On Blind Equalization of Biorthogonal Signaling

Andrew G. Klein, C. Richard Johnson, Jr., *Fellow, IEEE*, and Pierre Duhamel, *Fellow, IEEE*

***Abstract*—Motivated by increasing interest in energy efficient modulations, we provide the first look at adaptive equalization of biorthogonal signaling. While this modulation has historically been considered only for use in narrowband systems without intersymbol interference (ISI), recent attention has been given to its use in ISI channels. Due to the fact that biorthogonal modulation (BOM) results in a source that is not i.i.d., however, classical blind adaptive equalization techniques cannot be directly applied to equalization of BOM signals. We first examine minimum mean-squared error (MMSE) and least mean squares (LMS)-based equalizers, and identify some peculiarities that arise in equalization of BOM signals when compared to more traditional modulations like binary phase shift keying (BPSK). Next, we present two novel blind algorithms for the adaptive equalization of BOM signals: LTBOMB and TROMBONE. We discuss the convergence properties of these algorithms, and demonstrate their performance with numerical simulations.**

***Index Terms*—Adaptive, biorthogonal modulation (BOM), blind, energy efficient modulation, equalization, M-ary biorthogonal keying (MBOK).**

I. Introduction

INCREASING amounts of available unregulated spectrum, combined with an increasing demand for battery-operated wireless devices has spurred an interest in modulation schemes that give up some bandwidth efficiency in exchange for energy efficiency. Biorthogonal modulation (BOM) or M-ary biorthogonal keying (MBOK) is one such modulation scheme that has recently been considered for use in several consumer wireless standards, including the IEEE 802.11 WLAN standard [1] and the ultra wideband IEEE 802.15.3a WPAN standard [2]. Though BOM has been given serious consideration by standards bodies, twice making it to the final round of the selection process, little attention has been paid to BOM by the research community until recently (e.g., [3]–[5]). In the WLAN and WPAN applications where BOM has been considered, it is well known that intersymbol interference (ISI) will be present, and ISI is viewed to be a serious impairment to acceptable performance.

BOM and other energy efficient modulation schemes are certainly not new [6], but their use in frequency selective channels has only recently been considered. For these modulation schemes to ever be so deployed, some form of ISI compensation will be necessary. While the optimum detector in ISI is the maximum likelihood sequence estimator (MLSE), its complexity is usually too high for practical implementation, and, thus, suboptimal schemes are desirable. The only other work to consider compensation for ISI in BOM systems is [5], wherein the authors conduct a simulation study of a reduced state Viterbi equalizer for BOM. However, the equalizer used in [5] requires perfect channel knowledge, is not adaptive, and is still quite complex since it requires sequence estimation.

Because a BOM source is non-i.i.d. at the chip level, the adaptive equalization of non-i.i.d. sources is a related area of research. However, the majority of work on adaptive filtering of non-i.i.d. sources has been in the context of blind source separation of convolutive mixtures [7], [8]. In that application, the mixing matrix is prescribed to be tall, and the goal is to adapt a multichannel filter to recover the source data. The problem of interest here is fundamentally different since the channel matrix is fat and constrained to be Tœplitz, with the goal of adapting a single FIR filter. Also related is [9], where the effect of source statistics on an adaptive algorithm is analyzed; however, the focus there is on correlated sources rather than non-i.i.d. sources.

In this paper, we address the issue of linear equalization of BOM signals and propose two blind adaptive equalization algorithms specifically for BOM. In Section II, we review the basics of BOM, and present the system model and equalizer structure. In Section III we discuss the issues that arise in using classical approaches for adaptive equalization of BOM signals. In particular, we will present the design equations for the minimum mean-squared error (MMSE) equalizer, as well as the equations for trained and decision-directed (DD) least mean squares (LMS) adaptive algorithms. We also address the unsuitability in using the two most popular classical blind algorithms—the constant modulus algorithm (CMA) and the Shalvi-Weinstein algorithm (SWA). While the MMSE and LMS equalizers for BOM follow from straightforward application of Wiener filter theory, our main contribution in this paper is the invention of the first two blind algorithms beyond decision direction for BOM. In Section IV, we outline a strategy for crude assessment of blind algorithms, leading up to Sections V and VI which each propose a blind equalization algorithm for BOM, including a discussion of their characteristics and convergence. The first algorithm, called LTBOMB, is CMA-like in spirit, and we show that the zero-forcing (ZF) solutions are locally stable under ideal conditions. The second algorithm, called TROMBONE, was designed with a SWA-like philosophy in mind, and thus relies on a spectral prewhitener before equalization. We show that the ZF solutions are stationary points of TROMBONE. Section VII presents several numerical examples and simulations which demonstrate their performance relative to DD-LMS, and Section VIII concludes the paper.

Throughout this paper, we assume all signals are real. We use $^\top$ to denote matrix transpose, $[\boldsymbol{S}]_{i,j}$ to denote the i,jth entry of

Manuscript received October 6, 2005; revised May 12, 2006. The work of C. R. Johnson, Jr. was supported in part by NSF Grants CCF-0310023 and INT-0233127, Applied Signal Technology (Sunnyvale, CA), and a Fulbright Research Scholarship to France in 2005. The associate editor coordinating the review of this paper and approving it for publication was Prof. Philippe Loubaton.

A. G. Klein and P. Duhamel are with LSS, SUPELEC, Gif-sur-Yvette, France (e-mail: klein@lss.supelec.fr; pierre.duhamel@lss.supelec.fr).

C. R. Johnson, Jr. is with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: johnson@ece.cornell.edu).

Digital Object Identifier 10.1109/TSP.2006.889981

the matrix $\boldsymbol{S}$, and $[\boldsymbol{S}]_i$ to denote the ith column of the matrix $\boldsymbol{S}$. The unit vector consisting of a 1 in the ith location and zero everywhere else will be denoted $\boldsymbol{e}_i$, the $K \times K$ identity matrix is denoted $\boldsymbol{I}_K$, and the function $\delta[n]$ will be the Kronecker delta having value 1 when $n = 0$ and value zero when $n \neq 0$.

II. Background

A. Biorthogonal Modulation

An M-ary biorthogonal symbol contains $M/2$ chips, and is constructed by drawing from a set of $M/2$ orthonormal waveforms (each conveying $\log_2 M/2$ bits of information). The waveform is then modulated antipodally, which conveys another bit of information. Thus, there are M possible symbols, with each symbol representing $\log_2 M$ bits. Note that when $M = 2$, we have one chip per symbol, and BOM reduces to binary phase shift keying (BPSK). For convenience, we define $K \triangleq M/2$.

Let $\boldsymbol{S} \in \mathbb{R}^{K \times K}$ such that $\boldsymbol{S}S^\top = \boldsymbol{I}$ is an orthogonal matrix whose columns comprise the basis for the biorthogonal waveforms. Note that in this paper we only consider the case of a complete set of biorthogonal waveforms, thus requiring $\boldsymbol{S}$ to be square. We denote the symbol transmitted at time n by $\boldsymbol{x}[n]$ where $\boldsymbol{x}[n] \in \{\pm[\boldsymbol{S}]_0, \pm[\boldsymbol{S}]_1, \ldots, \pm[\boldsymbol{S}]_{K-1}\}$, and we assume all symbols are i.i.d. and equiprobable. We use a polyphase representation to describe the corresponding serial chip-rate process, denoting the ith chip of the nth symbol by $x[Kn - i]$, where $i \in \{0, \ldots, K-1\}$.

Since the columns of $\boldsymbol{S}$ are orthonormal, we have effectively fixed the *symbol* power to be unity, but the average power of each chip is therefore a function of K. Common choices for $\boldsymbol{S}$ include the Hadamard matrix [1] and the identity matrix [2]. Though we treat general biorthogonal bases throughout the paper, at times we stress the $\boldsymbol{S} = \boldsymbol{I}$ case which is sometimes referred to as *biorthogonal pulse position modulation* [3]. While the symbols are i.i.d., the chips are most certainly not. However, the chip-rate random process is cyclostationary with period K, and the second-order statistics are decorrelated as shown in the following lemma.

Lemma 1 (Decorrelation of Chip Statistics):

$$E[\boldsymbol{x}[n]\boldsymbol{x}^\top[m]] = \begin{cases} \frac{1}{K}\boldsymbol{I}_K & n = m \\ \boldsymbol{0}_{K\times K} & n \neq m. \end{cases}$$

Proof: Clearly $E[\boldsymbol{x}[n]] = \boldsymbol{0}_{K\times 1}$ due to the assumption of equiprobable symbols and the use of antipodal modulation. Thus, for $n \neq m, E[\boldsymbol{x}[n]\boldsymbol{x}^\top[m]] = \boldsymbol{0}_{K\times K}$ since the symbols are i.i.d. and zero mean. For $n = m$, averaging over all possible symbols gives $E[\boldsymbol{x}[n]\boldsymbol{x}^\top[n]] = (1/K)\sum_{k=0}^{K-1}[\boldsymbol{S}]_k[\boldsymbol{S}]_k^\top = (1/K)\boldsymbol{S}S^\top = (1/K)\boldsymbol{I}_K$. ∎

We note that the use of a code matrix in our model suggests that BOM shares some similarity with CDMA. Indeed, a BOM system can be viewed as a synchronous orthogonal-code CDMA system with $M/2$ antipodally modulated codes, with the caveat that only one i.i.d. randomly selected code is in use during each symbol period. This caveat, however, precludes the use of traditional CDMA demodulation techniques since the choice of code in use at a given instant is precisely what bears the information, and is, therefore, unknown to the receiver.

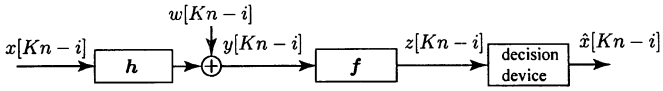


Fig. 1. System model.

B. System Model

The baseband-equivalent system model is shown in Fig. 1. While in general a sampled baseband equivalent model such as ours is represented by complex signals, we only consider real signals (i.e., as in pulse-amplitude modulated systems that do not employ quadrature reception)[1]. First, the BOM chips $x[Kn - i]$ are transmitted serially through a causal linear time-invariant channel with finite impulse response $\boldsymbol{h} = [h[0], h[1], \ldots, h[N_h - 1]]^\top$ where $h[k] = 0$ for $k \notin \{0, \ldots, N_h - 1\}$. We note that $\boldsymbol{h}$ includes the effects of pulse shaping, imperfect chip timing acquisition, and imperfect carrier phase recovery. The channel also contributes additive white Gaussian noise $w[Kn - i]$ of variance σ_w^2. The received chip stream $y[Kn - i]$ is passed through a linear equalizer with impulse response $\boldsymbol{f} = [f[0], f[1], \ldots, f[N_f - 1]]^\top$; the equalizer output is fed into the decision device, yielding $\hat{x}[Kn - i]$ which is an estimate of the transmitted chips.

To describe the system operation, we employ a formulation based on Tœplitz and Hankel matrices which permits us to isolate the equalizer vector $\boldsymbol{f}$. We encapsulate the channel impulse response in the Tœplitz matrix $\boldsymbol{\mathcal{H}} \in \mathbb{R}^{N_f \times N_c}$ defined as $[\boldsymbol{\mathcal{H}}]_{i,j} = h[j - i]$. The regressor matrix $\boldsymbol{Y}[n] \in \mathbb{R}^{N_f \times K}$ of received chips can be written as

$$\boldsymbol{Y}[n] = \boldsymbol{\mathcal{H}}\boldsymbol{X}[n] + \boldsymbol{W}[n] \tag{1}$$

where $N_c \triangleq N_f + N_h - 1, \boldsymbol{X}[n] \in \mathbb{R}^{N_c \times K}$ is the Hankel matrix of transmitted chips defined as $[\boldsymbol{X}[n]]_{i,j} = x[Kn - i - j], \boldsymbol{W}[n] \in \mathbb{R}^{N_f \times K}$ is the Hankel matrix of noise samples defined as $[\boldsymbol{W}[n]]_{i,j} = w[Kn - i - j]$, and $\boldsymbol{Y}[n]$ is therefore also a Hankel matrix. Then, the symbol output by the equalizer at time n can be written as

$$\boldsymbol{z}[n] = \boldsymbol{Y}^\top[n]\boldsymbol{f} \tag{2}$$

where $\boldsymbol{z}[n] \in \mathbb{R}^K$. From time to time, we will also use the combined channel/equalizer response $\boldsymbol{c} = [c[0], c[1], \ldots, c[N_c - 1]]^\top$, defined as $\boldsymbol{c} \triangleq \boldsymbol{\mathcal{H}}^\top\boldsymbol{f}$.

The decision device assumed in this paper is the naïve memoryless Euclidean distance detector, which is essentially a correlation detector [10]. This can be implemented very simply by forming the correlation $\boldsymbol{S}^\top\boldsymbol{z}[n]$, and then deciding in favor of the component with largest magnitude; the polarity of the decision is then given by the sign of this largest component. We have chosen this detector for its simplicity and low latency. When the equalizer is operating correctly, the decision device output is $\hat{\boldsymbol{x}}[n] \approx \boldsymbol{x}[n - \Delta]$ where Δ is the symbol delay through the combined channel and equalizer.

Last, we note that some means of symbol timing (i.e., aligning the chips so the block decision device operates on the symbol

[1]The results derived in this paper apply only to real-valued implementations, though we expect the extension to complex-valued signals to be fairly straightforward.

boundary) will be necessary. Our model does not assume that symbol timing has been accounted for before equalization, and thus we implicitly put the burden of symbol timing on the equalizer. That is, the equalizer needs to adjust its delay so that the total delay through the channel and equalizer is a multiple of K; otherwise, the decision device will not operate on a symbol boundary. In the sequel, we will see that the symbol timing ambiguity can pose a problem for blind equalization algorithms in some cases.

III. Classical Equalization Applied to BOM

A. MMSE Equalizer

In this section, we consider the classical approaches for calculating the equalizer taps $\boldsymbol{f}$, including direct MMSE calculation, LMS adaptation, and blind approaches. We begin by deriving the MMSE equalizer, a standard benchmark of equalizer performance. Let $\boldsymbol{e}_{K\Delta} \in \{0,1\}^{N_c}$ be the unit vector consisting of a 1 in the $(K\Delta)$th location, where Δ is a design parameter in this case. Note that $\boldsymbol{e}_{K\Delta}$ represents the desired combined channel/equalizer impulse response, and by choosing the delay to be a multiple of K we are assured that the chips are aligned on the symbol boundary before passing through the decision device. The MSE is given by

$$\begin{aligned} J_{\mathrm{mse}}(\boldsymbol{f}) &= E\left[||\boldsymbol{z}[n] - \boldsymbol{x}[n-\Delta]||_2^2\right] \\ &= E\left[||\boldsymbol{z}[n] - \boldsymbol{X}^\top[n]\boldsymbol{e}_{K\Delta}||_2^2\right]. \end{aligned}$$

Assuming the noise and data are uncorrelated, and using the facts that $E[\boldsymbol{X}\boldsymbol{X}^\top] = \boldsymbol{I}$ from Lemma 1 and for AWGN $E[\boldsymbol{W}\boldsymbol{W}^\top] = K\sigma_w^2\boldsymbol{I}$, the orthogonality principle gives the MMSE equalizer

$$\boldsymbol{f}_{\mathrm{mse}} = \left(\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^\top + K\sigma_w^2\boldsymbol{I}\right)^{-1}\boldsymbol{\mathcal{H}}\boldsymbol{e}_{K\Delta}. \tag{3}$$

We note that (3) is independent of the underlying orthonormal basis vectors $\boldsymbol{S}$, and coincides with the MMSE equalizer for BPSK modulation.

B. LMS Equalizer

As calculation of (3) requires perfect knowledge of the channel coefficients, we seek other means of calculating the MMSE equalizer. Since the MSE is quadratic in $\boldsymbol{f}$, we can use the LMS algorithm to calculate the MMSE equalizer adaptively when training data is available. The LMS algorithm is a stochastic gradient descent algorithm which uses the instantaneous gradient [11] of the MSE as an estimate for the true gradient (i.e., by ignoring the expectation operator). This results in the LMS update equation

$$\begin{aligned} \boldsymbol{f}_{\mathrm{lms}}[n+1] &= \boldsymbol{f}_{\mathrm{lms}}[n] - \mu\hat{\nabla}J_{\mathrm{mse}} \\ &= \boldsymbol{f}_{\mathrm{lms}}[n] - \mu\boldsymbol{Y}[n](\boldsymbol{z}[n] - \boldsymbol{x}[n-\Delta]) \end{aligned}$$

where μ is a small positive step-size which serves to average out the noise in the gradient estimate. With a small step-size, the algorithm exhibits mean transient and steady-state behavior very close to that of the exact gradient descent [11]. Note that the presence of $\boldsymbol{x}[n-\Delta]$ in the update term implies the availability of training data. When training data is unavailable, it is common to feed back the output of the decision device $\hat{\boldsymbol{x}}[n]$ instead, arriving at the DD-LMS update equation

$$\boldsymbol{f}_{DD}[n+1] = \boldsymbol{f}_{DD}[n] - \mu\boldsymbol{Y}[n](\boldsymbol{z}[n] - \hat{\boldsymbol{x}}[n])$$

which arrives from the cost function

$$J_{DD}(\boldsymbol{f}) = E\left[||\boldsymbol{z}[n] - \hat{\boldsymbol{x}}[n]||_2^2\right]. \tag{4}$$

Decision directed adaptive equalizers are notoriously sensitive to initialization [12], and generally require a nearly open eye initialization to ensure a sufficiently low symbol error rate. Consequently, DD algorithms are not a good choice for cold startup of BPSK equalizers [12], and the situation may only be worse for BOM signals since the open eye region is even smaller [13].

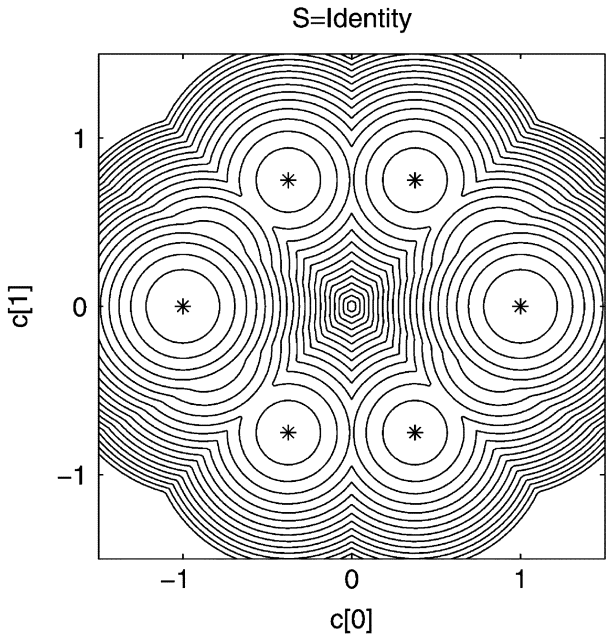


Fig. 2. Example cost surface for DD-LMS.

C. Example 1: False Local Minima in DD-LMS

To illustrate the fact that DD-LMS is indeed sensitive to initialization with BOM signals, we consider a specific low-dimensional example which permits us to visualize the DD-LMS cost surface as a function of $\boldsymbol{c}$, the combined channel/equalizer coefficients. We ignore the AWGN, we let $\boldsymbol{S} = \boldsymbol{I}_2$, and we consider only two taps so $N_c = 2$. Though this choice of system parameters may seem too simplistic to be practical, analysis of decision-directed equalizers is particularly difficult due to the presence of discontinuous functions (i.e., the decision device) [12], and this example is only meant to demonstrate that the DD-LMS algorithm has false local minima even in low-dimensional noiseless scenarios. We compute the DD cost from (4) by averaging over all possible channel inputs for this 2-tap example, and the resulting cost surface is plotted in the combined channel/equalizer space in Fig. 2, with the minima indicated by

asterisks. Ideally, this cost function would only have minima at single spike solutions; indeed, we note the appearance of local minima at the desired location $\boldsymbol{c} = \pm[1, 0]^\top$. However, we also note that there are false minima at $\boldsymbol{c} = \pm[3/8, 3/4]^\top$ and $\boldsymbol{c} = \pm[3/8, -3/4]^\top$, which verifies our claim that DD-LMS is sensitive to initialization. As the length of the channel/equalizer space is increased, the number of false local minima of the highly faceted DD-LMS cost surface only grows. This motivates the search for blind methods of equalizer adaptation for BOM other than decision directed algorithms.

D. Classical Blind Approaches

When faced with the task of designing a blind adaptive equalization algorithm for a new modulation scheme, a natural path is to consider the use of classical approaches to blind equalization. The focus of this paper is on stochastic gradient descent-based schemes, and the two most common classical algorithms that fall into this category are the constant modulus algorithm (CMA) [14], [15], and the Shalvi-Weinstein algorithm (SWA) [16].

As mentioned in Section II-A, BOM reduces to BPSK when $K = 1$. Likewise, when $K = 2$ and $\boldsymbol{S}$ is chosen to be the Hadamard matrix, the BOM chips have statistics identical to a BPSK source. Thus, we can expect the classical blind algorithms to work fine for some very particular choices of K and $\boldsymbol{S}$. However, the classical blind algorithms are unsuitable in general, as we now explain.

The CMA performs gradient descent of the cost $J_{\mathrm{CMA}}(\boldsymbol{f}) = E[(z^2[Kn - i] - 1)^2]$ where the underlying random source $x[Kn - i]$ is assumed to be i.i.d. Clearly, the i.i.d. assumption does not hold in general for a cyclostationary BOM source, and so traditional known results [17] of CMA are not applicable. As the CMA cost depends on i, we might consider an extension of CMA suitable for use on cyclostationary BOM sources, e.g.

$$J'_{\mathrm{CMA}}(\boldsymbol{f}) = \sum_{i=0}^{K-1} E[(z^2[Kn - i] - 1)^2]. \tag{5}$$

However, using local convergence analysis of this algorithm (such as that used in Sections V and VI), it can be shown that, for some choices[2] of $\boldsymbol{S}$, the zero-forcing solutions are not stable points of the CMA nor its cyclostationary extension. That is, even if the algorithm is initialized near a desired solution with no ISI, it will move away from the desired solution. Clearly, this is an undesirable feature which motivates the search for cost functions that are more suitable for cyclostationary BOM sources.

The SWA, on the other hand, maximizes the magnitude of the kurtosis of the equalizer output,

$$J_{\mathrm{SWA}}(\boldsymbol{f}) = |\mathcal{C}_4(z[Kn - i])| \tag{6}$$

where the kurtosis (or fourth-order cumulant) of a real-valued i.i.d. random process is defined as

$$\mathcal{C}_4(z[Kn - i]) \triangleq E[z^4[Kn - i]] - 3E[z^2[Kn - i]]^2.$$

[2]For example, the choice $\boldsymbol{S} = \boldsymbol{I}_K$ with $K > 3$ (which, incidentally, yields leptokurtic source statistics known to pose problems to CMA [9]).

In addition, the SWA algorithm requires a unit-norm constraint on the equalizer taps to avoid the trivial solution, and it has the requirement that prewhitening be performed before equalization so that the effective channel is white. The rationale for this criterion (6) is based on the fact (see theorem in [16]) that when the equalizer output power equals the power of the source process, the magnitude of the channel/equalizer output kurtosis is less than or equal to the magnitude of the source kurtosis, or $|\mathcal{C}_4(z[Kn - i])| \leq |\mathcal{C}_4(x[Kn - i])|$. Equality occurs when ISI has been eliminated, and so constrained maximization of (6) seems like a sensible approach. As is the case with the CMA, however, the standard analysis [16] of the SWA also requires that the source is i.i.d., and the cyclostationary nature of the BOM signal renders the SWA unsuitable as a candidate algorithm.

While these two algorithms are both unsuitable for BOM signals, it is worth pointing out that conventional use of both of these algorithms (say, with BPSK) results in global convergence to the ZF solutions under the assumptions of an infinite length real-valued equalizer and the absence of noise, which is a desirable feature of any candidate algorithm.

IV. Methodology for Assessment of Blind Algorithms

We now describe a general strategy for the value assessment of blind adaptive equalization algorithms based on gradient descent of multimodal cost functions. Blind algorithms of this kind typically rely on some form of property restoral, by attempting to recover properties of the original source signal that are altered or destroyed by the communication channel. Thus, a candidate cost function is chosen to penalize deviation from these desired properties. However, there is no guarantee that arbitrary cost functions chosen in this way will exhibit acceptable performance. And while algorithm performance can be verified with simulation to some degree, a purely simulation-based value assessment provides little knowledge of algorithm behavior.

Analysis of the candidate cost function in the most ideal situations is a typical first step, since any resulting algorithm that performs poorly in the most idealized scenarios would not generally be expected to perform any better in more practical scenarios. Thus, in the early stages of algorithm analysis, it is reasonable to ignore the effects of AWGN and to assume that the equalizer is sufficiently long so that the desired points in the channel/equalizer space are reachable with arbitrary precision. In most cases, including BOM, the desired points in the absence of noise amount to single-spike impulse responses (i.e., those with only one non-zero tap), which we call the zero-forcing (ZF) solutions. Yet another assumption is that the algorithm step-size is sufficiently small so that the stochastic gradient descent algorithm exhibits mean transient and steady-state behavior very close to that of the exact gradient descent.

Under these ideal assumptions, then, the first stage of the algorithm assessment is concerned with whether the ZF solutions are indeed stationary points of the candidate cost function, and secondly that they are *stable* stationary points of the cost function. Such assessment is done by examining the gradient and Hessian of the cost function, or by using perturbation analysis. Any candidate algorithm that does not exhibit local stability around the desired solutions should be modified or discarded.

The next phase of algorithm assessment evaluates the existence of other stationary points, particularly those that are false local minima. Typically, such a search for stationary points of the algorithm is conducted under the assumption of an infinitely long equalizer [18], so that we can invoke the assumption of an invertible relationship between the combined channel/equalizer response and the equalizer coefficients. Unfortunately, identification of all classes of stationary points of arbitrary length is not always tractable. When such intractability persists, we are forced to resort to searching for classes of stationary points that occur as impulse responses with finite time support, i.e., with some finite number of contiguous nonzero taps. Under this approach, we still assume that the equalizer has infinite length, but we set all but a finite number N_c of combined taps in the gradient to be zero; this allows us to work with systems of polynomials with a finite number of terms whose solutions can be calculated exactly, for example, with Gröbner bases [19].

We thus start by considering the class of all channel/equalizer combinations with, say, $N_c = 2$ nonzero coefficients. We can quite easily calculate the location of all stationary points, and can then classify them as maxima, minima, or saddle points by examining the eigenvalues of the Hessian. Assuming no false minima are found, we proceed by successively increasing the number of nonzero taps in the channel/equalizer combination, and continue categorizing the stationary points. However, as N_c increases beyond a certain point, we exhaust the memory and computational resources required by exact polynomial solvers, and so we will be forced to use a numerical search for stationary points.

In general, stationary points found in low-dimensional examples will persist into higher dimensions, and additionally, delayed versions obtained by adding zeros to the front of such impulse responses will generally be stationary points, as well. The fact that stationary points found in low-dimensional examples persist into higher dimensions is fairly obvious, since any finite length impulse response can be made into a longer response by simply appending zeros to the *end*. For a cyclostationary source with period K, the statistics of the received signal are unaffected by a K-chip delay; consequently, causal impulse responses corresponding to stationary points with K zeros added to the *front* are also stationary points.

Examination of finite-length channel/equalizer combined responses may not directly lead to a statement about global convergence behavior. However, in lieu of a global convergence proof, the technique can be used to build confidence that the algorithm exhibits good behavior. If, in the search for stationary points, we observe false local minima, we can possibly use their character to propose a fix to the algorithm.

If false local minima are observed, we would like to have some idea about their regions of attraction in comparison to other candidate algorithms. One technique for comparing algorithms' regions of attraction is to initialize the algorithms to the ZF solution, and then gradually expand the initializations in a sphere around the ZF solution. The superior algorithm will be the one that succeeds in converging to the ZF solution when the sphere of possible initializations is largest.

After assessing the location and character of algorithm stationary points, the next stage of algorithm assessment is simulation in a practical scenario—on practical channels with AWGN. Further assessment of the candidate algorithm would include the effects of AWGN, the end effects caused by using a finite-length equalizer, and studies of the regions of attraction of false minima.

V. The LTBOMB Algorithm

A. Algorithm Description

Here, we propose the first of two gradient descent-based blind algorithms. We rely on ideas from decades of research on classical blind equalization algorithms, but we try to exploit structure that they may ignore. For example, classical blind algorithms were not designed to exploit the fact that symbols are drawn from a biorthogonal set. Most blind algorithms inherently depend on higher order statistics, and this will be the case for our algorithms. As seen from Lemma 1, the second order statistics and hence the MMSE equalizer for a BOM signal are independent of $\boldsymbol{S}$. However, the fourth-order statistics of $x[Kn-i]$ will *not* be independent of $\boldsymbol{S}$, and, thus, we expect the shape of the cost surface of any candidate algorithm to depend on $\boldsymbol{S}$.

While samples of the chip process $x[Kn-i]$ could assume a range of dispersive values (depending on the choice of $\boldsymbol{S}$), the power of the BOM symbols is a constant. Thus, as the cost function for the first blind algorithm, termed "LTBOMB" (for **L**inear **T**ransversal equalizer adaptation for **BiO**rthogonal **M**odulation, **B**lindly), we choose to penalize the dispersion of the symbol power at the equalizer output

$$J_{\mathrm{LTB}}(\boldsymbol{f}) = E\left[\left(\|\boldsymbol{z}[n]\|_2^2 - 1\right)^2\right]. \tag{7}$$

Due to the invariance of the ℓ_2 norm to orthogonal transforms, we could equivalently penalize the dispersion of the correlator output power (i.e., since $\|\boldsymbol{S}^{\top}\boldsymbol{z}[n]\|_2 = \|\boldsymbol{z}[n]\|_2$). Taking the instantaneous gradient of (7) gives the update equation

$$\boldsymbol{f}_{\mathrm{LTB}}[n+1] = \boldsymbol{f}_{\mathrm{LTB}}[n] - \mu \boldsymbol{Y}[n]\left(\|\boldsymbol{z}[n]\|_2^2 - 1\right)\boldsymbol{z}[n].$$

While at first glance this simple cost function seems to ignore a lot of structure which is present in the BOM signal, we show that this is not the case and we draw connections to several other blind algorithms. The form of (7) looks much like the CMA, and not surprisingly reduces to the CMA when $K = 1$ since the vectors become scalars. The cost function shares even more similarity with the Vector CMA [20], though it is distinct in that our algorithm is driven by data that is not i.i.d., and it operates only once every K chips. Because of these two facts, the cost surface and algorithm performance will be quite different from the Vector CMA. Borrowing an idea from [21], and noting that $\boldsymbol{z}[n] = [z[Kn]\ldots z[Kn-K+1]]^{\top}$, we see that the cost function can be expanded as

$$J_{\mathrm{LTB}}(\boldsymbol{f}) = (1-K) + \sum_{i=0}^{K-1} E[(z^2[Kn-i]-1)^2] + \sum_{i=0}^{K-1}\sum_{j=0, j\neq i}^{K-1} E[z^2[Kn-i]z^2[Kn-j]]. \tag{8}$$

This gives an interesting interpretation since the second term is exactly the cyclostationary extension of the CMA cost (5), while the third term represents a penalty of the cross-correlation of

the squared equalizer output. Again, due to the invariance of the ℓ_2 norm to orthogonal transforms, we can equivalently replace $z[Kn-i]$ with the ith correlator output at time n, yielding the interpretation that the third term effectively penalizes the lack of biorthogonality in the signal.

B. Cumulant of a Vector Random Process

Before proceeding with the analysis of the algorithm, we first introduce the concept of the cumulant of a vector random process. Cumulants are useful in expanding the expectation in the LTBOMB cost function, resulting in an expression involving the channel/equalizer coefficients and the source statistics. While the chip-level random process in a BOM system is not i.i.d., the symbols are i.i.d., however, which motivates an analysis comprised of symbol vectors. Cumulants can be described for a vector random process in an analogous way to those for scalar random processes [22]. For a zero-mean i.i.d. vector random process $\boldsymbol{x}[n]$, the first order cumulant $\mathcal{C}_1(\boldsymbol{x}) = E[\boldsymbol{x}[n]]$ is simply the mean vector, the second-order cumulant $\mathcal{C}_2(\boldsymbol{x}) = E[\boldsymbol{x}[n]\boldsymbol{x}^\top[n]]$ is the symmetric covariance matrix, while the fourth-order cumulant is a supersymmetric fourth-order tensor. The fourth-order cumulant tensor for the BOM symbol $\boldsymbol{x}[n]$ can be expressed in terms of its corresponding chips as

$$\begin{aligned}
&[\mathcal{C}_4(\boldsymbol{x})]_{i_1,i_2,i_3,i_4} \\
&\quad \triangleq E[x[Kn-i_1]x[Kn-i_2]x[Kn-i_3]x[Kn-i_4]] \\
&\qquad - E[x[Kn-i_1]x[Kn-i_2]]E[x[Kn-i_3]x[Kn-i_4]] \\
&\qquad - E[x[Kn-i_1]x[Kn-i_3]]E[x[Kn-i_2]x[Kn-i_4]] \\
&\qquad - E[x[Kn-i_1]x[Kn-i_4]]E[x[Kn-i_3]x[Kn-i_2]].
\end{aligned} \tag{9}$$

Note that the cumulants themselves do not depend on n since the vectors are i.i.d. As with cumulants for scalar random processes, cumulants for vector random processes obey a linearity property [22]. For an i.i.d vector random process $\boldsymbol{x}[n]$ and a set of matrices $\boldsymbol{A}[n]$, the linearity property of cumulants is

$$\begin{aligned}
&\mathcal{C}_4\left(\sum_m \boldsymbol{A}[m]\boldsymbol{x}[m]\right) \\
&\qquad = \sum_m [\mathcal{C}_4(\boldsymbol{x}) \times_1 \boldsymbol{A}[m] \times_2 \boldsymbol{A}[m] \times_3 \boldsymbol{A}[m] \times_4 \boldsymbol{A}[m]]
\end{aligned}$$

where the symbol $\times_k$ denotes the k-mode tensor product [22]. We can use this property to express the cumulant of the equalizer output in terms of the cumulant of the source data. In the absence of AWGN, the equalizer output is $\boldsymbol{z}[n] = \boldsymbol{X}^\top[n]\boldsymbol{c}$. Using the linearity property and expanding the tensor product gives the individual tensor elements

$$\begin{aligned}
[\mathcal{C}_4(\boldsymbol{z})]_{i_1,i_2,i_3,i_4} = &\sum_{j_1,j_2,j_3,j_4} [\mathcal{C}_4(\boldsymbol{x})]_{j_1,j_2,j_3,j_4} \\
&\cdot \sum_m c[Km-i_1+j_1]c[Km-i_2+j_2] \\
&\cdot c[Km-i_3+j_3]c[Km-i_4+j_4].
\end{aligned} \tag{10}$$

In Appendix I, we express the cumulant tensor $\mathcal{C}_4(\boldsymbol{x})$ of a BOM source sequence in terms of $\boldsymbol{S}$.

C. Stability of ZF Solutions

We define the ZF solutions as those where the combined channel/equalizer response $\boldsymbol{c} = \boldsymbol{e}_{K\Delta}$. Note that, unlike more traditional equalization problems, our definition of the ZF solutions only includes responses with delays that are a multiple of K. In situations where the delay is not a multiple of K, the decision device will not be operating on the symbol boundary; as such a situation is undesirable, we exclude these impulse responses from the set of ZF solutions. As we will discuss later in Section V.F, the issues of equalization and symbol timing are tightly intertwined.

To investigate the ZF solutions, we first assume that there is no AWGN, so $\sigma_w^2 = 0$. We express the cost function (7) in terms of cumulants, and then use (10) to express the cost function in terms of the channel/equalizer coefficients, giving

$$\begin{aligned}
&J_{\mathrm{LTB}}(\boldsymbol{f}) \\
&\quad = \sum_{j_1,j_2} [\mathcal{C}_4(\boldsymbol{z})]_{j_1,j_1,j_2,j_2} + 2\sum_{j_1,j_2} [\mathcal{C}_2(\boldsymbol{z})]_{j_1,j_2}^2 \\
&\qquad + \left[\left(\sum_i [\mathcal{C}_2(\boldsymbol{z})]_{i,i}\right) - 1\right]^2 \\
&\quad = \sum_{i_1,i_2,i_3,i_4} [\mathcal{C}_4(\boldsymbol{x})]_{i_1,i_2,i_3,i_4} \sum_{j_1,j_2}\sum_p c[Kp-j_1+i_1] \\
&\qquad \cdot c[Kp-j_1+i_2]c[Kp-j_2+i_3]c[Kp-j_2+i_4] \\
&\qquad + \frac{2}{K^2}\sum_{j_1,j_2}\left(\sum_p c[p-j_1]c[p-j_2]\right)^2 \\
&\qquad + \left[\left(\sum_p c^2[p]\right) - 1\right]^2
\end{aligned} \tag{11}$$

where $i_1, i_2, i_3, i_4, j_1, j_2 \in \{0, \ldots, K-1\}$. The stationary points of the algorithm, i.e., those values of $\boldsymbol{f}$ that are minima, maxima, and saddle points, are the points where the derivative of the cost function with respect to $\boldsymbol{f}$ equals zero. We take the derivative of (11) to obtain

$$\frac{1}{4}\frac{\partial J_{\mathrm{LTB}}(\boldsymbol{f})}{\partial f[n]} = \sum_m h[m-n]\Lambda[m] \tag{12}$$

where

$$\begin{aligned}
\Lambda[m] \triangleq &\sum_{i_1,i_2,i_3,i_4} [\mathcal{C}_4(\boldsymbol{x})]_{i_1,i_2,i_3,i_4} \sum_{j_1,j_2}\sum_p \delta[m-Kp+j_1-i_1] \\
&\cdot c[Kp-j_1+i_2]c[Kp-j_2+i_3]c[Kp-j_2+i_4] \\
&+ \frac{2}{K^2}\sum_{j_1,j_2} c[m+j_1-j_2]\left(\sum_p c[p-j_1]c[p-j_2]\right) \\
&+ c[m]\left[\left(\sum_p c^2[p]\right) - 1\right]
\end{aligned} \tag{13}$$

and we have used the supersymmetry of the cumulant in the simplification. We can then write the gradient compactly as $\nabla J_{\mathrm{LTB}}(\boldsymbol{f}) = 4\boldsymbol{\mathcal{H}}\boldsymbol{\Lambda}$ where $\boldsymbol{\Lambda} = [\Lambda[0], \ldots, \Lambda[N_c-1]]^\top$,

so the stationary points are the solutions to the equation $\boldsymbol{\mathcal{H}\Lambda} = \mathbf{0}_{N_f-1}$. Continuing with the second derivative, we have

$$\frac{1}{4}\frac{\partial^2 J_{\mathrm{LTB}}(\boldsymbol{f})}{\partial f[n_1]\partial f[n_2]} = \sum_{m_1,m_2} h[m_1-n_1]h[m_2-n_2][\boldsymbol{\Psi}]_{m_1,m_2} \tag{14}$$

where

$$\begin{aligned}
&[\boldsymbol{\Psi}]_{m_1,m_2} \\
&\quad\triangleq \sum_{i_1,i_2,i_3,i_4} [\mathcal{C}_4(\boldsymbol{x})]_{i_1,i_2,i_3,i_4} \\
&\qquad\cdot \sum_{j_1,j_2}\sum_{p} [\delta[Kp-j_1+i_1-m_1]\delta[Kp-j_1+i_2-m_2] \\
&\qquad\cdot c[Kp-j_2+i_3]c[Kp-j_2+i_4] \\
&\qquad+ 2\delta[Kp-j_1+i_1-m_1]\delta[Kp-j_2+i_2-m_2] \\
&\qquad\cdot c[Kp-j_1+i_3]c[Kp-j_2+i_4]] \\
&\qquad+ \frac{2}{K^2}\sum_{j_1,j_2}\delta[j_1-j_2+m_1-m_2] \\
&\qquad\cdot \left(\sum_p c[p-j_1]c[p-j_2]\right) \\
&\qquad+ \frac{2}{K^2}\sum_{j_1,j_2} c[m_2+j_1-j_2] \\
&\qquad\cdot (c[m_1+j_1-j_2]+c[m_1-j_1+j_2]) \\
&\qquad+ 2c[m_1]c[m_2] + \delta[m_1-m_2]\left[\left(\sum_p c^2[p]\right)-1\right]
\end{aligned}$$

which allows us to write the Hessian matrix as $4\boldsymbol{\mathcal{H}\Psi\mathcal{H}}^{\top}$.

Before demonstrating that the ZF solutions are minima, we first prove two lemmas.

Lemma 2 (Property of Cumulants of Biorthogonal Signals): For a vector random process $\boldsymbol{x}[n] \in \mathbb{R}^K$ where the vectors are drawn i.i.d. from a complete biorthogonal set, and some $i_1, i_2 \in \{0,\ldots,K-1\}$

$$\sum_{j=0}^{K-1}[\mathcal{C}_4(\boldsymbol{x})]_{i_1,i_2,j,j} = \begin{cases} -\frac{2}{K^2} & \text{for } i_1 = i_2 \\ 0 & \text{otherwise.} \end{cases}$$

Proof: Let the underlying orthogonal basis be described by the columns of the square matrix $\boldsymbol{S}$. From the cumulant definition (9) and Lemma 1 we have

$$\begin{aligned}
&\sum_{j=0}^{K-1}[\mathcal{C}_4(\boldsymbol{x})]_{i_1,i_2,j,j} \\
&\quad= \sum_{j=0}^{K-1}(E[x[-i_1]x[-i_2]x^2[-j]] \\
&\qquad- E[x[-i_1]x[-i_2]]E[x^2[-j]] \\
&\qquad- 2E[x[-i_1]x[-j]]E[x[-i_2]x[-j]]) \\
&\quad= -\left(\frac{2}{K^2}+\frac{1}{K}\right)\delta[i_1-i_2] \\
&\qquad+ \frac{1}{K}\sum_{j,k=0}^{K-1}[\boldsymbol{S}]_{i_1,k}[\boldsymbol{S}]_{i_2,k}[\boldsymbol{S}]_{j,k}^2 \\
&\quad= -\frac{2}{K^2}\delta[i_1-i_2]
\end{aligned}$$

where the last line uses the orthonormality of columns of $\boldsymbol{S}$. ■

Lemma 3 (Sum of Delta Functions): The sum

$$\sum_{i_1,i_2=0}^{K-1}\delta[K\Delta+i_1-i_2-m] \le K$$

with equality only when $m = K\Delta$.

Proof: Via substitution, we can write the left-hand side (LHS) as

$$\sum_{i_1=0}^{K-1}\sum_{i_2=m-K\Delta}^{m-K\Delta+K-1}\delta[i_1-i_2]$$

which is simply a measure of the number of common elements in the two length-K sequences $\{0,\ldots,K-1\}$ and $\{m-K\Delta,\ldots,m-K\Delta+K-1\}$. Clearly, the largest possible number of common elements is K which is only possible when $m = K\Delta$. ■

Theorem 1: The ZF solutions $\boldsymbol{c} = \boldsymbol{e}_{K\Delta}$ are stationary points of the LTBOMB cost function (7).

Proof: Substituting $\boldsymbol{c} = \boldsymbol{e}_{K\Delta}$ into (12) gives the gradient at the ZF solutions as

$$\begin{aligned}
&\frac{1}{4}\frac{\partial J_{\mathrm{LTB}}(\boldsymbol{f})}{\partial f[n]}\bigg|_{\boldsymbol{c}=\boldsymbol{e}_{K\Delta}} \\
&\quad= \frac{2}{K}h[K\Delta-n] \\
&\qquad+ \sum_{i_1,i_2,i_3}[\mathcal{C}_4(\boldsymbol{x})]_{i_1,i_2,i_3,i_3}h[K\Delta-i_2+i_1-n] \\
&\quad= 0
\end{aligned}$$

where the last line follows from Lemma 2. ■

Theorem 2: The ZF solutions $\boldsymbol{c} = \boldsymbol{e}_{K\Delta}$ are stable minima of the LTBOMB cost function (7).

Proof: Having already established that the ZF solutions are stationary points in Theorem 1, we show they are minima by considering the definiteness of the Hessian matrix $4\boldsymbol{\mathcal{H}\Psi\mathcal{H}}^{\top}$. Since $\boldsymbol{\mathcal{H}}$ is full rank, we only need to consider the positive definiteness of $\boldsymbol{\Psi}$. That is, for any $\boldsymbol{q} \in \mathbb{R}^{N_c-1}$ where $\boldsymbol{q} \neq \mathbf{0}_{N_c-1}$, we need to show that $\boldsymbol{q}^{\top}\boldsymbol{\Psi q} > 0$. Substituting $\boldsymbol{c} = \boldsymbol{e}_{K\Delta}$ and (20) into (15) gives

$$\begin{aligned}
&\boldsymbol{q}^{\top}\boldsymbol{\Psi q}|_{\boldsymbol{c}=\boldsymbol{e}_{K\Delta}} \\
&\quad= \frac{2}{K}\sum_{\ell}\left[\sum_m\sum_{i_1,i_2} q[m][\boldsymbol{S}]_{i_1,\ell}\right. \\
&\qquad\left.\cdot\ [\boldsymbol{S}]_{i_2,\ell}\delta[K\Delta+i_1-i_2-m]\right]^2 \\
&\qquad+ \frac{2}{K^2}\sum_m q^2[m]\left[K-\sum_{i_1,i_2}\delta[K\Delta+i_1-i_2-m]\right].
\end{aligned}$$

The first term is obviously nonnegative, and the second term is nonnegative via Lemma 3. To show strict positive definiteness, we only need to show that the two terms cannot be simultaneously zero. From Lemma 3, we see that the bracketed quantity in the second term can only be zero when $m = K\Delta$; hence, the entire second term can only be zero when $\boldsymbol{q} = \alpha\boldsymbol{e}_{K\Delta}$ for any $\alpha \neq 0$ (to avoid the trivial case). For this choice of $\boldsymbol{q}$, we have $\boldsymbol{q}^{\top}\boldsymbol{\Psi q} = 2\alpha^2 > 0$. Thus, $\boldsymbol{\Psi}$ is positive definite, so the Hessian is positive definite and the ZF solutions are minima.

TABLE I
LTBOMB STATIONARY POINTS FOR EXAMPLE 2

$\boldsymbol{c}^\top$	type
$[0,0,0]$	maximum
$\pm[1,0,0], \pm[0,0,1]$	minima
$\pm[0,\sqrt{2/3},0]$	saddle points
$\pm[\sqrt{1/3},0,-\sqrt{1/3}]$	degenerate saddle points

D. Example 2: Case of Global Convergence

While we have demonstrated local convergence, we cannot make any claims about the global performance of the algorithm in general. Similar insurmountable difficulties were encountered in the global convergence analysis of the Vector CMA, due to the presence of the cross terms. It is precisely these cross terms, i.e., the last term in (8), which complicate analysis of our algorithm.

In light of this difficulty, we resort to considering a low-dimensional numerical example, as part of the strategy outlined in Section IV. Again, since this is purely an illustrative example, we operate exclusively in the combined channel/equalizer domain $\boldsymbol{c}$, thereby avoiding end effects that are known to plague finite-length chip-rate blind equalizers. We recall that stationary points appearing in this low-dimensional example will persist into higher dimensions. Furthermore, any of these stationary points delayed by a multiple of K will also be stationary points.

Letting $\boldsymbol{S} = \boldsymbol{I}_2$ and $N_c = 3$, the stationary points can be found by setting the non-trivial elements of the resulting gradient to zero, resulting in the system of equations

$$
\begin{aligned}
&c^3[0] + \frac{3}{2}c^2[1]c[0] + \frac{1}{2}c^2[1]c[2] + 2c^2[2]c[0] - c[0] = 0\\
&c[0]c[1]c[2] + \frac{3}{2}c^2[2]c[1] + \frac{3}{2}c^2[0]c[1] + \frac{3}{2}c^3[1] - c[1] = 0\\
&c^3[2] + 2c^2[0]c[2] + \frac{1}{2}c^2[1]c[0] + \frac{3}{2}c^2[1]c[2] - c[2] = 0\\
&c[0]c[2]c[1] = 0.
\end{aligned}
$$

Using Gröbner bases, we can solve for the locations of all stationary points exactly, and they have been tabulated in Table I. For this particular example, we see that we can expect global convergence to the desired solution, as minima occur only at the ZF solutions. In spite of the fact that the single-spike impulse response $\boldsymbol{c} = \pm[0,1,0]^\top$ results in a situation with no ISI, we note that it is not a minimum of the algorithm. Such an impulse response represents a delay that is not a multiple of K, and therefore does not appropriately align the chips to the symbol boundary. We also note the appearance of degenerate saddle points. A degenerate saddle point is one where the Hessian is singular, which implies the cost surface is very flat, and the adaptive algorithm will likely suffer convergence speed problems as it passes through this region. In contrast to the DD-LMS example in Section IV, we observe no undesirable local minima here, and for this low-dimensional example the only minima of LTBOMB cost surface are the ZF solutions.

E. Example 3: Case of False Local Minima

Though Theorem 2 established that all ZF solutions are minima, we would ultimately like to know if all minima are ZF solutions. While Example 2 provided some hope, we now investigate stationary points that arise in impulse responses with larger lengths of contiguous nonzero taps while maintaining $\boldsymbol{S} = \boldsymbol{I}_2$. For larger values of N_c, the system of equations is too complex for the use of Gröbner bases in calculating all the stationary points. However, during the course of experimentation with the algorithm, we noticed that once the number of nonzero taps grew to $N_c \geq 6$ we would occasionally observe poor algorithm performance near impulse responses of the form $\boldsymbol{c}^* = \pm[0,\alpha,\beta,\gamma,-\beta,\alpha]^\top$. The symmetry of this class of impulse responses enabled us to reduce the 6-parameter problem to a 3-parameter problem so that we could solve for the exact locations of stationary points using Gröbner bases. Upon categorizing the exact location and character of all the stationary points in this class, we found one that was a minimum. The exact expressions for α,β,γ at this stationary point are unwieldy, but their approximate values are $(\alpha,\beta,\gamma) \approx (0.1741, 0.4718, 0.5859)$, and we note that the Hessian was found to have strictly positive eigenvalues at this point. As this stationary point is clearly not a ZF solution, this impulse response is indeed a false local minimum of the algorithm.

The low-dimensionality of Example 2 enabled us to calculate all of the stationary points for impulse responses with 3 or fewer contiguous nonzero taps. Since the only local minima in Example 2 were ZF solutions, Example 2 provided some hope of global convergence to a ZF solution. For longer impulse responses, we can make no such claims due to the inability to solve larger system of polynomials. In spite of the fact that categorization of all stationary points is not possible for longer impulse responses, however, we have stumbled upon a false local minimum that appears once N_c is increased beyond 6, and we recall that this minimum will persist in higher dimensions. This counterexample shows that we cannot expect global convergence to a ZF solution.

F. On the Interaction of Symbol Timing and Equalization

The problems of symbol timing (i.e., finding the symbol boundary within a chip stream) and equalization are tightly related. Equalization algorithms with training data, like LMS, are effectively given the symbol timing information via the training data. Blind algorithms, however, do not have such information at their disposal.

As mentioned in Section V-C, the ZF solutions are those responses that amount to a delay that is a multiple of K. When such a response is attained by the equalizer, symbol timing has effectively been acquired. If these ZF solutions were the only minima of the algorithm, we could rely on the blind algorithm to acquire the symbol timing and perform equalization simultaneously. However, we know this is not the case, as was shown by the false local minima of Example 3.

We conjecture that the false local minima of the algorithm are caused in part by the symbol timing ambiguity. Thus, we suggest an *ad hoc* scheme that attempts to simultaneously acquire the symbol timing and avoid the false minima. The rationale for the scheme is based on the fact that, in the combined

channel/equalizer domain, any impulse response that is a (desired or undesired) minimum of the algorithm—shifted by K taps—is also a local minimum. Conversely, tap shifts that are *not* a multiple of K are typically *not* stable points of the algorithm. Thus, we consider operating K equalizers in parallel, each being updated once per symbol, but each operating on a different one of the K possible symbol boundary hypotheses. We begin adaptation of the K equalizers, and after it appears that some or all of the equalizers have settled (i.e., once the magnitude of the update terms are sufficiently small), we select the one equalizer setting with lowest estimated cost by computing a sample average of (7), and we discard the other $K-1$ equalizers. While heuristically it seems this may help mitigate the symbol timing ambiguity problem, we cannot verify the validity of this scheme without investigating the regions of attraction of the algorithm, which is far beyond the scope of this paper. We will, however, show simulations of this scheme in Section VII.

VI. The Trombone Algorithm

A. Algorithm Description

As aforementioned, the two most popular classical blind equalization algorithms are the CMA and the SWA. Since the LTBOMB algorithm draws largely from the spirit of the CMA, a sensible next step is to consider how we might apply the SWA philosophy to equalization of BOM signals.

When the channel has been appropriately equalized, the correlator output $\boldsymbol{S}^\top \boldsymbol{z}[n]$ should be a canonical unit vector (modulo sign). While the previous algorithm dealt exclusively with the ℓ_2 norm of the correlator output, we now consider other norms of the correlator output. In particular, we observe that for any ℓ_p norm, we desire $\|\boldsymbol{S}^\top \boldsymbol{z}[n]\|_p = 1$. As shown in [23], for any $p < q$, any exponent m, and any vector $\boldsymbol{x}$, we have

$$\|\boldsymbol{x}\|_p^m \geq \|\boldsymbol{x}\|_q^m$$

with equality when $\boldsymbol{x}$ is a canonical unit vector. This fact is the motivation for our next algorithm, termed **T**he **R**ecovery **O**f $\boldsymbol{M}$-ary **Bi****O**rthogonal signals via p-**N**orm **E**quivalence (TROMBONE) and having cost function

$$J_{\mathrm{TRO}}(\boldsymbol{f}) = E\left[\|\boldsymbol{S}^\top \boldsymbol{z}[n]\|_p^m - \|\boldsymbol{S}^\top \boldsymbol{z}[n]\|_q^m\right]. \tag{16}$$

First, we note that when the equalizer is operating correctly so that the correlator outputs are "perfect," the cost will be zero as hoped. We also note that the trivial solution has zero cost, so to avoid this solution we need to impose a constraint on the algorithm. Here, we choose to constrain the equalizer output power to be $E[\|\boldsymbol{z}[n]\|_2^2] = E[\|\boldsymbol{x}[n]\|_2^2]$, which amounts to $\boldsymbol{c}^\top \boldsymbol{c} = 1$ in the absence of noise. Note that other choices of m, p, q with $p < q$ may also lead to suitable algorithms, but we focus on the case $p = 2, m = q = 4$ due to its similarity with the LTBOMB algorithm, and its relative ease of implementation. Expanding the cost in terms of cumulants, the cost becomes

$$\begin{aligned}
&J_{\mathrm{TRO}}(\boldsymbol{f}) \\
&\quad = E\left[\|\boldsymbol{S}^\top \boldsymbol{z}[n]\|_2^4 - \|\boldsymbol{S}^\top \boldsymbol{z}[n]\|_4^4\right] \qquad (17)\\
&\quad = \sum_{\substack{i,j,\\ i\neq j}} [\mathcal{C}_4(\boldsymbol{S}^\top \boldsymbol{z})]_{i,i,j,j} + 2\sum_{i,j} [\mathcal{C}_2(\boldsymbol{S}^\top \boldsymbol{z})]_{i,j}^2 \\
&\qquad + \left(\sum_i [\mathcal{C}_2(\boldsymbol{S}^\top \boldsymbol{z})]_{i,i}\right)^2 - 3\sum_i [\mathcal{C}_2(\boldsymbol{S}^\top \boldsymbol{z})]_{i,i}^2 \\
&\quad = \sum_{i_1,i_2,i_3,i_4} [\mathcal{C}_4(\boldsymbol{x})]_{i_1,i_2,i_3,i_4} \sum_{j_1,j_2}\sum_p c[Kp+i_1-j_1] \\
&\qquad \cdot c[Kp+i_2-j_1]c[Kp+i_3-j_2]c[Kp+i_4-j_2] \\
&\qquad - \sum_{i_1,i_2,i_3,i_4} [\mathcal{C}_4(\boldsymbol{x})]_{i_1,i_2,i_3,i_4} \\
&\qquad \cdot \sum_{j_1,j_2,j_3,j_4} \sum_k [[\boldsymbol{S}]_{j_1,k}[\boldsymbol{S}]_{j_2,k}[\boldsymbol{S}]_{j_3,k}[\boldsymbol{S}]_{j_4,k} \\
&\qquad \cdot \sum_p c[Kp+i_1-j_1]c[Kp+i_2-j_2]c[Kp+i_3-j_3] \\
&\qquad \cdot c[Kp+i_4-j_4]\Bigg] \\
&\qquad - \frac{3}{K^2}\sum_j \left[\sum_p \sum_{i_1,i_2} [\boldsymbol{S}]_{i_1,j}[\boldsymbol{S}]_{i_2,j} c[p-i_1]c[p-i_2]\right]^2 \\
&\qquad + \frac{2}{K^2}\sum_{i_1,i_2}\left[\sum_p c[p-i_1]c[p-i_2]\right]^2 + \left(\sum_p c^2[p]\right)^2.
\end{aligned}$$

The constraint that we have imposed, i.e., $\boldsymbol{c}^\top \boldsymbol{c} = 1$, is a function of the combined channel/equalizer response. In practice, we do not have knowledge of $\boldsymbol{c}$. However, if we assume spectral prewhitening has been performed before equalization as in [16], thereby assuming the effective channel is white, the constraint becomes

$$\boldsymbol{c}^\top \boldsymbol{c} = 1 \Rightarrow \boldsymbol{f}^\top \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^\top \boldsymbol{f} = \boldsymbol{f}^\top \boldsymbol{f} = 1$$

so that normalization of the equalizer taps ensures that we will meet the constraint. The instantaneous gradient gives the algorithm update equation with a normalization step as

$$\begin{aligned}
&\boldsymbol{f}'_{\mathrm{TRO}}[n+1] \\
&\quad = \boldsymbol{f}_{\mathrm{TRO}}[n] - \mu \boldsymbol{Y}[n](\boldsymbol{z}[n]^\top \boldsymbol{z}[n]\boldsymbol{I} \\
&\qquad - \boldsymbol{S}\mathrm{diag}(\boldsymbol{S}^\top \boldsymbol{z}[n])^2 \boldsymbol{S}^\top)\boldsymbol{z}[n] \\
&\boldsymbol{f}_{\mathrm{TRO}}[n+1] \\
&\quad = \boldsymbol{f}'_{\mathrm{TRO}}[n+1] / \sqrt{\boldsymbol{f}'^\top_{\mathrm{TRO}}[n+1]\boldsymbol{f}'_{\mathrm{TRO}}[n+1]}
\end{aligned}$$

where $\mathrm{diag}(\boldsymbol{x})$ is the square diagonal matrix having $\boldsymbol{x}$ along its diagonal.

We see this algorithm does appear to have some vague similarities with the SWA in that we have a constrained cost function, motivated by the fact that $\|\boldsymbol{S}^\top \boldsymbol{z}[n]\|_2^4 \geq \|\boldsymbol{S}^\top \boldsymbol{z}[n]\|_4^4$ with equality when ISI has been eliminated. The algorithm also shares some similarity with the Shtrom-Fan algorithms [24] in that it involves a difference of two ℓ_p-norms. Beyond this similarity, however, our algorithm is fundamentally different from the Shtrom-Fan algorithms. Our cost function is a function of the difference of two norms of the correlator output data. The Shtrom-Fan cost function, on the other hand, involves the difference of two norms of the combined channel/equalizer response, and relies on properties of scalar cumulants to map

the cost function from the combined channel/equalizer space to the equalizer space. Thus, the Shtrom-Fan class of algorithms implicitly requires the data to be i.i.d. at the chip level, which is not the case for BOM.

We will now show that the ZF solutions are stationary points. Taking the derivative of the unconstrained TROMBONE cost function gives

$$\frac{1}{4}\frac{\partial J_{\mathrm{TRO}}(\boldsymbol{f})}{\partial f[n]} = \sum_m h[p-n]\Lambda[m]$$

where

$$\begin{aligned}\Lambda[m] = &\sum_{i_1,i_2,i_3,i_4} [\mathcal{C}_4(\boldsymbol{x})]_{i_1,i_2,i_3,i_4} \\ &\cdot \sum_{k_1,k_2}\sum_p \delta[Kp+i_1-k_1-m] \\ &\cdot c[Kp+i_2-k_1]c[Kp+i_3-k_2]c[Kp+i_4-k_2] \\ &- \sum_{i_1,i_2,i_3,i_4} [\mathcal{C}_4(\boldsymbol{x})]_{i_1,i_2,i_3,i_4} \\ &\cdot \sum_{k_1,k_2,k_3,k_4}\sum_j \Bigg[[\boldsymbol{S}]_{k_1,j}[\boldsymbol{S}]_{k_2,j}[\boldsymbol{S}]_{k_3,j}[\boldsymbol{S}]_{k_4,j} \\ &\cdot \sum_p \delta[Kp+i_1-k_1-m]c[Kp+i_2-k_2] \\ &\cdot c[Kp+i_3-k_3]c[Kp+i_4-k_4]\Bigg] \\ &- \frac{3}{K^2}\sum_{i_1,i_2,i_3,i_4}\sum_j [\boldsymbol{S}]_{i_1,j}[\boldsymbol{S}]_{i_2,j}[\boldsymbol{S}]_{i_3,j}[\boldsymbol{S}]_{i_4,j} \\ &\cdot \sum_p c[m+i_1-i_2]c[p-i_3]c[p-i_4] \\ &+ \frac{2}{K^2}\sum_p\sum_{k_1,k_2} c[m+k_1-k_2]c[p-k_1]c[p-k_2] \\ &+ c[m]\left(\sum_p c^2[p]\right)\end{aligned}$$

Theorem 3: The ZF solutions $\boldsymbol{c} = \boldsymbol{e}_{K\Delta}$ are stationary points of the TROMBONE cost function (17).

Proof: First, note that $\boldsymbol{c} = \boldsymbol{e}_{K\Delta}$ satisfies the unit-norm constraint. We have

$$\begin{aligned}\Lambda[p] = &-\sum_{i_1,i_2,i_3,i_4} [\mathcal{C}_4(\boldsymbol{x})]_{i_1,i_2,i_3,i_4} \\ &\cdot \sum_{j,k} [\boldsymbol{S}]_{k,j}[\boldsymbol{S}]_{i_2,j}[\boldsymbol{S}]_{i_3,j}[\boldsymbol{S}]_{i_4,j}\delta[K\Delta+i_1-k-p] \\ &+ \left(1-\frac{3}{K}\right)\delta[K\Delta-p] \\ = &\ 0\end{aligned}$$

where we substituted (20) in Appendix I for the cumulant tensor. ∎

TABLE II
TROMBONE STATIONARY POINTS FOR EXAMPLE 4

$\boldsymbol{c}^{\mathsf{T}}$	type
$\pm[1,0,0], \pm[0,0,1]$	minima
$\pm[\sqrt{1/2},\ 0,\ -\sqrt{1/2}], \pm[\sqrt{1/2},\ 0,\ \sqrt{1/2}]$	saddle points
$\pm[0,\ 1,\ 0]$	degenerate saddle pts

We could have equivalently used the method of Lagrange multipliers, but since the ZF solutions are stationary points and simultaneously satisfy the unit-norm constraint, the Lagrange multiplier is zero. Simple examination of the Hessian eigenvalues of constrained algorithms does not allow us to generically classify these stationary points as stable minima; we would need to resort to perturbation analysis or reparameterization of the cost function as we will now show in a low-dimensional example. From our observations through simulation, we believe the stationary points at the ZF solutions are locally stable, as we have tested this claim on thousands of prewhitened channels.

B. Example 4: Stationary Points in Low Dimensions

As we did in Example 2 for LTBOMB, we now classify all the stationary points for the same numerical example. Again working in the combined channel/equalizer domain $\boldsymbol{c}$, we let $\boldsymbol{S} = \boldsymbol{I}_2$ and $N_c = 3$. Because the unit norm tap constraint is difficult to apply in the $\boldsymbol{c}$ domain, we need a transform to reparameterize the cost function in a coordinate system that permits us to easily apply the constraint. We can reparameterize the function of $\boldsymbol{c}$ in polar coordinates, having $N_c - 1$ rotation angles $\{\theta_0, \ldots, \theta_{N_c-2}\}$ and one radius r. One possibility [25] for parameterizing $\boldsymbol{c}$ in polar form is to choose (18), shown at the bottom of the page, so that by fixing $r = 1$, any arbitrary unit-norm $\boldsymbol{c}$ may be reached.

Substituting (18) into the TROMBONE cost function with $\boldsymbol{S} = \boldsymbol{I}_2$, setting the gradient to zero, and zeroing all but $N_c = 3$ taps results in a system of equations in two parameters, θ_0 and θ_1. As before, we can solve for the locations of all stationary points exactly, and they have been tabulated in Table II. The results here for the TROMBONE algorithm are very similar to those for the LTBOMB algorithm. We again see that we can expect global convergence to the desired solution, as minima occur only at the ZF solutions. Furthermore, we again note the appearance of degenerate saddle points, though they are in a slightly different location.

C. Example 5: False Minima of TROMBONE

As was done for the LTBOMB in Section V-E, we also conducted a (nonexhaustive) numerical search for stationary points of the TROMBONE algorithm. Again, choosing $\boldsymbol{S} = \boldsymbol{I}_2$, we found that indeed, at least for this example,

$$c[n] = \begin{cases} r\sin(\theta_n)\prod_{j=0}^{n-1}\cos(\theta_j) & \text{for } 0 \le n \le N_c-2 \\ r\cos(\theta_{N_c-1})\prod_{j=0}^{N_c-2}\cos(\theta_j) & \text{for } n = N_c-1 \end{cases} \tag{18}$$

the ZF solutions are minima of the algorithm. In addition, we found false local minima with impulse response $\boldsymbol{c} = \pm[0, 0.2973, 0.5425, 0.4844, -0.5425, 0.2973]^\top$ which also appear every 2-tap shift. This impulse response is very similar to the false minimum for LTBOMB, and so we see that both of these blind algorithms have an inherent problem. This is peculiar since, at least on the surface, the two cost functions attempt to restore different properties of the BOM signal. The fact that two blind algorithms, both designed with different criteria in mind, suffer from very similar spurious local minima raises questions about the existence of an algorithm exhibiting global convergence. Again, we conjecture that these false minima arise in part due to the symbol timing ambiguity. Use of the parallel equalizer scheme described in Section V-F may help circumvent this issue.

VII. Numerical Examples

A. Visualizing the LTBOMB Cost Surface

We once again consider the noiseless case with $\boldsymbol{S} = \boldsymbol{I}_2$ and $N_c = 2$, and we plot a two–dimensional (2-D) slice[3] of the cost surface contours. The cost surface is shown in Fig. 3, where we observe the presence of a maximum at the origin and only 2 minima, those at the ZF solution $\boldsymbol{c} = \pm[1, 0]^\top$. As expected minima do not also occur at $\boldsymbol{c} = \pm[0, 1]^\top$, but instead there are saddle points in that region. The fact that these are not minima implies that the proposed algorithm can acquire the symbol timing since, as hoped, minima only occur for delays that are a multiple of K. We refer the reader back to Fig. 2, where the cost surface for DD-LMS with the same system parameters exhibited false local minima.

Since the cost function depends on the cumulants of the underlying signal basis $\boldsymbol{S}$, we expect the cost surface to look different for different choices of $\boldsymbol{S}$. If we change the underlying orthogonal basis so that $\boldsymbol{S}$ is the 2×2 Hadamard matrix, we arrive at the cost surface also shown in Fig. 3. In addition to the maximum at the origin and the minima at the ZF solutions, we observe that minima appear at $\boldsymbol{c} = \pm[0, 1]^\top$, which corresponds to a delay that is not a multiple of K. Note that for this choice of $\boldsymbol{S}$, the chip statistics are identical to BPSK since the symbols become $\pm[1, 1]/\sqrt{2}, \pm[1, -1]/\sqrt{2}$. Thus, when $\boldsymbol{S}$ is chosen to be the Hadamard matrix, the algorithm has no hope of recovering the symbol timing. Nevertheless, we emphasize that all choices of $\boldsymbol{S}$ will exhibit local convergence to ZF solutions as proven by Theorem 2.

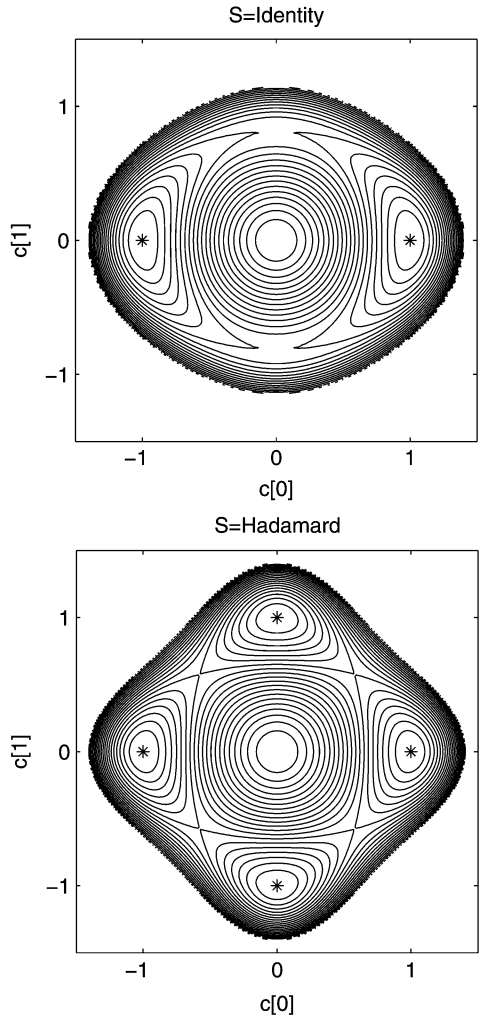


Fig. 3. LTBOMB cost surface for different choices of $\boldsymbol{S}$.

B. Visualizing the Trombone Cost Surface

Again considering the noiseless case with $\boldsymbol{S} = \boldsymbol{I}_2$ and $N_c = 2$, we have plotted the unconstrained TROMBONE cost surface in Fig. 4. The unit norm constraint will force the algorithm to stay on contour indicated by the dotted circle. For the case of $\boldsymbol{S} = \boldsymbol{I}$, the cost surface is a single trough with a bulb at the origin, and the cost is zero along the $c[0]$ axis. For the case of the Hadamard matrix, there are two troughs, with zero cost along both axes.

Identifying the stationary points from the unconstrained cost, however, is not easy. Thus, similar to what was done in Section VI-B, we transform the 2-parameter cost plot into polar coordinates with a single rotation angle. Looking at the cost function in polar coordinates, where the angle $\theta = \tan^{-1}(c[1]/c[0])$, we can more easily see the stationary points as shown in Fig. 5. We see that both examples result in minima at the ZF solutions, as well as minima at $\boldsymbol{c} = \pm[0, 1]^\top$, though this latter minimum is quite shallow for the case of $\boldsymbol{S} = \boldsymbol{I}$.

C. Simulation Near Zero-Forcing Solutions

To provide verification that our analysis of the local behavior of these algorithms is valid, we consider a simulation where we initialize the algorithms in a ball around a ZF solution. We operate in a noiseless scenario, a channel with impulse response $\boldsymbol{h} = [-0.4, 0.84, 0.336, 0.1344, 0.0538, 0.0215]^\top$, and we have chosen $N_f = 30$ equalizer taps. As before, we choose the signal bases $\boldsymbol{S} = \boldsymbol{I}_2$. The equalizer corresponding to the ZF solution at the chosen delay of $\Delta = 3$ has approximate impulse response shown at the bottom of the next page. Note that the channel response is approximately white (i.e., $\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^\top \approx \boldsymbol{I}$), and the corresponding ZF equalizer satisfies $\boldsymbol{f}_{\mathrm{ZF}}^\top \boldsymbol{f}_{\mathrm{ZF}} = 1$.

[3]While in general, stationary points of slices may change their character in higher dimensions, those shown here do not.

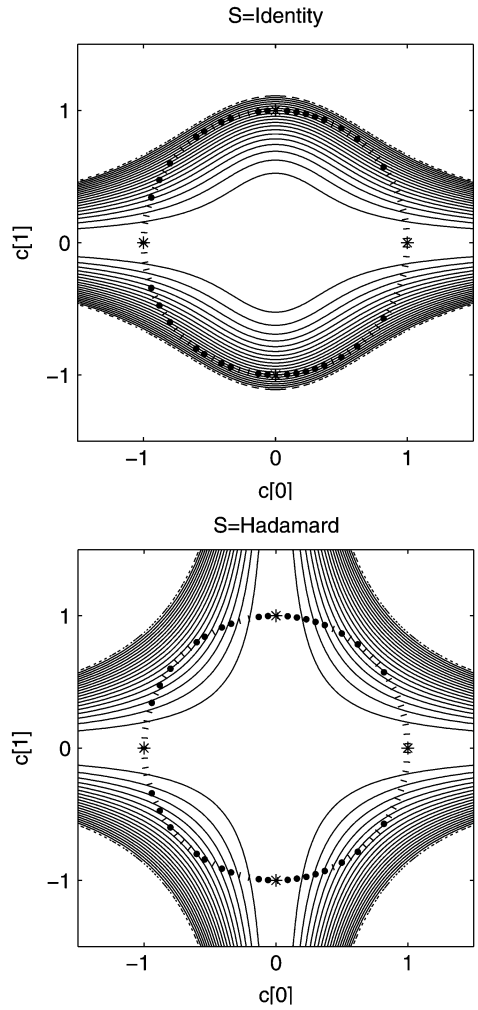


Fig. 4. Unconstrained TROMBONE cost surface for different choices of $\boldsymbol{S}$.

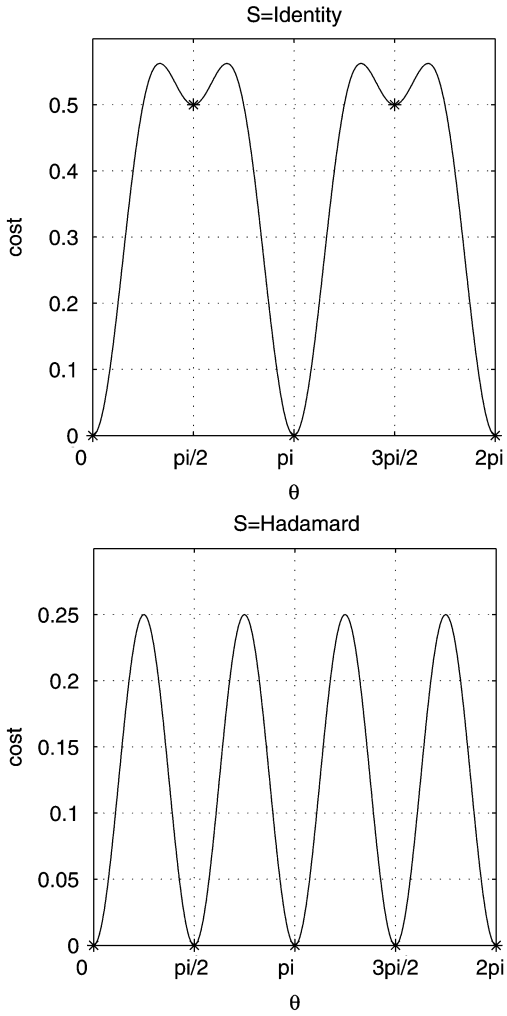


Fig. 5. Polar representation of TROMBONE cost surface.

As an initialization, we chose 1000 points uniformly distributed in a ball around the ZF solution, and we observed the ability of the algorithms to converge to a ZF solution as the radius of the ball is increased. We ran the algorithms for 10 000 symbols at each of the 1000 initializations, and declared the algorithm to have converged if the MSE was less than 10^{-3}. As shown in Fig. 6, all of the algorithms converge to the ZF solution when the size of the ball of initializations around the ZF solution is less than 1, thus, verifying our analysis of local convergence. For larger radii beyond 1, which we can hardly consider to be "local" to the ZF solutions, we see that all algorithms still converge with a fairly high percentage, though possibly to a ZF solution corresponding to some other delay. Furthermore, the blind algorithms both outperform DD-LMS. The curves are all monotone non-increasing, albeit with some jumps due to the nonlinear nature of the cost surfaces, and for very large radii we note that the convergence percentages eventually reach a constant value. This is because, beyond a sufficiently large initialization ball radius, there is no notion of locality; the ball grows to encompass the entire space, and so increasing the ball radius further has no effect on convergence percentage.

We note that the lack of a good adaptive algorithm for prewhitening and the added complexity of prewhitening suggests the practical superiority of LTBOMB over TROMBONE; this matches the current wisdom regarding conventional blind equalization of BPSK via the CMA versus the SWA.

D. Simulation of a Practical Situation

Continuing with the strategy for algorithm assessment outlined in Section IV, we now consider the use of LTBOMB in a practical channel, which serves to justify our claims of local convergence, as well as the superior performance of our algorithm over decision-directed LMS. We focus on LTBOMB here because, as aforementioned, TROMBONE requires a prewhitening filter and may not be as useful in practice. It

$$\boldsymbol{f}_{\mathrm{ZF}} \approx [0.009, 0.022, 0.054, 0.134, 0.336, 0.840, -0.399, 0.002, 0.003, 0.005, 0.006, -0.003, 0, \ldots, 0]^{\top}.$$

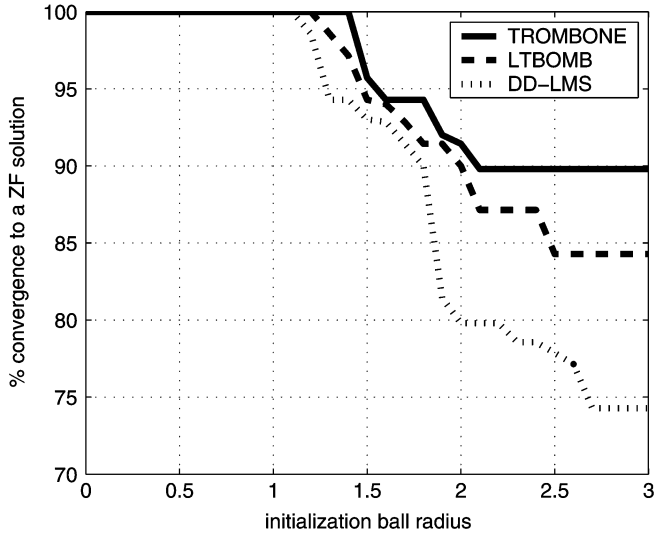


Fig. 6. Convergence percentage vs. initialization distance from ZF solution.

should be noted that the analysis up to now has ignored the effects of noise and has largely considered the cost surface in the combined channel/equalizer space, so these simulations will also provide some faith that our algorithm performs well in the presence of real channels with AWGN. As a channel model, we choose one based on the Saleh-Valenzuela [26] model of an indoor channel. The IEEE 802.153a committee has constructed a set of such channels [27] based on this model, and we have selected to use their channel model CM3 which models a nonline of sight indoor environment over distances of 4–10 m. To convert the channels to a baseband equivalent tapped-delay line channel, we chose a carrier frequency of 3 GHz and a sample period of 10 ns, and performed low-pass filtering using a raised-cosine filter with a rolloff factor of 0.5. We then zeroed any leading or trailing taps with energy less than 20 dB below the peak, which resulted in a set of baseband equivalent channels with lengths ranging from $4 \leq N_h \leq 13$, with the average channel length being 7.8 taps. Note that a large number of these channels were nonminimum phase, and often had roots near the unit circle. Furthermore, these channels were most definitely not white.

For the BOM source signal, we chose $\boldsymbol{S} = \boldsymbol{I}_2$ so $K = 2$. Furthermore, the SNR was set at 8 dB, the equalizer had length $N_f = 30$, and we used a centered double spike initialization. While an analysis of initialization strategies and regions of convergence is beyond the scope of this paper, we observed that for a BOM system with K chips per symbol and $\boldsymbol{S} = \boldsymbol{I}_K$, a K-spike initialization seemed to improve convergence over a single-spike initialization.

After generating 1000 channel realizations, we ran LTBOMB on each channel, as well as DD-LMS. An equalizer was declared to have converged near the MMSE solution if its MSE was within 1 dB of the MSE of the nearest MMSE solution. In addition, we considered two equalizer setups: one with just a single equalizer, and another setup with two equalizers operating in parallel as described in Section V-F. In the setup with parallel equalizers, the equalizer with lower sample average cost was selected upon convergence, and compared with the nearest MMSE solution. The simulation results are shown in Table III. We see that the LTBOMB algorithm does quite well, consistently beating the decision-directed algorithm. In addition, we note that the use of two parallel equalizers does buy us some improvement in convergence, with the LTBOMB converging nearly always. The decision-directed algorithm, too, sees a benefit from the use of two parallel equalizers. As a test, we also ran the TROMBONE algorithm without the required prewhitening filter; unsurprisingly, it only converged 24.3% of the time with a single equalizer, and 36.7% in the dual equalizer setup. Again, we stress that LTBOMB may be more useful in practice due to the lack of need for a blind prewhitener.

TABLE III
SIMULATION RESULTS

Algorithm	percent converged to MMSE (single eq.)	percent converged to MMSE (dual eq.)
DD-LMS	64.3%	94.3%
LTBOMB	88.4%	99.2%

TABLE IV
DEPENDENCE ON K

K	DD-LMS percent converged to MMSE (single eq.)	LTBOMB percent converged to MMSE (single eq.)
2	64.3%	88.4%
4	85.3%	85.9%
8	90.8%	77.2%
16	92.1%	75.3%

Finally, we consider the performance of the algorithm as K is increased. We performed the same simulation as before, but increased the alphabet size to investigate the effect of K on algorithm convergence rate. The results when using a single equalizer are shown in Table IV, where we see that the DD-LMS algorithm improves as K is increased, while LTBOMB algorithm performance diminishes for large K. When operating in the parallel equalizer setup, both algorithms converged more than 99% of the time for $K \in \{4, 8, 16\}$. It is well known [10], at least for the AWGN channel, that the symbol-error-rate in BOM decreases as the alphabet size $M \triangleq 2K$ is increased. Euclidean distance arguments suggest similar behavior in ISI channels, which we have observed to be the case. Thus, for larger values of K, we can expect the decision device to make fewer errors, thereby improving the performance of decision-directed LMS. The blind algorithms do not benefit from an increase in K since they do not rely on correct decisions.

VIII. CONCLUSION

While energy efficient modulations are being given serious attention by industry for use in environments with ISI, little at-

tention has been given to equalization of such signals. We have provided the first look at adaptive equalization of BOM signals. Due to the non-i.i.d. nature of these signals, many difficulties arise in the application of classical equalization techniques. We have attempted to address this difficulties, and have proposed two novel blind algorithms for the equalization of BOM signals. We then addressed convergence issues of these algorithms, and demonstrated their performance.

In spite of the fact that our algorithm performed quite well, there are still many issues to resolve in equalization of BOM signals. Future work could investigate improvements of these algorithms to improve their convergence behavior, perhaps by exploiting oversampling or MIMO techniques. In addition, the issue of joint symbol timing acquisition and equalization needs to be fully resolved. We plan to investigate blind equalization of other energy efficient modulations, as well, such as orthogonal modulation (including PPM and FSK) and simplex modulation.

APPENDIX I
FOURTH-ORDER CUMULANT TENSOR OF BOM SOURCE

Here, we provide an expression for the fourth-order cumulant tensor in terms of $\boldsymbol{S}$. First, we consider the choice $\boldsymbol{S} = \boldsymbol{I}$. Due to the supersymmetry of the cumulant tensor, we can consider an ordering $i_1 \geq i_2 \geq i_3 \geq i_4$ without loss of generality. From (9), averaging over all possible symbols gives the cumulant tensor for $\boldsymbol{S} = \boldsymbol{I}$ as

$$[\mathcal{C}_4(\boldsymbol{x})]_{i_1,i_2,i_3,i_4} = \begin{cases} \frac{1}{K} - \frac{3}{K^2} & \text{for } i_1 = i_2 = i_3 = i_4 \\ -\frac{1}{K^2} & \text{for } i_1 = i_2 > i_3 = i_4 \\ 0 & \text{otherwise.} \end{cases} \tag{19}$$

Next, note that for any choice of $\boldsymbol{S}$, we can express the source symbol $\boldsymbol{x}[n] = \boldsymbol{S}\boldsymbol{x}'[n]$ where $\boldsymbol{x}'[n]$ is a BOM symbol for the particular choice $\boldsymbol{S} = \boldsymbol{I}$. Thus, exploiting the linearity property of cumulants and using (19), we have for all $i_1, i_2, i_3, i_4 \in \{0, \ldots, K-1\}$

$$\begin{aligned} &[\mathcal{C}_4(\boldsymbol{x})]_{i_1,i_2,i_3,i_4} \\ &\quad = [\mathcal{C}_4(\boldsymbol{S}\boldsymbol{x}')]_{i_1,i_2,i_3,i_4} \\ &\quad = \sum_{k_1,k_2,k_3,k_4} [\mathcal{C}_4(\boldsymbol{x}')]_{k_1,k_2,k_3,k_4} [\boldsymbol{S}]_{i_1,k_1} \\ &\qquad \times [\boldsymbol{S}]_{i_2,k_2} [\boldsymbol{S}]_{i_3,k_3} [\boldsymbol{S}]_{i_4,k_4} \\ &\quad = \frac{1}{K} \sum_{\ell} [\boldsymbol{S}]_{i_1,\ell} [\boldsymbol{S}]_{i_2,\ell} [\boldsymbol{S}]_{i_3,\ell} [\boldsymbol{S}]_{i_4,\ell} \\ &\qquad - \frac{1}{K^2} \delta[i_1 - i_3] \delta[i_2 - i_4] \\ &\qquad - \frac{1}{K^2} \delta[i_1 - i_2] \delta[i_3 - i_4] - \frac{1}{K^2} \delta \\ &\qquad \times [i_1 - i_4] \delta[i_2 - i_3]. \end{aligned} \tag{20}$$

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their comments and keen observations which helped improve this manuscript.

REFERENCES

[1] M. Webster, Proposal for a high speed PHY for the 2.4 GHz band IEEE P802.1 1-98/47, Jan. 1998.

[2] R. Fisher, DS-UWB physical layer submission to 802.15 task group 3a IEEE P802.1 5-04/0 1373r3, Jul. 2004.

[3] H. Zhang and T. A. Gulliver, "Biorthogonal pulse position modulation for time-hopping multiple access UWB communications," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1 154–1 162, May 2005.

[4] N. Boubaker and K. B. Letaief, "Biorthogonal pulse shape modulation for ultra-wideband wireless communications," in *Proc. IEEE Global Telecommun. Conf. (IEEE GLOBECOM'04)*, Nov. 2004, vol. 6, pp. 3493–3497.

[5] K. Takizawa and R. Kohno, "Low -complexity RAKE reception and equalization for MBOK DS-UWB systems," in *Proc. IEEE Global Telecommun. Conf. (IEEE GLOBECOM'04)*, Nov. 2004, vol. 2, pp. 1249–1253.

[6] A. J. Viterbi, "On coded phase-coherent communications," *IRE Trans. Space Electron. Telemetry*, vol. SET-7, pp. 3–14, Mar. 1961.

[7] C. Simon, P. Loubaton, C. Vignat, C. Jutten, and G. d'Urso, "Blind source separation of convolutive mixtures by maximization of fourth-order cumulants: The non i.i.d. case," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Nov. 1998, vol. 2, pp. 1584–1588.

[8] M. Castella, J.-C. Pesquet, and A. P. Petropulu, "New contrasts for blind separation of non iid sources in the convolutive case," in *Proc. Eur. Signal Process. Conf. (EUSIPCO'02)*, Sept. 2002, vol. 2, pp. 107–110.

[9] J. P. LeBlanc, I. Fijalkow, and C. R. Johnson, Jr., "CMA fractionally spaced equalizers: Stationary points and stability under IID and temporally correlated sources," *Intl. J. Adaptive Contr. Signal Process.*, vol. 12, pp. 135–155, Mar. 1998.

[10] J. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2000.

[11] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ: Prentice-Hall, 2001.

[12] O. Macchi and E. Eweda, "Convergence analysis of self-adaptive equalizers," *IEEE Trans. Inf. Theory*, vol. 30, pp. 161–176, Mar. 1984.

[13] A. Klein, C. R. Johnson, Jr., and P. Duhamel, "On blind equalization of M-ary bi-orthogonal signaling," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP'05)*, Mar. 2005.

[14] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-31, pp. 459–472, Apr. 1983.

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

[16] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. Inf. Theory*, vol. 36, pp. 312–321, Mar. 1990.

[17] C. R. Johnson, Jr., "Blind equalization using the constant modulus criterion: A review," *Proc. IEEE*, vol. 86, pp. 1927–1950, Oct. 1998.

[18] G. Foschini, "Equalizing without altering or detecting data," *Bell Syst. Tech. J.*, vol. 64, pp. 1885–1911, Oct. 1985.

[19] D. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms*. Berlin, Germany: Springer-Verlag, 1992.

[20] V. Yang and D. Jones, "A vector constant modulus algorithm for shaped constellation equalization," *IEEE Signal Process. Lett.*, vol. 5, pp. 89–91, Apr. 1998.

[21] A. Touzni, L. Tong, R. A. Casas, and C. R. Johnson, Jr., "Vector-CM stable equilibrium analysis," *IEEE Signal Process. Lett.*, vol. 7, pp. 31–33, Feb. 2000.

[22] L. De Lathauwer, "Signal processing based on multilinear algebra," Ph.D. dissertation, Katholieke Univ. Leuven, Leuven, Belgium, Sep. 1997.

[23] G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities*. London, U.K.: Cambrige Univ. Press, 1934.

[24] V. Shtrom and H. H. Fan, "New class of zero-forcing cost functions in blind equalization," *IEEE Trans. Signal Process.*, vol. 46, no. 10, pp. 2674–2683, Oct. 1998.

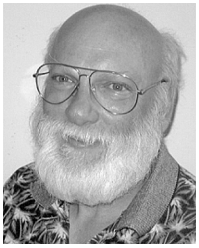
[25] P. A. Regalia, "On the equivalence between the godard and Shalvi-Weinstein schemes of blind equalization," *Signal Process. (Elsevier)*, vol. 73, pp. 185–190, Feb. 1999.

[26] A. Salch and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE J. Sel. Areas Commun.*, vol. SAC-5, pp. 128–137, Feb. 1987.

[27] Channel Modeling Sub-Committee Report Final IEEE P802.15-02/490, Feb. 2003, J. Foerster, Ed..

Andrew G. Klein received the B.S. degree in electrical engineering from Cornell University, Ithaca, NY, in 1998, the M.S. degree in electrical engineering from the University of California, Berkeley, in 2000., and the Ph.D. degree from Cornell University in 2005.

He was with several Bay Area wireless startup companies, while pursuing his degrees. He is currently a Chateaubriand Fellow conducting postdoctoral research at LSS/Supelec, Paris, France. His areas of interest are statistical signal processing and adaptive parameter estimation for wireless communication systems.



C. Richard Johnson, Jr. (S'74–M'77–F'89) received the Ph.D. degree in electrical engineering, along with the first Ph.D. minor in art history, from Stanford University, Stanford, CA, in 1977.

He joined the faculty of Cornell University, Ithaca, NY, in 1981, where he is a Stephen H. Weiss Presidential Fellow and a professor of electrical and computer engineering. His primary research interest has been blind, adaptive equalization for mitigating intersymbol interference in communication systems.

Dr. Johnson has received several teaching awards, including selection by Eta Kappa Nu as the C. Holmes MacDonald Outstanding Teacher among young professors of electrical engineering in the United States in 1983, and selection in 2004 as a Stephen H. Weiss Presidential Fellow.

Pierre Duhamel (M'87–SM'87–F'98) was born in France in 1953. He received the Eng. degree in electrical engineering from the National Institute for Applied Sciences (INSA) Rennes, France, in 1975, and the Dr.Eng. and the Doctorat és Sciences degrees from Orsay University, Orsay, France, in 1978 and 1986, respectively.

From 1975 to 1980, he was with Thomson-CSF, Paris, France, where his research interests were in circuit theory and signal processing, including digital filtering and analog fault diagnosis. In 1980, he joined the National Research Center in Telecommunications (CNET), Issy les Moulineaux, France, where his research activities were first concerned with the design of recursive charge-coupled device (CCD) filters. Later, he worked on fast algorithms for computing Fourier transforms and convolutions and applied similar techniques to adaptive filtering, spectral analysis, and wavelet transforms. From 1993 to September 2000, he was Professor with the Ecole Nationale Supérieure des Télécommunications (ENST, National School of Engineering in Telecommunications), Paris, with research activities focused on signal processing for communications, where he headed the Signal and Image Processing Department from 1997 to 2000. He is now with CNRS/LSS (Laboratoire de Signaux et Systemes), Gif-sur-Yvette, France, where he is developing studies in signal processing for communications (including equalization, iterative decoding, multicarrier systems) and signal/image processing for multimedia applications, including source coding, joint source/channel coding, watermarking, and audio processing.

Dr. Duhamel was Chairman of the DSP committee from 1996 to 1998 and a Member of the SP for Com Committee until 2001. He was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1989 to 1991, an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS, and a Guest Editor for the Special Issue on wavelets of the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He was Distiguished Lecturer, IEEE, for 1999, and was Co-general Chair of the 2001 International Workshop on Multimedia Signal Processing, Cannes, France. The paper on subspace-based methods for blind equalization, which he coauthored, received the Best Paper Award from the IEEE TRANSACTIONS ON SIGNAL PROCESSING in 1998. He was awarded the "Grand Prix France Telecom" by the French Science Academy in 2000.