \Pi_{i=1}^{L}(1 + r_i z^{-1})$$

$$= \Pi_{i=1}^{L_1}(1 + r_i z^{-1})\Pi_{i=L_1+1}^{L}(1 + r_i z^{-1})$$

where $1 \leq L_1 \leq L$, $|r_i| > 1, i = 1, 2, ..., L_1$ *and* $|r_i| < 1, i = L_1 + 1, ..., L$. *Then, there exists a* l_1*-stable* $w(z) = \sum_{i=-\infty}^{\infty} w_i z^i$ *such that*

$$1 = w(z)h(z) \quad and \quad |w_i| \leq M\rho^{|i|}$$

for some constants $M > 0$ *and* $0 \leq \rho < 1$.

Proof: For any input $s(z)$, let

$$x_1(z) = \Pi_{i=L_1+1}^{L}(1 + r_i z^{-1})s(z), \quad y(z) = \Pi_{i=1}^{L_1}(1 + r_i z^{-1})x_1(z) = h(z)s(z).$$

To prove $w(z)h(z) = 1$, it suffices to show that for any $s(z)$ and $y(z) = h(z)s(z)$, $s(z) = w(z)y(z)$. Now, note that $|r_i| < 1$, $i = L_1 + 1, ..., L$, we have

$$s(z) = \frac{1}{\Pi_{i=L_1+1}^{L}(1 + r_i z^{-1})}x_1(z) = (\sum_{i=0}^{\infty} \alpha_i z^{-i})x_1(z)$$

for some α_i such that $|\alpha_i| \leq M_1 \rho_1^i$, where $M_1 > 0$ and $0 \leq \rho_1 < 1$. Similarly, because of $|r_i| > 1$, $i = 1, 2..., L_1 + 1$, it follows that

$$y(z) = \Pi_{i=1}^{L_1}(1 + r_i z^{-1})x_1(z) = r_1...r_{L_1} z^{-L_1}\Pi_{i=1}^{L_1}(1 + z/r_i)x_1(z).$$

Thus,

$$x_1(z) = \frac{z^{L_1}}{r_1...r_{L_1}}\frac{1}{\Pi_{i=1}^{L_1}(1+z/r_i)}y(z) = \frac{z^{L_1}}{r_1...r_{L_1}}(1-z/r_1+z^2/r_1^2-...)...(1-z/r_{L_1}+z^2/r_{L_1}^2-...)y(z)$$

$$= z^{L_1}(\sum_{i=0}^{\infty}\beta_i z^i)y(z)$$

for some $|\beta_i| \leq M_2\rho_2^i$ where $M_2 > 0$ and $0 \leq \rho_2 < 1$. Now,

$$s(z) = \sum_{k=0}^{\infty}\alpha_k z^{-k}x_1(z) = z^{L_1}\sum_{k=0}^{\infty}\alpha_k z^{-k}\sum_{m=0}^{\infty}\beta_m z^m y(z) = [z^{L_1}\sum_{l=-\infty}^{\infty}w_l z^l]y(z)$$

with

$$w_l = \sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\alpha_k\beta_m, \quad m = k+l \ or \ k = m-l.$$

This shows that $w(z)h(z) = 1$. Moreover, if $l \geq 0$, $m = k+l \geq 0$, we have

$$|w_l| \leq M_3^2\sum_{k=0}^{\infty}\sum_{m=k+l}^{\infty}\rho^k\rho^m = M_3^2\sum_{k=0}^{\infty}\rho^k\frac{\rho^{k+l+1}}{1-\rho} = \frac{\rho M_3^2}{(1-\rho)(1-\rho^2)}\rho^l = M\rho^l$$

where $M_3 = max(M_1, M_2)$ and $\rho = max(\rho_1, \rho_2)$. Now, if $l < 0$, $m \geq 0$ implies $k = m - l = m + |l|$. Hence,

$$|w_l| \leq M_3^2\sum_{m=0}^{\infty}\sum_{k=m+|l|}^{\infty}\rho^k\rho^m = M_3^2\sum_{m=0}^{\infty}\rho^m\frac{\rho^{m+|l|+1}}{1-\rho} = M\rho^{|l|}.$$

This shows that $w(z)$ is l_1-stable and completes the proof.

The above two theorems provide a foundation for our blind channel equalization approach. Suppose the channel $h(z)$ does not have any zero on the unit circle, then the blind channel equalization problem can be solved by solving the optimization problem (2.2).

Unfortunately, the above optimization is only theoretically solvable because (1) A complete observation of the output sequence $\{y(k)\}_1^\infty$ is needed to solve (2.2). (2) The filter $w(z)$ is IIR which has an infinitely long coefficient sequence w_i that is impossible to calculate in practice. To make the approach practically efficient, we will make two approximations. First, we observe that the filter $w(z)$ is stable and the coefficient w_i converges to zero exponentially. Therefore, for a reasonable $l > 0$, $\sum_{i=-l}^{l} w_i z^i$ is close to $\sum_{i=-\infty}^{\infty} w_i z^i$. In fact, the difference between $\sum_{i=-l}^{l} w_i z^i$ and $\sum_{i=-\infty}^{\infty} w_i z^i$ converges to zero in any norm as l gets larger. Secondly, We replace the infinite summation from $k=1$ to $k=\infty$ by a finite summation from $k=N_0+1$ to $k=N_0+N$. This approximation also has a clear interpretation in terms of the minimum mean square error (MMSE) equalizer. To this end, note that for large N,

$$\min_{w_i, \bar{s}(k)\in\mathcal{A}} \sum_{k=N_0+1}^{N_0+N} |\sum_{i=-\infty}^{\infty} w_i y(k-i) - \bar{s}(k)|^2 \Longrightarrow \min_{w_i, \bar{s}(k)\in\mathcal{A}} \sum_{k=N_0+1}^{N_0+N} \frac{1}{N}|\sum_{i=-\infty}^{\infty} w_i y(k-i) - \bar{s}(k)|^2$$

$$\Longrightarrow \min_{w_i, \bar{s}(k)\in\mathcal{A}} \mathbf{E}|\sum_{i=-\infty}^{\infty} w_i y(k-i) - \bar{s}(k)|^2$$

where $\mathbf{E}$ is the expectation operator. This is exactly the MMSE equalizer.

After two approximations, the optimization problem (2.2) is reduced to

$$\min_{w_i, \bar{s}(k)\in\mathcal{A}} \sum_{k=N_0+1}^{N_0+N} |\sum_{i=-l}^{l} w_i y(k-i) - \bar{s}(k)|^2. \tag{2.3}$$

This optimization problem is clearly practically feasible, although its computational complexity is still high. In fact, the complexity increases exponentially as a function of N and l, and this makes the problem difficult. To overcome this difficulty, we note that there has been an incredible progress in numerical optimization literature on ϵ-approximation algorithms. Many NP hard problems actually have polynomial complexity if an ϵ-approximation solution for some $0<\epsilon<1$ is needed. The optimization (2.3) belongs to this class [13][19].

We now consider two input alphabets in detail.

2.1 Real channels and binary inputs

In this section, we consider the case that the channel coefficients h_i's are real valued and the input $s(k)$ is binary, i.e., $\mathcal{A}=\{1,-1\}$, and $\text{Prob}\{s(k)=1\}=\text{Prob}\{s(k)=-1\}=0.5$. Equal probability is not necessary and is assumed only to simplify the problem. Now, if

h_i's are real valued, the equalizer coefficients w_i's are also real valued. For chosen l and N, let

$$\bar{S}' = (\bar{s}(N_0+1), ..., \bar{s}(N_0+N)),$$

$$W' = (w_{-l}, ..., w_0, ..., w_l),$$

$$Y = \begin{pmatrix} y(N_0+1+l) & \dots & y(N_0+1) & \dots & y(N_0+1-l) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y(N_0+N+l) & \dots & y(N_0+N) & \dots & y(N_0+N-l) \end{pmatrix} \tag{2.4}$$

and

$$J = \sum_{k=N_0+1}^{N_0+N} |\sum_{i=-l}^{l} w_i y(k-i) - \bar{s}(k)|^2,$$

the optimization problem of (2.3) can be re-written as

$$\min_{w_i, \bar{s}(k)=\pm 1} J = \min_{w_i, \bar{s}(k)=\pm 1} \{\|YW - \bar{S}\|^2\}. \tag{2.5}$$

The following theorem provides a solution to (2.5).

Theorem 2.3 *Consider a real valued channel and binary inputs $s(k) = \pm 1$. Then,*

1. *$Y'Y$ is invertible for large N.*
2. *If $Y'Y$ is invertible, a solution W and $\bar{S}$ of the optimization problem (2.5) is given by*

$$W = (Y'Y)^{-1} Y' \bar{S}$$

where $\bar{S}$ is a solution of

$$\min_{\bar{s}(k)=\pm 1} \bar{S}' Q \bar{S}$$

with

$$Q = I - Y(Y'Y)^{-1} Y'. \tag{2.6}$$

Proof: For part 1, observe

$$Y = \underbrace{\begin{pmatrix} s(N_0+1+l) & \dots & s(N_0+1-l-L) \\ \vdots & \ddots & \vdots \\ s(N_0+N+l) & \dots\dots & s(N_0+N-l-L) \end{pmatrix}}_{\tilde{S}} \underbrace{\begin{pmatrix} h_0 & & \\ \vdots & \ddots & \\ h_L & & h_0 \\ & \ddots & \vdots \\ & & h_L \end{pmatrix}}_{\tilde{h}}. \tag{2.7}$$

The matrix $\tilde{h}$ is full column rank and also for large N,

$$\frac{1}{N}\tilde{S}'\tilde{S} \to I$$

where I is the identity matrix with proper dimension. It follows that for large N,

$$\frac{1}{N}Y'Y = \frac{1}{N}\tilde{h}'\tilde{S}'\tilde{S}\tilde{h} \to \tilde{h}'\tilde{h}$$

is non-singular and thus invertible.

To show the second part, note the optimization (2.5) involves two variables W and $\bar{S}$. They are not independent and related. A necessary condition for W to be optimal is that

$$0 = \frac{\partial J}{\partial W} = 2Y'YW - 2Y'\bar{S}.$$

If $(Y'Y)^{-1}$ exists, solution exists and in fact is unique

$$W = (Y'Y)^{-1}Y'\bar{S}. \tag{2.8}$$

Since the maximum exists, the uniqueness of the solution (2.8) guarantees the optimality of W provided that $\bar{S}$ is optimal. Now, by substituting $W = (Y'Y)^{-1}Y'\bar{S}$ back into the equation, we have

$$\begin{aligned} \min_{\bar{s}(k)=\pm 1} J &= \min_{\bar{s}(k)=\pm 1} \{\|(Y(Y'Y)^{-1}Y' - I)\bar{S}\|^2\} \\ &= \min_{\bar{s}(k)=\pm 1} \bar{S}'Q\bar{S} \end{aligned}$$

where

$$Q = (I - Y(Y'Y)^{-1}Y')'(I - Y(Y'Y)^{-1}Y') = I - Y(Y'Y)^{-1}Y'.$$

This completes the proof.

Clearly, the matrix Q is formed by the received signal $y(k)$ only and is available. This is a quadratic programming problem with binary quadratic constraints.

2.2 Complex channels and QAM inputs

In this case, $\mathcal{A} = \{\pm 1 \pm j\}$ and Prob$\{Re(s(k)) = \pm 1\} =$ Prob$\{Im(s(k)) = \pm 1\} = 0.5$. It is assumed that the real part and the imaginary part of $s(k)$ are statistically independent. Now, because all w_i, $y(k)$ and $\bar{s}(k)$ are complex, we have

$$w_i = Re(w_i) + jIm(w_i) \Longrightarrow W = (w_{-l}, ..., w_l)' = ReW + jImW$$

$$\bar{s}(k) = Re\bar{s}(k) + jIm\bar{s}(k) \Longrightarrow \bar{S} = (\bar{s}(N_0+1), ..., \bar{s}(N_0+N))' = Re\bar{S} + jIm\bar{S}$$

$$y(k) = Rey(k) + jImy(k) \Longrightarrow$$

$$Y = \begin{pmatrix} y(N_0+1+l) & \dots & y(N_0+1) & \dots & y(N_0+1-l) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y(N_0+N+l) & \dots & y(N_0+N) & \dots & y(N_0+N-l) \end{pmatrix} = ReY + jImY.$$

Moreover,

$$\bar{s}(k) \in \mathcal{A} = \{\pm 1 \pm j\} \Leftrightarrow Re\bar{s}(k) = \pm 1, \ Im\bar{s}(k) = \pm 1.$$

Now,

$$YW - \bar{S} = (ReY + jImY)(ReW + jImW) - (Re\bar{S} + jIm\bar{S})$$

$$= [(ReY, -ImY)\begin{pmatrix} ReW \\ ImW \end{pmatrix} - Re\bar{S}] + j[(ImY, ReY)\begin{pmatrix} ReW \\ ImW \end{pmatrix} - Im\bar{S}].$$

Therefore, the optimization problem (2.3) can be re-written as

$$\min_{w_i, \bar{s}(k) \in \mathcal{A}} \sum_{k=N_0+1}^{N_0+N} |\sum_{i=-l}^{l} w_i y(k-i) - \bar{s}(k)|^2 = \min_{w_i, \bar{s}(k) \in \mathcal{A}} \{\|YW - \bar{S}\|^2\} = \min_{w_i, \bar{s}(k) \in \mathcal{A}} J \tag{2.9}$$

$$= \min_{w_i, \bar{s}(k) \in \mathcal{A}} \{(ReW', ImW') \underbrace{\begin{pmatrix} ReY'ReY + ImY'ImY & ImY'ReY - ReY'ImY \\ ReY'ImY - ImY'ReY & ReY'ReY + ImY'ImY \end{pmatrix}}_{P1} \begin{pmatrix} ReW \\ ImW \end{pmatrix}$$

$$-2(ReW', ImW')\underbrace{\begin{pmatrix} ReY' & ImY' \\ +ImY' & ReY' \end{pmatrix}}_{P2}\begin{pmatrix} Re\bar{S} \\ Im\bar{S} \end{pmatrix} + \|\bar{S}\|^2\}.$$

Similar to the binary case, a solution W and $\bar{S}$ of (2.9) is provided by the following theorem.

Theorem 2.4 *Consider complex channels and QAM inputs. Then,*

1. *$P1 = P2 \cdot P2'$ is invertible for large N.*

2. *If $P1$ is invertible, (2.9) is solved by*

$$\begin{pmatrix} ReW \\ ImW \end{pmatrix} = P1^{-1}P2\begin{pmatrix} Re\bar{S} \\ Im\bar{S} \end{pmatrix}$$

where $\begin{pmatrix} Re\bar{S} \\ Im\bar{S} \end{pmatrix}$ *is a solution of*

$$\min_{Re\bar{s}(k)=\pm 1, Im\bar{s}(k)=\pm 1} (Re\bar{S}', Im\bar{S}')Q\begin{pmatrix} Re\bar{S} \\ Im\bar{S} \end{pmatrix}$$

with

$$Q = I - P2'(P2 \cdot P2')^{-1}P2. \tag{2.10}$$

Proof: To show the first part, we note again equation (2.7). However, for QAM inputs, Y, $\tilde{S}$ and $\tilde{h}$ are complex

$$ReY + jImY = (Re\tilde{S} + jIm\tilde{S})(Re\tilde{h} + jIm\tilde{h})$$

$$= (Re\tilde{S}Re\tilde{h} - Im\tilde{S}Im\tilde{h}) + j(Im\tilde{S}Re\tilde{h} + Re\tilde{S}Im\tilde{h}).$$

Note also that the matrix $\tilde{h}$ is full column rank. Therefore, for any non-zero $(2l+1)$-dimensional complex vector ξ,

$$0 \neq \tilde{h}\xi = (Re\tilde{h} + jIm\tilde{h})(Re\xi + jIm\xi) = (Re\tilde{h}Re\xi - Im\tilde{h}Im\xi) + j(Im\tilde{h}Re\xi + Re\tilde{h}Im\xi). \tag{2.11}$$

Since $P1 = P2 \cdot P2'$, the full column rankness of $P2'$ guarantees the invertibility of $P1$. Then, it suffices to show that $(ReY, -ImY)$ is full column rank for large N. To this end, note that

$$(ReY, -ImY) = (Re\tilde{S}, -Im\tilde{S}) \begin{pmatrix} Re\tilde{h} & -Im\tilde{h} \\ Im\tilde{h} & Re\tilde{h} \end{pmatrix}.$$

The matrix $\begin{pmatrix} Re\tilde{h} & -Im\tilde{h} \\ Im\tilde{h} & Re\tilde{h} \end{pmatrix}$ is full column rank. Otherwise, there exists a non-zero vector $\xi = Re\xi + jIm\xi$ such that

$$\begin{pmatrix} Re\tilde{h} & -Im\tilde{h} \\ Im\tilde{h} & Re\tilde{h} \end{pmatrix} \begin{pmatrix} Re\xi \\ Im\xi \end{pmatrix} = 0.$$

This implies

$$Re\tilde{h}Re\xi - Im\tilde{h}Im\xi = 0, \quad Im\tilde{h}Re\xi + Re\tilde{h}Im\xi = 0$$

which contradicts (2.11). Now, for large N,

$$\frac{1}{N}(ReY, -ImY)'(ReY, -ImY) = \frac{1}{N} \begin{pmatrix} Re\tilde{h} & -Im\tilde{h} \\ Im\tilde{h} & Re\tilde{h} \end{pmatrix}' (Re\tilde{S}, -Im\tilde{S})'(Re\tilde{S}, -Im\tilde{S}) \begin{pmatrix} Re\tilde{h} & -Im\tilde{h} \\ Im\tilde{h} & Re\tilde{h} \end{pmatrix}$$

$$\to \begin{pmatrix} Re\tilde{h} & -Im\tilde{h} \\ Im\tilde{h} & Re\tilde{h} \end{pmatrix}' \begin{pmatrix} Re\tilde{h} & -Im\tilde{h} \\ Im\tilde{h} & Re\tilde{h} \end{pmatrix}$$

which is invertible. This shows that $(ReY, -ImY)$ is full column rank for large N and hence $P1$ is invertible. To prove the second part, we observe again that W can be calculated in terms of $\bar{S}$ by setting

$$0 = \frac{\partial J}{\partial(ReW', ImW')'} = P1 \begin{pmatrix} ReW \\ ImW \end{pmatrix} - P2 \begin{pmatrix} Re\bar{S} \\ Im\bar{S} \end{pmatrix},$$

which gives rise to

$$\begin{pmatrix} ReW \\ ImW \end{pmatrix} = P1^{-1}P2 \begin{pmatrix} Re\bar{S} \\ Im\bar{S} \end{pmatrix} \tag{2.12}$$

if $P1$ is invertible. By substituting W in terms of $\bar{S}$, we finally have

$$\min_{Re\bar{s}(k)=\pm 1, Im\bar{s}(k)=\pm 1} (Re\bar{S}', Im\bar{S}')Q \begin{pmatrix} Re\bar{S} \\ Im\bar{S} \end{pmatrix}$$

with

$$Q = I - P2'(P1)^{-1}P2 = I - P2'(P2 \cdot P2')^{-1}P2.$$

This finished the proof.

The optimization problem is again a quadratic problem with binary quadratic constraints.

2.3 General input sets

In the previous subsections, 2 specific alphabets were discussed. The idea is to formulate the blind equalization problem into a quadratic optimization problem with binary constraints so that efficient ϵ approximation algorithms can be applied. We remark that this formulation is not limited to the alphabets discussed in the previous subsections and applies to general alphabets. We only give a simple example here to illustrate the idea. Consider a real valued multi-level alphabet $\mathcal{A} = \{-2M, -2(M-1), ..., -2, 0, 2, ..., 2(M-1), 2M\}$. Then, for every $s(k) \in \mathcal{A}$, we have

$$s(k) = s^1(k) + s^2(k) + ... + s^{2M}(k)$$

where $s^i(k) = \pm 1$. Therefore,

$$\begin{aligned} &\min_{w_i, \bar{s}(k) \in \mathcal{A}} \sum_{k=N_0+1}^{N_0+N} | \sum_{i=-l}^{l} w_i y(k-i) - \bar{s}(k)|^2 \\ = \min\{ &\sum_{k=N_0+1}^{N_0+N} | \sum_{i=-l}^{l} w_i y(k-i) - [\bar{s}^1(k) + ... + \bar{s}^{2M}(k)]|^2 \} \end{aligned}$$

s.t.

$$\bar{s}^i(k) = \pm 1.$$

This is again a quadratic problem with binary constraints. Admittedly, this characterization of the alphabet $\mathcal{A}$ into the summation of $s^1(k) + ... + s^{2M}(k)$ is not the most efficient way of characterization and there are several more efficient ways of characterization. It does, however, serve the purpose that blind channel equalization can be formulated into a quadratic optimization problem with binary constraints for a large class of multi-level alphabets.

3 ϵ-approximation solutions

We have shown that blind channel equalization problem can be formulated as a quadratic optimization problem with binary constraints

$$\min_{x\in\mathcal{C}} J(x) = x'Qx \tag{3.1}$$

where $x = (x_1, ..., x_N)' \in R^N$, $Q = Q' \in R^{N\times N}$ and $\mathcal{C}$ is the constraint set

$$\mathcal{C} = \{x \in R^N \;:\; x_k^2 = 1,\; i = 1, ..., N\}.$$

If a global optimum is needed, the computational complexity of the above problem is exponential. However, the problem can be solved in polynomial time if only an ϵ-approximation is required. Formally, an ϵ-approximation solution can be defined as follows. Let $\underline{J}$ and $\bar{J}$ denote the minimal and maximal objective values for $x \in \mathcal{C}$. Any $x \in \mathcal{C}$ is said to be an ϵ-approximation solution if

$$\frac{J(x) - \underline{J}}{\bar{J} - \underline{J}} \leq \epsilon.$$

Recently, some significant results have been obtained concerning the ϵ-approximation solutions. Several ϵ-approximation algorithms with polynomial time complexity have been developed [7, 13, 19]. The minimum ϵ achieved to have a polynomial time algorithm in the literature is $\epsilon = 4/7 = 0.571$ [13, 19]. The main steps of [13, 19] can be easily summarized. Let $X \in R^{N\times N}$ be a symmetric positive semi-definite matrix and let x_{ii}'s denote its diagonal components. Now, consider a positive semi-definite programming (SDP) relaxation problem

$$\min_X \mathrm{Trace}(Q'X)$$

$$\text{subject to} \;\; x_{kk} = 1, i = 1, ..., N$$

$$x_{kk} \leq 1, \quad X \geq 0.$$

where $X \geq 0$ means that X is positive semidefinite.

Let X be a solution of the SDP relaxation problem. Since X is positive semi-definite, we write

$$X = V'V$$

with

$$V = (V_1, V_2, ..., V_N) \in R^{N \times N}.$$

Further, let η be a random vector uniformly distributed on the N-dimensional unit ball (or simply the standard normal vector), and then assign

$$\hat{x} = \mathrm{diag}(\sqrt{x_{11}}, ..., \sqrt{x_{NN}})\ sign(V'\eta) \tag{3.2}$$

where $sign(x)$ is the vector whose components are

$$sign(x_i) = \begin{cases} 1 & if\ x_i \geq 0 \\ -1 & if\ x_i < 0 \end{cases}$$

The following result is from [19].

Theorem 3.1 *Consider the optimization problem of (3.1). Then, $\hat{x}$ given by (3.2) satisfies*

$$\frac{E_\eta[J(\hat{x})] - \underline{J}}{\bar{J} - \underline{J}} \leq 4/7 = 0.571, \tag{3.3}$$

where E_η is the expectation operator. Moreover, $\hat{x}$ can be obtained in polynomial time.

Because of the special structure of $Q = I - A(A'A)^{-1}A'$ for some full rank matrix of A in our blind channel equalization setting, we will show that ϵ can be further reduced when $\bar{J}(\leq N)$ is large.

Theorem 3.2 *Consider the optimization problem of (3.1) where $Q = I - A(A'A)^{-1}A'$ for some full rank matrix of A, and $\mathcal{C} = \{x \in R^N \ :\ x_i^2 = 1,\ i = 1, ..., N\}$. Then, $\hat{x}$ given by (3.2) satisfies*

$$\frac{E_\eta[J(\hat{x})] - \underline{J}}{N - \underline{J}} \leq 1 - \frac{2}{\pi} = 0.363. \tag{3.4}$$

Moreover, $\hat{x}$ can be obtained in polynomial time.

Proof: Note that

$$E_\eta[J(\hat{x})] = E_\eta[\hat{x}'Q\hat{x}] = E_\eta[\hat{x}'\hat{x}] - E_\eta[\hat{x}'(A(A'A)^{-1}A')\hat{x}] = N - E_\eta[\hat{x}'(A(A'A)^{-1}A')\hat{x}].$$

It is known from [7, 13, 19] that

$$E_\eta[\hat{x}'(A(A'A)^{-1}A')\hat{x}] = \mathrm{Trace}((A(A'A)^{-1}A')(\frac{2}{\pi}\arcsin[X]))$$

where $\arcsin[X]$ is the matrix whose i,jth element equal $\arcsin(X_{i,j})$ for $i,j = 1, ..., N$, and X is a solution of the corresponding SDP relaxation problem. Moreover,

$$\underline{J} \geq \mathrm{Trace}((I - A(A'A)^{-1}A')X) = N - \mathrm{Trace}((A(A'A)^{-1}A')X),$$

or

$$\mathrm{Trace}((A(A'A)^{-1}A')X) \geq N - \underline{J}.$$

On the other hand, it is proved in [13] that $\arcsin[X] - X$ is positive semidefinite. Since $(A(A'A)^{-1}A')$ is also positive semidefinite, we have

$$\begin{aligned}\mathrm{Trace}((A(A'A)^{-1}A')(\frac{2}{\pi}\arcsin[X])) &= \frac{2}{\pi}\mathrm{Trace}((A(A'A)^{-1}A')\arcsin[X]) \\ &\geq \frac{2}{\pi}\mathrm{Trace}((A(A'A)^{-1}A')X) \geq \frac{2}{\pi}(N - \underline{J}).\end{aligned}$$

Therefore,

$$E_\eta[J(\hat{x})] - \underline{J} \leq N - \underline{J} - \frac{2}{\pi}(N - \underline{J}) = (1 - \frac{2}{\pi})(N - \underline{J})$$

which completes the proof.

We make two remarks here:

Remark 1: Note that for large N, $\bar{J} \to N$. Thus, the above theorem basically says that for large N,

$$\frac{E_\eta[J(\hat{x})] - \underline{J}}{\bar{J} - \underline{J}} \leq 1 - \frac{2}{\pi} = 0.363.$$

$\epsilon = 0.363$ is much smaller than the best available result $\epsilon = 4/7 = 0.571$ in the literature. The reason why ϵ can be reduced significantly in our setting is because of the special structure $Q = I - A(A'A)^{-1}A'$ while in [13, 19], no special structure on Q is assumed.

Remark 2: In our algorithm, the error $\epsilon = 1 - \frac{2}{\pi}$ actually is not that small. However, in the setting of blind channel equalization, simulation seems to show that the approximation algorithm of provides a good result in most cases. If further improvement is preferable we may apply some local search algorithms taking the ϵ-approximation solution $\hat{x}$ as the initial point. In such a case, there is always a balance between accuracy and complexity. To maintain a polynomial complexity, we can bound the maximum number of iterations for a local search programming by, say one thousand. A large number of simulation shows satisfactory results when the number of iteration is limited by 500.

4 Algorithms and simulations

Based on the previous discussion, the algorithm of solving the blind channel equalization problem by using optimization approach can be summarized.

Algorithm for solving the blind channel by using optimization approach.

Given an unknown channel which can be non-minimum phase and the input alphabet.

1. Determine N and the equalizer order l. The choice of N and l depends on how much prior information we have on the channel and input. If there is no prior information available, N and l can be chosen large. In theory, large N and l would give a more accurate result at expense of computational complexity.

2. Collect received signal $y(k)$ and construct the matrices Y as in (2.4) and Q. Q depends on the input alphabet. For instance, for real channels and binary inputs, Q is given by (2.6). For complex channels and QAM inputs, Q is provided by (2.10).

3. Call for SDP programming to calculate an ϵ-approximation solution $\hat{x}$ of (3.2). Note for the binary input case, $\hat{x} = \bar{S}$ and in the QAM input case, $\hat{x} = (Re\bar{S}', Im\bar{S}')'$.

4. If further refinement is preferable, apply any local search algorithm taking $\hat{x}$ as the initial point. Limit the maximum number of local iteration to 500.

5. Let $\hat{s}(k)$ be the estimate of the unknown transmitted signal $s(k)$ which is obtained by

$$\hat{s}(k) = \mathcal{F}(\sum_{i=-l}^{l} w_i y(k-i))$$

where $W' = (w_{-l}, ..., w_0, ..., w_l)$ is the equalizer coefficient vector and $\mathcal{F}$ is the decision block which maps the output of the equalizer $\sum_{i=-l}^{l} w_i y(k-i)$ to the nearest element of the input alphabet.

Remark 3: It is interesting to note that the inequality (3.3) or (3.4) holds in an average sense. Therefore, if we could calculate 2 or 3 ϵ approximation solutions independently and choose the one with minimum J as a solution to our optimization problem, the equalization result can be again improved at a price of increased computational load. Simulations seem to suggest that this is a very effective way.

We now consider two numerical simulations. For all simulations, results were the average of 50 Monte Carlo calculations with $N = 100$. Local search was also implemented using the Tabu method [14] after the ϵ-solution is obtained and the maximum number of local iteration is restricted to 500.

The first simulation considers a real valued channel

$$h(z) = 0.0190 + 0.0210z^{-1} - 0.0720z^{-2} + 0.1390z^{-3} + 0.5010z^{-4} + 0.3310z^{-5}$$

$$-0.0110z^{-6} - 0.1010z^{-7} - 0.1040z^{-8}$$

which has five unstable zeros. Figure 1 shows the impulse response of $h(z)$. The input $s(k)$ is binary, equally and independently distributed in $\mathcal{A} = \{\pm 1\}$. Let $s(k) - \hat{s}(k)$ be the symbol error. Figure 2 shows symbol error rate with respect to SNR for several l using the proposed approach.

The second simulation considers a QAM input with $\mathcal{A} = \pm 1 \pm j$ and equal probability. The channel is from [5] that is complex valued

$$h(z) = 1 + (-1.28 - 0.301j)z^{-1} + (1.6717 + 2.385j)z^{-2} + (0.178 + 0.263j)z^{-3}.$$

The impulse response is shown in Figure 3. For this simulation, we choose the better one in very two consecutive ϵ-approximation solutions. Figure 4 is the equalized channel $w(z)h(z)$ for $l = 5$ and $snr = 40$. In the ideal case, the real part of the equalized channel should be a discrete time delta function which has value 1 when $k = 0$ and zero otherwise. The imaginary part should be identically zero. We see from the figure, the equalized is close to the ideal case. Signals $y(k)$ and $\bar{s}(k)$ are shown in Figure 5. Figure 6 is the symbol error rate with respect to SNR for several l.

5 Concluding remarks

The proposed approach of blind channel equalization can be summarized into two steps: (1) Formulate the blind channel equalization problem into an optimization problem where efficient ϵ-approximation algorithms exist. (2) Apply efficient ϵ-approximation algorithms to solve the problem. Comparing to existing methods in the literature, the proposed approach has some unique features: (1) It does not need to calculate any higher order statistics including the second orders. (2) It equalizes the channel without faster sampling or multiple antennas. (3) It does not depends on the exact order of $h(z)$ and thus is robust. In fact, $h(z)$ can be IIR. (4) Because of a two sided filter $w(z)$, the unknown channel can be minimum phase, maximum phase or mixed phase.

Our approach can also be easily extended to blind equalization of single-input multi-output channels that share some possibly common unstable zeros. Let the output be de-

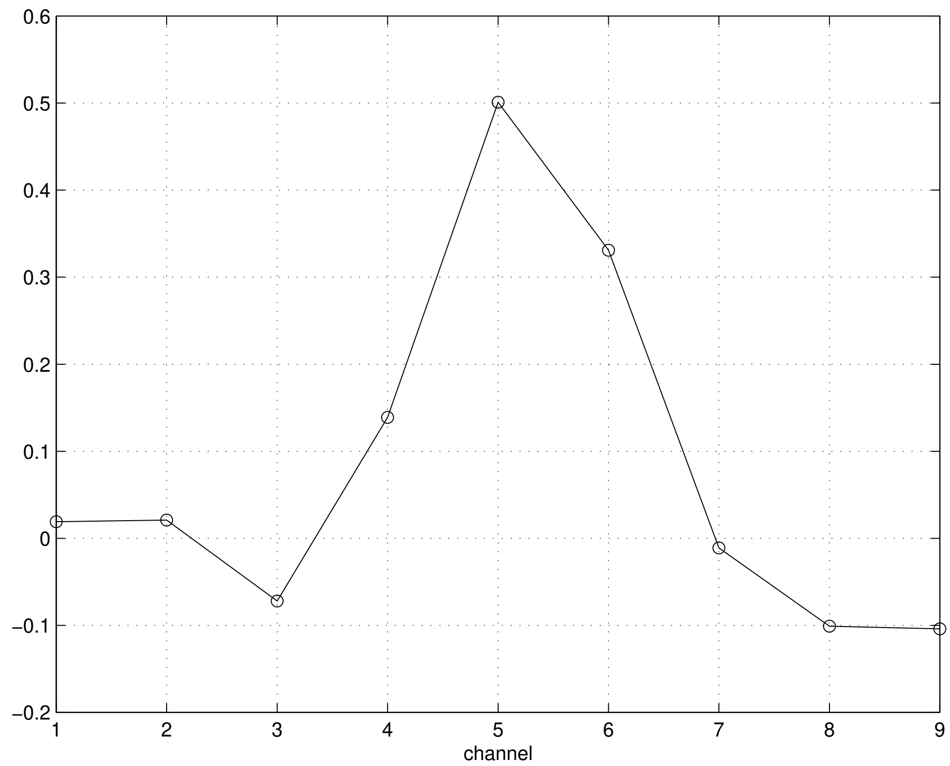

Figure 1: Channel impulse response

scribed by

$$y(z) = \begin{pmatrix} y_1(z) \\ \vdots \\ y_m(z) \end{pmatrix} = \begin{pmatrix} h_1(z) \\ \vdots \\ h_m(z) \end{pmatrix} s(z) = \begin{pmatrix} \bar{h}_1(z) \\ \vdots \\ \bar{h}_m(z) \end{pmatrix} c(z)s(z)$$

where $c(z)$ contains all the common zeros of $h_i(z)$. Suppose $c(z)$ does not have any zero on the unit circle. Then, there exist some l_1-stable $\bar{w}(z)$ and $\bar{w}_i(z),\ i = 1, ..., m$ such that

$$c(z)\bar{w}(z) = 1, \quad (\bar{w}_1(z), ..., \bar{w}_m(z)) \begin{pmatrix} \bar{h}_1(z) \\ \vdots \\ \bar{h}_m(z) \end{pmatrix} = 1.$$

Since $y_i(z)$'s are available, $\bar{w}(z)\bar{w}_i(z)$ can be calculated by applying the proposed approach, although the method is probably not efficient because it treats a SIMO channel into separate

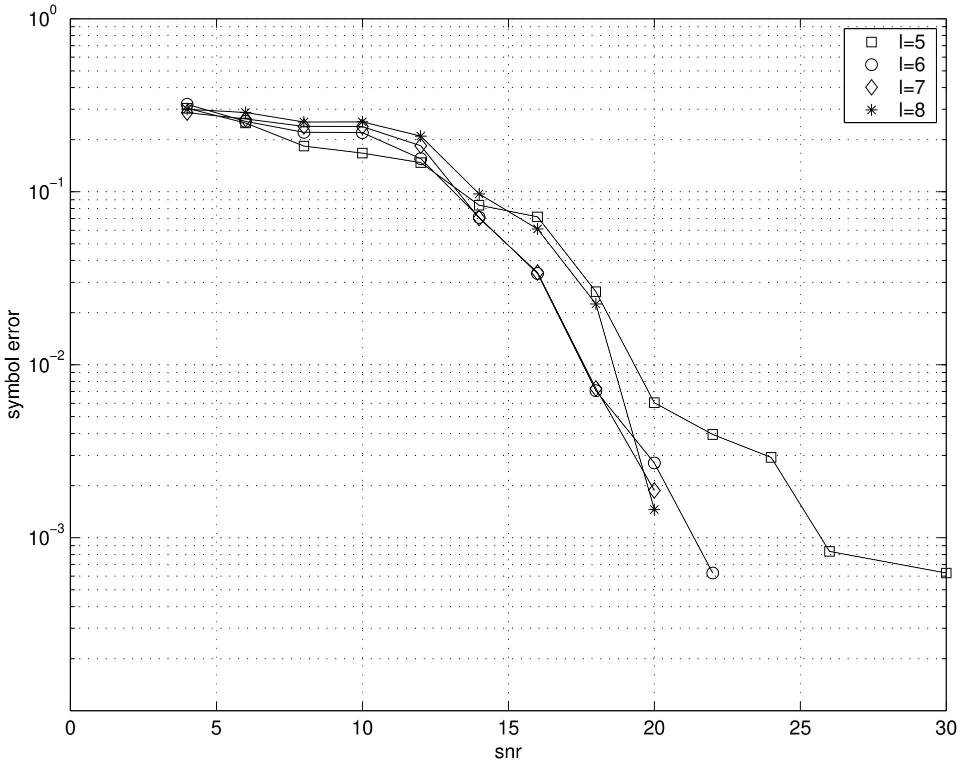

Figure 2: Symbol error rate versus SNR

individual channels. A more efficient way is to define an (unknown) signal $\bar{s}(z)$,

$$\bar{s}(z) = c(z)s(z), \quad y(z) = \begin{pmatrix} \bar{h}_1(z) \\ \vdots \\ \bar{h}_m(z) \end{pmatrix} \bar{s}(z).$$

Observe that $\bar{s}(z)$ can be determined, e.g. by using the method of [18], and once $\bar{s}(z)$ is available, $s(z)$ can be obtained by applying the proposed approach. Of course, the above steps are very sketchy. This is one of our current research topics.

It is also worth mentioning that the input $s(\cdot)$ is assumed to be stationary in the paper. However, Theorem 2.1 is valid for nonstationary input. This implies that our approach is applicable to the case where the statistics of the input is time-varying. This is possibly another further research topic.

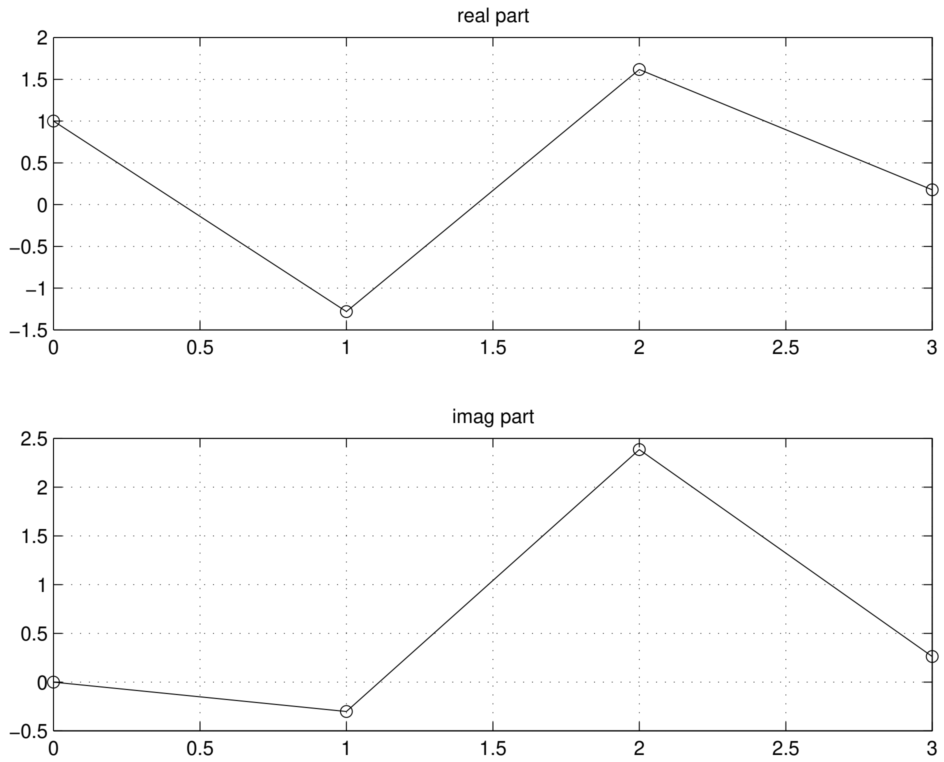

Figure 3: Real and imaginary parts of the complex channel

References

[1] E.W. Bai and M. Fu, "Blind system identification and channel equalization of IIR systems without statistical information", *IEEE Trans. on SP*, Vol. 47, pp.1910-1921, 1999

[2] A. Benvenisteand M Goursat, "Blind equalizer", *IEEE Trans. on Comm.*, Vol.32, pp.871-883, 1984

[3] Q. Cheng, "Maximum standardized cumulant deconvolution of non-Gaussian linear processes", *The Annual of Statistics*, Vol. 18, pp.1774-1783, 1990

[4] Z. Ding, "Matrix outer product decomposition method for blind multiple channel identification", *IEEE Trans. on SP*, Vol. 45, pp.3053-3061, 1997

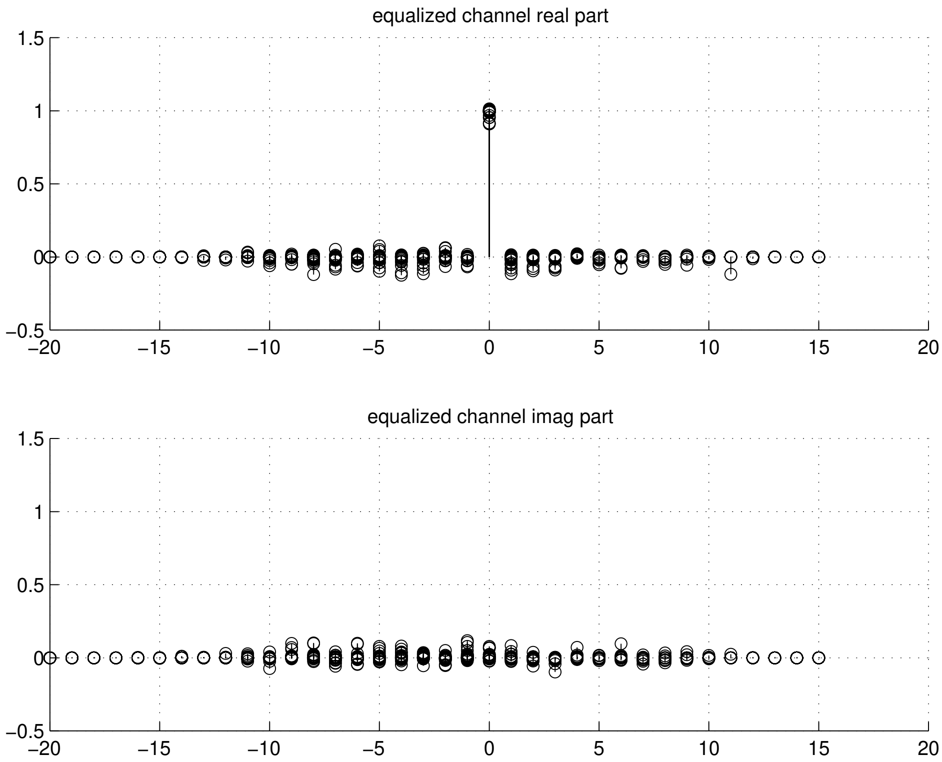

Figure 4: Real and imaginary parts of the complex channel

[5] G. Giannakis, "Direct blind equalization of multiple FIR Channels: a deterministic approach", *IEEE Trans. on SP*, Vol.47, pp.62-74, 1999

[6] D. Godard, "Self-recovering equalization and carrier tracking in two-dimension data communication system", *IEEE Trans. on Comm.*, Vol. 28, pp.1867-1875, 1980

[7] M. Goemans and D. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problem using semidefinite programming", *J. ACM*, Vol. 42, pp.1115-1145, 1995

[8] F. Gustafsson and B. Wahlberg, "Blind equalization by direct examination of the input sequences", *IEEE Trans. on Comm.*, Vol.43, pp.2213-2222, 1995

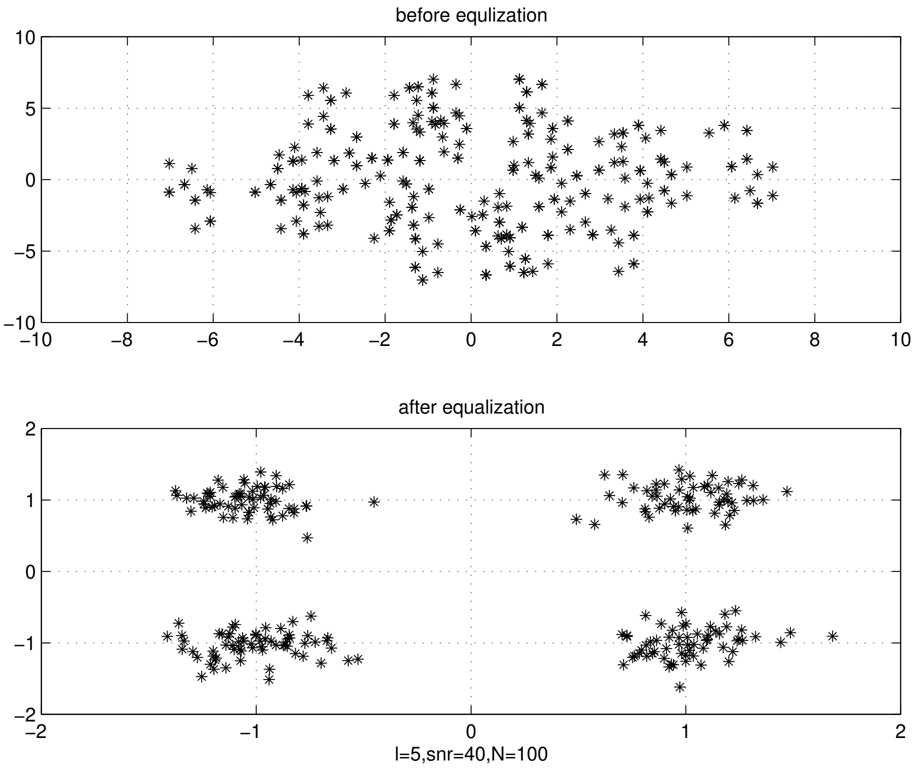

Figure 5: Signals before and after equalization

[9] R. Johnson, "Admissibility in blind adaptive channel equalization", *IEEE Contr. Sys. Mag.* Vol. 11, pp3-15, 1991

[10] T. Li, "Blind deconvolution of linear systems with multilevel nonstationary inputs", *The Annals Statistics*, Vol. 23, No.2, pp.690-704, 1995

[11] T. Li and K. Mbarek, "A blind equalizer for nonstationary discrete-valued signals", *IEEE Trans. on SP*, Vol.45, No.1, pp.247-254, 1997

[12] R. Liu and L. Tong (Eds), SPECIAL ISSUE ON BLIND IDENTIFICATION AND EQUALIZATION, *Proc. of IEEE*, Vol.86, No. 10, 1998

[13] Y. E. Nesterov, "Semidefinite relaxation and nonconvex quadratic optimization," *Optimization Methods and Software* Vol. 9, pp.141-160, 1998.

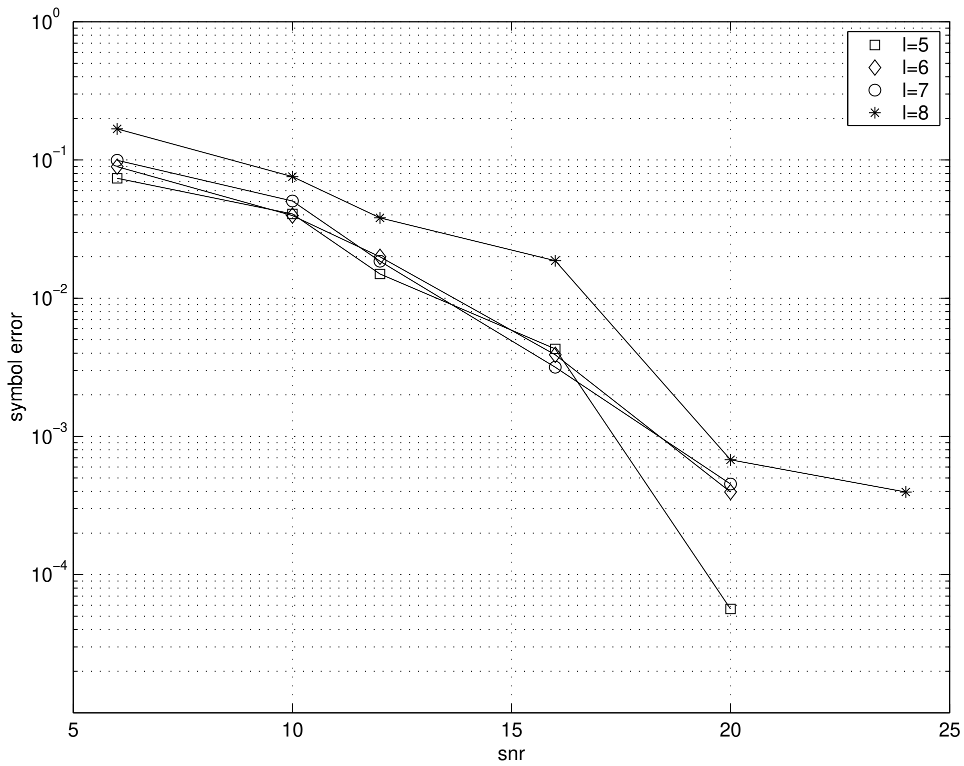

Figure 6: Symbol error rate versus SNR

[14] R. Reeves (Ed), ADVANCE TOPICS IN COMPUTER SCIENCE, pp.70-150, Blackwell Sci Publications, 1993

[15] Y. Sato, "A method of self-recovering equalization for multilevel amplitude modulation", *IEEE Trans. on Comm.* Vol. 23, pp.679-682, 1975

[16] L. Tong, G. Xu and T. Kailath, "Blind identification and equalization based on second-order statistics: a time domain approach", *IEEE Trans, on Info. Theory*, Vol. 40, pp.340-349, 1994

[17] J. Tugnait and B. Huang, "Second-order statistics-based blind equalization of IIR single-input multi-output channels with common zero", *IEEE Trans. on SP*, Vol. 47, pp.147-157, 1999

[18] G. Xu, H. Liu, L. Tong and T. Kailath, "A least squares approach to blind channel identification", *IEEE Trans. on SP*, Vol. 43, pp.2982-2993, 1995

[19] Y. Ye, "Approximating quadratic programming with bound and quadratic constraints", *Math Program.* Vol. 84, pp.219-226, 1999