LOW COMPLEXITY BLIND EQUALIZATION FOR OFDM SYSTEMS WITH GENERAL CONSTELLATIONS

Tareq Y. Al-Naffouri[*1], *Ala Dahman*[1], *Muhammad S. Sohail*[1], *Weiyu Xu*[2] *and Babak Hassibi*[3]

[1] Department of Electrical Engineering, King Fahd University of Petroleum and Minerals, Saudi Arabia
[2] Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA
[3] Department of Electrical Engineering, California Institute of Technology, Pasadena, CA

Abstract—This paper proposes a low-complexity algorithm for blind equalization of data in OFDM-based wireless systems with general constellation. The proposed algorithm is able to recover data even when the channel changes on a symbol-by-symbol basis, making it suitable for fast fading channels. The proposed algorithm does not require any statistical information of the channel and thus does not suffer from latency normally associated with blind methods. We also demonstrate how to reduce the complexity of the algorithm, which becomes especially low at high SNR. Specifically, we show that in the high SNR regime, the number of operations is of the order $O(LN)$, where L is the cyclic prefix length and N is the total number of subcarriers. Simulation results confirm the favorable performance of our algorithm.

Index Terms—OFDM, channel estimation, maximum-likelihood detection, maximum a posteriori detection and recursive least squares.

I. Introduction

Modern wireless communication systems are expected to meet an ever increasing demand for high data rates. A major hindrance for such high data rate systems is multipath fading. Orthogonal frequency division multiplexing (OFDM), owing to its robustness to multipath fading, has been incorporated in many existing standards (e.g., IEEE 802.11, IEEE 802.16, DAB, DVB, HyperLAN, ADSL etc.) and is also a candidate for future wireless standards (e.g., IEEE 802.20). All current standards use pilot symbols to obtain channel state information (needed to perform coherent data detection). This reduces the bandwidth available for data transmission, e.g., the IEEE 802.11n standard uses 4 subcarriers for pilots, that is 7.1% of the available bandwidth, of the 56 subcarriers available for transmission. Blind equalization methods are advantageous as they do not require regular training/pilots symbols, thus freeing up valuable bandwidth.

Several works exist in literature on blind channel estimation and equalization. A brief classification of these works based on a few commonly used constraints/assumptions is given in Table I (note that this list is not exhaustive). Broadly speaking, the literature on blind channel estimation can be classified into maximum-likelihood (ML) methods and non-ML methods.

The non-ML methods include approaches based on subspace techniques [1]-[10], second-order statistics [11], [12], [13], Cholesky factorization [14], iterative methods [15], virtual carriers [16] real signal characteristics [17] and linear precoding [12], [18]. Subspace-based methods [1]-[5], [7]-[10] generally have lower complexity but suffer from slow convergence as they require many OFDM symbols to get an accurate estimate of the channel autocorrelation matrix. Blind methods based on second-order statistics [11], [12], [13] also require the channel to be strictly stationary over several OFDM blocks. More often than not, this condition is not fulfilled in wireless scenarios (e.g., as in WLAN and fixed wireless applications). Methods based on Cholesky's factorization [14] and iterative techniques [15] suffer from high computational complexity.

Several ML-based blind methods have been proposed in literature [20], [19], [21]-[35], [37]. Although they incur a higher computational cost, their superior performance and faster convergence is very attractive. These characteristics make this class of algorithms suitable for block fading scenarios with short channel coherence time. Usually, suboptimal approximations are used to reduce the computational complexity of ML-based methods. Though these methods reduce the complexity of the exhaustive ML search, they still incur a significantly high computational cost. Some methods like [21], [23], [24] are sensitive to initialization parameters, while others work only for specific constellations (see Table I). A few ML-based algorithms allow the channel to change on a symbol-by-symbol basis (e.g., [26], [37]), however, these algorithms are only able to deal with constant modulus constellations.

To the best of our knowledge no blind algorithm in literature is able to deal with channels that change from one OFDM symbol to another when the data symbols are drawn from a general constellation. Contrast this with the equalization algorithm presented in this paper. The key features of the blind equalization algorithm presented in this paper are that it
1) works with an arbitrary constellation,
2) can deal with channels that change for one symbol to the next,
3) does not assume any statistical information about the channel.

In addition, we propose a low-complexity implementation of the algorithm by utilizing the special structure of partial FFT matrices and prove that the complexity becomes especially low in the high SNR regime.

The paper is organized as follows. Section II describes the system model and Section III describes the blind equalization

TABLE I
LITERATURE CLASSIFICATION

Constraint	Limited by	Not limited by
Channel constant over M symbols, $M > 1$	[1], [2], [3], [5], [6], [7], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [20], [21], [24], [27], [28]	[26], [37]
Uses pilots to resolve phase ambiguity	[1], [5], [9], [10], [11], [14], [15], [16], [18], [20], [21], [28], [25], [26], [37]	[36]
Constant modulus constellation	[2], [3], [6], [9], [12], [13], [15], [16], [20], [21], [24], [26], [27], [36], [37]	[1], [5], [7], [11], [14] [17], [19], [28]

algorithm. Section IV presents an approximate method to reduce the computational complexity of the algorithm, while Section V evaluates this complexity in the high SNR regime. Section VI presents the simulation results and Section VII gives the concluding remarks.

A. Notation

We denote scalars with small-case letters, x, vectors with small-case boldface letters, $\boldsymbol{x}$, while the individual entries of a vector $\boldsymbol{h}$ are denoted by $h(l)$. Upper case boldface letters, $\boldsymbol{X}$, represent matrices while calligraphic notation, $\boldsymbol{\mathcal{X}}$, is reserved for vectors in the frequency domain. A hat over a variable indicates an estimate of the variable, e.g., $\hat{\boldsymbol{h}}$ is an estimate of $\boldsymbol{h}$. $(.)^{\mathrm{T}}$ and $(.)^{\mathrm{H}}$ denote the transpose and Hermitian operations, while the notation $\odot$ stands for element-by-element multiplication. The discrete Fourier transform (DFT) matrix is denoted by $\boldsymbol{Q}$ and defined as $q_{l,k} = e^{-j\frac{2\pi}{N}(l-1)(k-1)}$ with $k, l = 1, 2, \cdots, N$ (N is the number of subcarriers in the OFDM symbol), while the invrse DFT (IDFT) is denoted as $\boldsymbol{Q}^{\mathrm{H}}$. The notation $\|\boldsymbol{a}\|_{\boldsymbol{B}}^2$ represents the weighted norm defined as $\|\boldsymbol{a}\|_{\boldsymbol{B}}^2 \triangleq \boldsymbol{a}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{a}$ for some vector $\boldsymbol{a}$ and matrix $\boldsymbol{B}$.

II. SYSTEM MODEL

Consider an OFDM system where all the N available subcarriers are modulated by data symbols chosen from an arbitrary constellation. The frequency-domain OFDM symbol $\boldsymbol{\mathcal{X}}$, of size $N \times 1$, undergoes an IDFT operation to produce the time-domain symbol $\boldsymbol{x}$, i.e.

$$\boldsymbol{x} = \sqrt{N}\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{\mathcal{X}}. \tag{1}$$

The transmitter then appends a length L cyclic prefix (CP) to $\boldsymbol{x}$ and transmits it over the channel. The channel $\boldsymbol{h}$, of maximum length $L+1 < N$, is assumed to be constant for the duration of a single OFDM symbol, but could change from one symbol to the next. The received signal is a convolution of the transmitted signal with the channel observed in additive white circularly symmetric Gaussian noise $\mathbf{n} \sim \mathcal{N}(0, \boldsymbol{I})$. The CP converts the linear convolution relationship to circular convolution, which, in the frequency domain, reduces to an element-by-element operation. Discarding the CP, the frequency-domain received symbol is given by

$$\boldsymbol{\mathcal{Y}} = \sqrt{\rho}\,\boldsymbol{\mathcal{H}} \odot \boldsymbol{\mathcal{X}} + \boldsymbol{\mathcal{N}}, \tag{2}$$

where ρ is the signal to noise ratio (SNR) and $\boldsymbol{\mathcal{Y}}, \boldsymbol{\mathcal{H}}, \boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{N}}$, are the N-point DFT's of $\boldsymbol{y}$, $\boldsymbol{h}$, $\boldsymbol{x}$, and additive noise $\boldsymbol{n}$

respectively, i.e.

$$\boldsymbol{\mathcal{H}} = \boldsymbol{Q}\left[\begin{array}{c} \boldsymbol{h} \\ \boldsymbol{0} \end{array}\right], \quad \boldsymbol{\mathcal{X}} = \frac{1}{\sqrt{N}}\boldsymbol{Q}\boldsymbol{x},$$
$$\boldsymbol{\mathcal{N}} = \frac{1}{\sqrt{N}}\boldsymbol{Q}\boldsymbol{n}, \text{ and } \boldsymbol{\mathcal{Y}} = \frac{1}{\sqrt{N}}\boldsymbol{Q}\boldsymbol{y}. \tag{3}$$

Note that $\boldsymbol{h}$ is zero padded before taking its N-point DFT. Let $\boldsymbol{A}^{\mathrm{H}}$ consist of first $L+1$ columns of $\boldsymbol{Q}$ (i.e., $\boldsymbol{A}$ consist of first $L+1$ rows of $\boldsymbol{Q}^{\mathrm{H}}$), then

$$\boldsymbol{\mathcal{H}} = \boldsymbol{A}^{\mathrm{H}}\boldsymbol{h} \quad \text{and} \quad \boldsymbol{h} = \boldsymbol{A}\boldsymbol{\mathcal{H}}. \tag{4}$$

This allows us to rewrite (2) as

$$\boldsymbol{\mathcal{Y}} = \sqrt{\rho}\,\mathrm{diag}(\boldsymbol{\mathcal{X}})\boldsymbol{A}^{\mathrm{H}}\boldsymbol{h} + \boldsymbol{\mathcal{N}}. \tag{5}$$

III. BLIND EQUALIZATION APPROACH

Consider the input/output equation (5), which in its element by element form reads

$$\mathcal{Y}(j) = \sqrt{\rho}\,\mathcal{X}(j)\boldsymbol{a}_j^{\mathrm{H}}\boldsymbol{h} + \mathcal{N}(j) \tag{6}$$

where $\boldsymbol{a}_j$ is the jth column of $\boldsymbol{A}$. The problem of joint ML channel estimation and data detection for OFDM channels can be cast as the following minimization problem

$$
\begin{aligned}
J_{ML} &= \min_{h, \mathcal{X} \in \Omega^N} \|\boldsymbol{\mathcal{Y}} - \sqrt{\rho}\,\mathrm{diag}(\boldsymbol{\mathcal{X}})\boldsymbol{A}^{\mathrm{H}}\boldsymbol{h}\|^2 \\
&= \min_{h, \mathcal{X} \in \Omega^N} \sum_{i=1}^{N} |\mathcal{Y}(i) - \sqrt{\rho}\,\mathcal{X}(i)\boldsymbol{a}_i^{\mathrm{H}}\boldsymbol{h}|^2 \\
&= \min_{h, \mathcal{X} \in \Omega^N} \left\{ \sum_{j=1}^{i} |\mathcal{Y}(j) - \sqrt{\rho}\,\mathcal{X}(j)\boldsymbol{a}_j^{\mathrm{H}}\boldsymbol{h}|^2 + \right. \\
&\quad \left. \sum_{j=i+1}^{N} |\mathcal{Y}(j) - \sqrt{\rho}\,\mathcal{X}(j)\boldsymbol{a}_j^{\mathrm{H}}\boldsymbol{h}|^2 \right\}
\end{aligned}
\tag{7}
$$

where Ω^N denotes the set of all possible $N-$dimensional signal vectors. Let us consider a partial data sequence $\boldsymbol{\mathcal{X}}_{(i)}$ up to the time index i, i.e.[1]

$$\boldsymbol{\mathcal{X}}_{(i)} = [\mathcal{X}(1) \ \ \mathcal{X}(2) \ \ \cdots \ \ \mathcal{X}(i)]^{\mathrm{T}}$$

and define $M_{\mathcal{X}_{(i)}}$ as the corresponding cost function, i.e.

$$M_{\mathcal{X}_{(i)}} = \min_h \|\boldsymbol{\mathcal{Y}}_{(i)} - \sqrt{\rho}\,\mathrm{diag}(\boldsymbol{\mathcal{X}}_{(i)})\boldsymbol{A}_{(i)}^{\mathrm{H}}\boldsymbol{h}\|^2, \tag{8}$$

[1]Thus, for example $\boldsymbol{\mathcal{X}}_{(2)} = [\mathcal{X}(1), \mathcal{X}(2)]^{\mathrm{T}}$ and $\boldsymbol{\mathcal{X}}_{(N)} = [\mathcal{X}(1), \cdots, \mathcal{X}(N)]^{\mathrm{T}} \triangleq \boldsymbol{\mathcal{X}}$.

where $\boldsymbol{A}_{(i)}^{\mathrm{H}}$ consists of the first i rows of $\boldsymbol{A}^{\mathrm{H}}$.

In the following, we pursue an idea for blind equalization of single-input multiple-output systems first inspired by [19]. Let R be the optimal value for the objective function (7) (we show how to determine R in Section III-B further ahead). If $M_{\mathcal{X}_{(i)}} > R$, then $\boldsymbol{\mathcal{X}}_{(i)}$ can not be the first i symbols of the ML solution $\hat{\boldsymbol{\mathcal{X}}}^{\mathrm{ML}}$ to (7). To prove this, let $\hat{\boldsymbol{\mathcal{X}}}^{\mathrm{ML}}$ and $\hat{\boldsymbol{h}}^{\mathrm{ML}}$ denote the ML estimates and suppose that our estimate $\hat{\boldsymbol{\mathcal{X}}}_{(i)}$ satisfies

$$\hat{\boldsymbol{\mathcal{X}}}_{(i)} = \hat{\boldsymbol{\mathcal{X}}}_{(i)}^{\mathrm{ML}} \tag{9}$$

i.e. the estimate $\hat{\boldsymbol{\mathcal{X}}}_{(i)}$ matches the first i elements of the ML estimate. Then, we can write

$$\begin{aligned}
R &= \min_{h, \mathcal{X} \in \Omega^N} \|\boldsymbol{\mathcal{Y}} - \sqrt{\rho}\, \mathrm{diag}(\boldsymbol{\mathcal{X}})\boldsymbol{A}^{\mathrm{H}}\boldsymbol{h}\|^2 \\
&= \|\boldsymbol{\mathcal{Y}}_{(i)} - \sqrt{\rho}\, \mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}_{(i)}^{\mathrm{ML}})\boldsymbol{A}_{(i)}^{\mathrm{H}}\hat{\boldsymbol{h}}^{\mathrm{ML}}\|^2 \\
&\quad + \sum_{j=i+1}^{N} |\mathcal{Y}(j) - \sqrt{\rho}\, \hat{\mathcal{X}}^{\mathrm{ML}}(j)\boldsymbol{a}_j^{\mathrm{H}}\hat{\boldsymbol{h}}^{\mathrm{ML}}|^2 \\
&= \|\boldsymbol{\mathcal{Y}}_{(i)} - \sqrt{\rho}\, \mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}_{(i)}^{\mathrm{H}}\hat{\boldsymbol{h}}^{\mathrm{ML}}\|^2 \\
&\quad + \sum_{j=i+1}^{N} |\mathcal{Y}(j) - \sqrt{\rho}\, \hat{\mathcal{X}}^{\mathrm{ML}}(j)\boldsymbol{a}_j^{\mathrm{H}}\hat{\boldsymbol{h}}^{\mathrm{ML}}|^2, \tag{10}
\end{aligned}$$

where the last equation follows from (9). Now, clearly

$$\begin{aligned}
&\|\boldsymbol{\mathcal{Y}}_{(i)} - \sqrt{\rho}\, \mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}_{(i)}^{\mathrm{H}}\hat{\boldsymbol{h}}^{\mathrm{ML}}\|^2 \\
&\geq \min_h \|\boldsymbol{\mathcal{Y}}_{(i)} - \sqrt{\rho}\, \mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}_{(i)}^{\mathrm{H}}\boldsymbol{h}\|^2 \tag{11} \\
&= \|\boldsymbol{\mathcal{Y}}_{(i)} - \sqrt{\rho}\, \mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}_{(i)}^{\mathrm{H}}\hat{\boldsymbol{h}}\|^2, \tag{12}
\end{aligned}$$

where $\hat{\boldsymbol{h}}$ is the argument that minimizes the RHS of (11). Then

$$\begin{aligned}
R &= \|\boldsymbol{\mathcal{Y}}_{(i)} - \sqrt{\rho}\, \mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}_{(i)}^{\mathrm{H}}\hat{\boldsymbol{h}}^{\mathrm{ML}}\|^2 \\
&\quad + \sum_{j=i+1}^{N} |\mathcal{Y}(j) - \sqrt{\rho}\, \hat{\mathcal{X}}(j)\boldsymbol{a}_j^{\mathrm{H}}\hat{\boldsymbol{h}}^{\mathrm{ML}}|^2 \\
&\geq \min_h \|\boldsymbol{\mathcal{Y}}_{(i)} - \sqrt{\rho}\, \mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}_{(i)}^{\mathrm{H}}\boldsymbol{h}\|^2 \\
&= M_{\mathcal{X}_{(i)}}. \tag{13}
\end{aligned}$$

So, for $\hat{\boldsymbol{\mathcal{X}}}_{(i)}$ to correspond to the first i symbols of the ML solution $\hat{\boldsymbol{\mathcal{X}}}_{(i)}^{\mathrm{ML}}$, we should have $M_{\hat{\mathcal{X}}_{(i)}} < R$. Note that the above represents a necessary condition only. Thus if $\hat{\boldsymbol{\mathcal{X}}}_{(i)}$ is such that $M_{\hat{\mathcal{X}}_{(i)}} < R$, then this does not necessarily mean that $\hat{\boldsymbol{\mathcal{X}}}_{(i)}$ coincides with $\hat{\boldsymbol{\mathcal{X}}}_{(i)}^{\mathrm{ML}}$.

This suggests the following method for blind equalization. At each subcarrier frequency i, make a guess of the new value of $\mathcal{X}(i)$ and use that along with previous estimated values $\hat{\mathcal{X}}(1), ..., \hat{\mathcal{X}}(i-1)$ to construct $\hat{\boldsymbol{\mathcal{X}}}_{(i)}$. Estimate $\boldsymbol{h}$ so as to minimize $M_{\hat{\mathcal{X}}_{(i)}}$ in (13) and calculate the resulting minimum value of $M_{\hat{\mathcal{X}}_{(i)}}$. If the value of $M_{\hat{\mathcal{X}}_{(i)}} < R$, then proceed to $i+1$. Otherwise, backtrack in some manner and change the guess of $\mathcal{X}(j)$ for some $j \leq i$. A problem with this approach is that for $i \leq L+1$, given any choice of $\hat{\mathcal{X}}(i)$, $\boldsymbol{h}$ can always be

chosen by least-squares to make $M_{\hat{\mathcal{X}}_{(i)}}$ in (13) equal to zero[2]. Then, we will need at least $L+1$ pilots defying the blind nature of our algorithm. Alternatively, our search tree should be at least $L+1$ deep before we can obtain a nontrivial (i.e. nonzero) value for $M_{\hat{\mathcal{X}}_{(i)}}$.

An alternative strategy would be to find $\boldsymbol{h}$ using weighted regularized least squares. Specifically, instead of minimizing the objective function J_{ML} in equation (7), we minimize the maximum a posteriori (MAP) objective function

$$J_{MAP} = \min_{h, \mathcal{X} \in \Omega^N} \left\{ \|\boldsymbol{h}\|_{R_h^{-1}}^2 + \|\boldsymbol{\mathcal{Y}} - \sqrt{\rho}\, \mathrm{diag}(\boldsymbol{\mathcal{X}})\boldsymbol{A}^{\mathrm{H}}\boldsymbol{h}\|^2 \right\} \tag{14}$$

where $\boldsymbol{R}_h$ is the autocorrelation matrix of $\boldsymbol{h}$ (in Section IV, we modify the blind algorithm to avoid the need for channel statistics). Now the objective function in (14) can be decomposed as

$$J_{MAP} = \min_{h, \mathcal{X} \in \Omega^N} \left\{ \underbrace{\|\boldsymbol{h}\|_{R_h^{-1}}^2 + \sum_{j=1}^{i} |\mathcal{Y}(j) - \sqrt{\rho}\, \mathcal{X}(j)\boldsymbol{a}_j^{\mathrm{H}}\boldsymbol{h}|^2}_{=M_{\mathcal{X}_{(i)}}} + \sum_{j=1+1}^{N} |\mathcal{Y}(j) - \sqrt{\rho}\, \mathcal{X}(j)\boldsymbol{a}_j^{\mathrm{H}}\boldsymbol{h}|^2 \right\} \tag{15}$$

Given an estimate of $\hat{\boldsymbol{\mathcal{X}}}_{(i-1)}$, the cost function reads

$$M_{\hat{\mathcal{X}}_{(i-1)}} = \min_h \left\{ \|\boldsymbol{h}\|_{R_h^{-1}}^2 + \|\boldsymbol{\mathcal{Y}}_{(i-1)} - \sqrt{\rho}\, \mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}_{(i-1)})\boldsymbol{A}_{(i-1)}^{\mathrm{H}}\boldsymbol{h}\|^2 \right\} \tag{16}$$

with the optimum value (see [38], Chapter 12, pp. 671)

$$\hat{\boldsymbol{h}} = \sqrt{\rho}\, \boldsymbol{R}_h \boldsymbol{A}_{(i-1)} \mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}_{(i-1)}^{\mathrm{H}})$$
$$[\boldsymbol{I} + \rho\, \mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}_{(i-1)})\boldsymbol{A}_{(i-1)}^{\mathrm{H}}\boldsymbol{R}_h\boldsymbol{A}_{(i-1)}\mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}_{(i-1)}^{\mathrm{H}})]^{-1}\boldsymbol{\mathcal{Y}}_{(i-1)} \tag{17}$$

and corresponding minimum cost (MMSE error)

$$\mathrm{mmse} = [\boldsymbol{R}_h^{-1} + \rho\boldsymbol{A}_{(i-1)}\mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}_{(i-1)})^{\mathrm{H}}\mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}_{(i-1)})\boldsymbol{A}_{(i-1)}^{\mathrm{H}}]^{-1} \tag{18}$$

If we have a guess of $\mathcal{X}(i)$, we can update the cost function and obtain $M_{\hat{\mathcal{X}}_{(i)}}$. In fact, the cost function $M_{\hat{\mathcal{X}}_{(i)}}$ is the same as that of $M_{\hat{\mathcal{X}}_{(i-1)}}$ with the additional observation $\mathcal{Y}(i)$ and an additional regressor $\hat{\mathcal{X}}(i)\boldsymbol{a}_i^{\mathrm{H}}$, i.e.

$$M_{\hat{\mathcal{X}}_{(i)}} = \min_h \left\{ \|\boldsymbol{h}\|_{R_h^{-1}}^2 + \left\| \begin{bmatrix} \boldsymbol{\mathcal{Y}}_{(i-1)} \\ \mathcal{Y}(i) \end{bmatrix} - \sqrt{\rho} \begin{bmatrix} \mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}_{(i-1)})\boldsymbol{A}_{(i-1)}^{\mathrm{H}} \\ \hat{\mathcal{X}}(i)\boldsymbol{a}_i^{\mathrm{H}} \end{bmatrix} \boldsymbol{h} \right\|^2 \right\} \tag{19}$$

We can thus, recursively update the value $M_{\hat{\mathcal{X}}_{(i)}}$ based on $M_{\hat{\mathcal{X}}_{(i-1)}}$ using recursive least squares (RLS) [38], i.e.

$$M_{\hat{\mathcal{X}}_{(i)}} = M_{\hat{\mathcal{X}}_{(i-1)}} + \gamma(i)|\mathcal{Y}(i) - \sqrt{\rho}\, \hat{\mathcal{X}}(i)\boldsymbol{a}_i^{\mathrm{H}}\hat{\boldsymbol{h}}_{i-1}|^2 \tag{20}$$

---

[2]Since $\boldsymbol{A}_{(i)}^{\mathrm{H}}$ is full rank for $i \leq L+1$, $\mathrm{diag}(\boldsymbol{\mathcal{X}}_{(i)})\boldsymbol{A}_{(i)}^{\mathrm{H}}$ is full rank too for each choice of $\mathrm{diag}(\boldsymbol{\mathcal{X}}_{(i)})$ and so we will always find some $\boldsymbol{h}$ that will make the objective function in (13) zero (since $\boldsymbol{h}$ has $L+1$ degrees of freedom).

$$\hat{h}_i = \hat{h}_{i-1} + g_i \left(\mathcal{Y}(i) - \sqrt{\rho}\, \hat{\mathcal{X}}(i) a_i^{\mathrm{H}} \hat{h}_{i-1} \right) \tag{21}$$

where

$$g_i = \sqrt{\rho}\, \gamma(i) \hat{\mathcal{X}}(i)^{\mathrm{H}} \boldsymbol{P}_{i-1} \boldsymbol{a}_i \tag{22}$$

$$\gamma(i) = \frac{1}{1 + \rho |\hat{\mathcal{X}}(i)|^2 \boldsymbol{a}_i^{\mathrm{H}} \boldsymbol{P}_{i-1} \boldsymbol{a}_i} \tag{23}$$

$$\boldsymbol{P}_i = \boldsymbol{P}_{i-1} - \rho\, \gamma(i) |\hat{\mathcal{X}}(i)|^2 \boldsymbol{P}_{i-1} \boldsymbol{a}_i \boldsymbol{a}_i^{\mathrm{H}} \boldsymbol{P}_{i-1} \tag{24}$$

These recursions apply for all i and are initialized by

$$M_{\hat{\mathcal{X}}_{(-1)}} = 0, \quad \boldsymbol{P}_{-1} = \boldsymbol{R}_h, \quad \text{and} \quad \hat{h}_{-1} = \boldsymbol{0}$$

Now, let R be the optimal value for the regularized objective function in (14). If the value R can be estimated, we can restrict the search of the blind MAP solution $\hat{\mathcal{X}}$ to the offsprings of those partial sequences $\hat{\mathcal{X}}_{(i)}$ that satisfy $M_{\hat{\mathcal{X}}_{(i)}} < R$. This forms the basis for our exact blind algorithm described below.

A. Exact Blind Algorithm

In this subsection, we describe the algorithm used to find the MAP solution of the system. The algorithm employs the above set of iterations $(20)-(24)$ to update the value of the cost function $M_{\hat{\mathcal{X}}_{(i)}}$ which is then compared with the optimal value R. The input parameters for the algorithm are: the received channel output $\mathcal{Y}$, the initial search radius r, the modulation constellation[3] Ω and the $1 \times N$ index vector I.

The algorithm is described as follows (the algorithm is also described by the flowchart in Figure 1)

1) (**Initialize**) Set $i = 1$, $I(i) = 1$ and set $\hat{\mathcal{X}}(i) = \Omega(I(i))$.
2) (**Compare with bound**) Compute and store the metric $M_{\hat{\mathcal{X}}_{(i)}}$. If $M_{\hat{\mathcal{X}}_{(i)}} > r$, go to 3; else, go to 4;
3) (**Backtrack**) Find the largest $1 \le j \le i$ such that $I(j) < |\Omega|$. If there exists such j, set $i = j$ and go to 5; else go to 6.
4) (**Increment subcarrier**) If $i < N$ set $i = i+1$, $I(i) = 1$, $\hat{\mathcal{X}}(i) = \Omega(I(i))$ and go to 2; else store current $\hat{\mathcal{X}}_{(N)}$, update $r = M_{\hat{\mathcal{X}}_{(N)}}$ and go to 3.
5) (**Increment constellation**) Set $I(i) = I(i) + 1$ and $\hat{\mathcal{X}}(i) = \Omega(I(i))$. Go to 2.
6) (**End/Restart**) If a full-length sequence $\hat{\mathcal{X}}_{(N)}$ has been found in Step 4, output it as the MAP solution and terminate; otherwise, double r and go to 1.

The essence of the algorithm is to eliminate any choice of the input that increments the objective function beyond the radius r. When such a case is confronted, the algorithm backtracks (Step 3 then Step 5) to the nearest subcarrier whose alphabet has not been exhausted (the nearest subcarrier will be the current subcarrier if its alphabet set is not exhausted).

The other dimension the algorithm works on is properly sizing r; if r is too small such that we are not able to backtrack, the algorithm doubles r (Step 3 then Step 6). If on the other hand r is too large that we reach the last subcarrier too fast, the

algorithm reduces r to the most recent value of the objective function. ($r = M_{\mathcal{X}_{(N)}}$) and backtracks (Step 4 then Step 3).

Remark 1: The backtracking algorithm depends heavily on calculating the cost function using (20)-(24). In the constant modulus case, the values of $\rho |\hat{\mathcal{X}}(i)|^2$ in equations (23) and (24) become constant (equal to $\rho\, \mathcal{E}_{\mathcal{X}}$) for all i, and the values of $\gamma(i)$ and $\boldsymbol{P}_i$ become

$$\gamma(i) = \frac{1}{1 + \rho\, \mathcal{E}_{\mathcal{X}} \boldsymbol{a}_i^{\mathrm{H}} \boldsymbol{P}_{i-1} \boldsymbol{a}_i} \tag{25}$$

$$\boldsymbol{P}_i = \boldsymbol{P}_{i-1} - \rho\, \mathcal{E}_{\mathcal{X}} \gamma(i) \boldsymbol{P}_{i-1} \boldsymbol{a}_i \boldsymbol{a}_i^{\mathrm{H}} \boldsymbol{P}_{i-1}, \tag{26}$$

which are independent of the transmitted signal and thus can be calculated offline.

Remark 2: The algorithm can also be used for a pilot-based standard. In this case, when the algorithm reaches a pilot holding-subcarrier, no backtracking is performed as the value of the data carrier is known perfectly. In the presence of pilots, it is wise to execute the algorithms over the pilot-holding subcarriers first and subsequently move to the data subcarriers. For equispaced comb-type pilots, (semi)-orthogonality of regressors is still guaranteed.

Remark 3: Like all blind algorithms, we use one pilot bit to resolve the sign ambiguity (see references in Table I).

B. Determination of initial radius ρ, $\boldsymbol{R}_h$ and r

Our algorithm depends on ρ, $\boldsymbol{R}_h$ and r which we need to determine. The receiver can easily estimate ρ by measuring the additive noise variance at its side. As for the channel covariance matrix $\boldsymbol{R}_h$, our simulations show that with carrier reordering we can replace $\boldsymbol{R}_h$ with identity with essentially no effect on performance. This becomes especially true in the high SNR regime. It remains to obtain an initial guess of the search radius r. To this emd, note that if $\boldsymbol{h}$ and $\mathcal{X}$ are perfectly known (with $\boldsymbol{h}$ drawn from $\mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_h)$ but is known) then

$$\xi = \|\boldsymbol{h}\|_{R_h^{-1}}^2 + \|\mathcal{Y} - \sqrt{\rho}\, \mathrm{diag}(\mathcal{X}) \boldsymbol{A}^{\mathrm{H}} \boldsymbol{h}\|^2 \tag{27}$$

is a chi-square random variable with $k = 2(N+L+1)$ degrees of freedom[4]. Thus, the search radius should be chosen such that $P(\xi > r) \le \epsilon$, where $P(\xi > r) = 1 - F(r; k)$, and where $F(r; k)$ is the cumulative distribution function of the chi-square random variable given by

$$F(r; k) = \frac{\gamma(k/2, \, r/2)}{\Gamma(k/2)}, \tag{28}$$

Here, $\gamma(k/2, \, r/2)$ is the lower incomplete gamma function defined as

$$\gamma(k/2, \, r/2) = \int_0^{r/2} t^{k/2-1} e^{-t}\, \mathrm{d}t. \tag{29}$$

So, under this initial radius, we guarantee finding the MAP solution with probability at least $1 - \epsilon$. In case a solution is not found, the algorithm doubles the value of r and starts over. This process continues until a solution is found. For example,

---

[3]Examples of the modulation constellation are Ω are 4-QAM and 16-QAM. We use $|\Omega|$ to denote the constellation size and $\Omega(k)$ for the kth constellation point. For example, in 4-QAM $|\Omega| = 4$ and $\Omega(1), \cdots, \Omega(4)$ are the four constellation points of 4-QAM. The indicator $I(i)$ refers to the last constellation point visited by our search algorithm at the ith subcarrier.

[4]The first term on the right hand side has $2(L+1)$ degrees of freedom as $\boldsymbol{h}$ is Gaussian distributed while the second term has $2N$ degrees of freedom as $\mathcal{Y} - \sqrt{\rho}\, \mathrm{diag}(\mathcal{X}) \boldsymbol{A}^{\mathrm{H}} \boldsymbol{h}$ is just Gaussian noise.
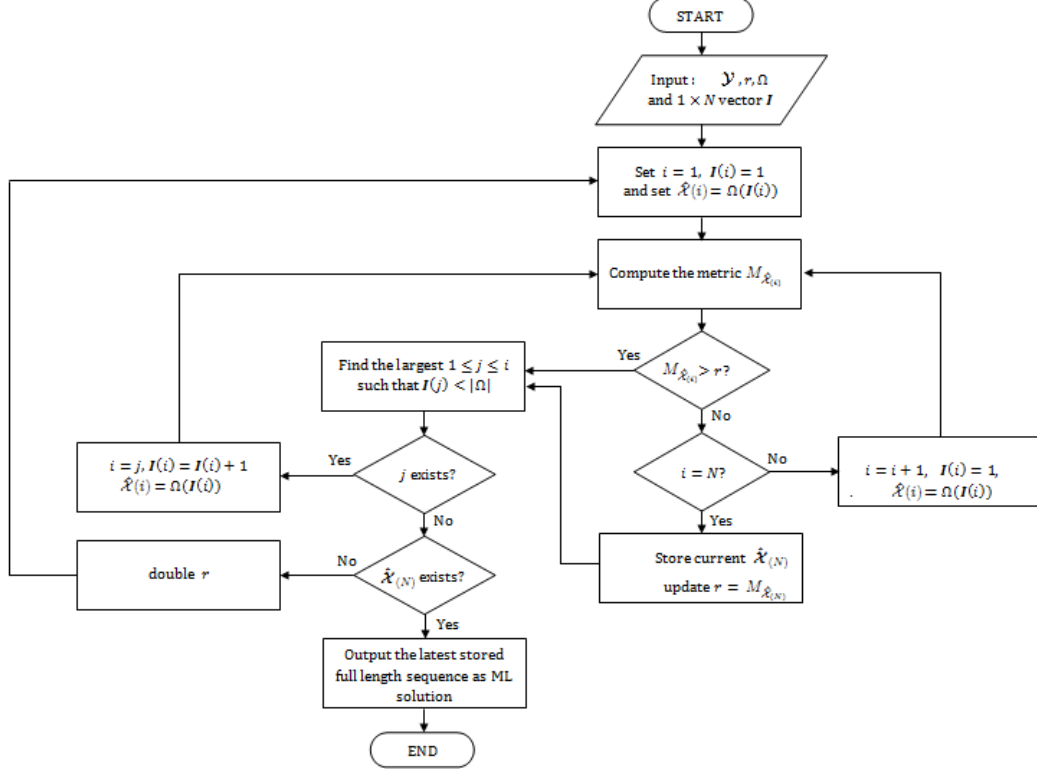
Fig. 1. Flowchart of the blind algorithm.

when $N = 64, L = 15$ and $\epsilon = 0.01$, the value of our radius should be set to 204.

IV. AN APPROXIMATE BLIND EQUALIZATION METHOD

There are two main sources that contribute to the complexity of the exact blind algorithm of Section III:

1) *Calculating P_i:* the second step of the blind algorithm requires updating the metric $M_{\hat{\mathcal{X}}_{(N)}}$. This metric depends heavily on operations involving the $(L+1) \times (L+1)$ matrix P_i which are the most computationally expansive (see Table II which estimates the computational complexity of the RLS).

2) *Backtracking:* When the condition $M_{\hat{\mathcal{X}}_{(i)}} \leq r$ is not satisfied, we need to backtrack and pursue another branch of the search tree. This represents a major source of complexity.

In the following, we show how we can avoid calculating P_i all together. We postpone the issue of backtracking to Section V.

A. Avoiding P_i

Note that in the RLS recursions (20)−(24), P_i always appears multiplied by a_i. Let's see how this changes if we set $P_{-1} = I$ and assume that the a_i's are orthogonal or, in

particular, if we assume that $a_i^{\mathrm{H}} a_{i+1} = a_i^{\mathrm{H}} a_{i+2} = 0$. With these assumptions note that

$$\gamma(0) = \frac{1}{1 + \rho \, |\hat{\mathcal{X}}(0)|^2 a_0^{\mathrm{H}} P_{-1} a_0} = \frac{1}{1 + \rho \, |\hat{\mathcal{X}}(0)|^2 (L+1)} \tag{30}$$

i.e., $\gamma(0)$ is independent of P_{-1}. Also note that

$$
\begin{aligned}
P_0 a_1 &= P_{-1} a_1 - \rho \, \gamma(0) |\hat{\mathcal{X}}(0)|^2 P_{-1} a_0 a_0^{\mathrm{H}} P_{-1} a_1 \\
&= a_1 - \rho \, \gamma(0) |\hat{\mathcal{X}}(0)|^2 a_0 a_0^{\mathrm{H}} a_1 \\
&= a_1. \tag{31}
\end{aligned}
$$

For a similar reason

$$P_0 a_2 = a_2. \tag{32}$$

From (31), it is also easy to conclude that

$$\gamma(1) = \frac{1}{1 + \rho \, |\hat{\mathcal{X}}(1)|^2 (L+1)} \tag{33}$$

i.e., $\gamma(1)$ is independent of P_0. Also, from (31) and (32) it follows that $P_i a_{i+1} = a_{i+1}$ and $P_i a_{i+2} = a_{i+2}$. We now investigate what happens to P_{i+1}.

$$
\begin{aligned}
&P_{i+1} a_{i+2} \\
&= P_i a_{i+2} - \rho \, \gamma(i+1) |\hat{\mathcal{X}}(i+1)|^2 P_i a_{i+1} a_{i+1}^{\mathrm{H}} P_i a_{i+2} \\
&= a_{i+2} - \rho \, \gamma(i+1) |\hat{\mathcal{X}}(i+1)|^2 a_{i+1} a_{i+1}^{\mathrm{H}} a_{i+2} \\
&= a_{i+2}. \tag{34}
\end{aligned}
$$

TABLE II
ESTIMATED COMPUTATIONAL COST PER ITERATION OF THE RLS
ALGORITHM

Term	$\times$	$+$	$\div$
$\sqrt{\rho}\,\hat{\mathcal{X}}(i)\boldsymbol{a}_i^{\mathrm{H}}\hat{\boldsymbol{h}}_{i-1}$	$2L+2$	L	
$\|\mathcal{Y}(i)-\sqrt{\rho}\,\hat{\mathcal{X}}(i)\boldsymbol{a}_i^{\mathrm{H}}\hat{\boldsymbol{h}}_{i-1}\|^2$	1	1	
$\rho\,\gamma(i)$	1		1
$M_{\hat{\mathcal{X}}_{(i)}}$	1	1	
$\hat{\boldsymbol{h}}_i$	$L+2$	$L+1$	1
$\boldsymbol{P}_{i-1}\boldsymbol{a}_i$	L^2+2L+1	L^2+L	
$\boldsymbol{g}_i$	$L+3$		
$\boldsymbol{a}_i^{\mathrm{H}}\boldsymbol{P}_{i-1}\boldsymbol{a}_i$	$L+1$	L	
$\gamma(i)$	3	1	1
$\boldsymbol{a}_i^{\mathrm{H}}\boldsymbol{P}_{i-1}$	L^2+2L+1	L^2+L	
$\boldsymbol{P}_i$	L^2+2L+2	L^2+2L+1	
Total per iteration	$3L^2+11L+17$	$2L^2+5L+4$	3

Similarly,

$$\boldsymbol{P}_{i+1}\boldsymbol{a}_{i+3} = \boldsymbol{a}_{i+3}. \tag{35}$$

So, by induction we see that each occurrence of $\boldsymbol{P}_i\boldsymbol{a}_i$ in the recursion set (20)-(23) can be replaced with $\boldsymbol{a}_i$. This allows us to discard (24), i.e.,

$$M_{\hat{\mathcal{X}}_{(i)}} = M_{\hat{\mathcal{X}}_{(i-1)}} + \gamma(i)|\mathcal{Y}(i)-\sqrt{\rho}\,\hat{\mathcal{X}}(i)\boldsymbol{a}_i^{\mathrm{H}}\hat{\boldsymbol{h}}_{i-1}|^2 \tag{36}$$

$$\hat{\boldsymbol{h}}_i = \hat{\boldsymbol{h}}_{i-1} + \boldsymbol{g}_i\left(\mathcal{Y}(i)-\sqrt{\rho}\,\hat{\mathcal{X}}(i)\boldsymbol{a}_i^{\mathrm{H}}\hat{\boldsymbol{h}}_{i-1}\right), \tag{37}$$

where

$$\boldsymbol{g}_i = \sqrt{\rho}\,\gamma(i)\hat{\mathcal{X}}(i)^{\mathrm{H}}\boldsymbol{a}_i \tag{38}$$

$$\gamma(i) = \frac{1}{1+\rho\,|\hat{\mathcal{X}}(i)|^2(L+1)}. \tag{39}$$

Thus, the approximate blind RLS algorithm is effectively running at LMS complexity. Table II summarizes the computational complexity incurred in the RLS calculation.

B. Avoiding $\boldsymbol{P}_i$ with Carrier Reordering

The reduction in complexity above is based on two assumptions. The first assumption is to set $\boldsymbol{P}_{-1} = \boldsymbol{I}$ (instead of $\boldsymbol{R}_h$) and the second is to assume that the consecutive $\boldsymbol{a}_i$'s are orthogonal. Note that the $\boldsymbol{a}_i$'s are columns of $\boldsymbol{A}$, i.e. they are partial FFT vectors. As such, strictly speaking, they are not orthogonal. Notice, however, that for $i \neq i'$,

$$\boldsymbol{a}_i^{\mathrm{H}}\boldsymbol{a}_{i'} = \sum_{k=0}^{L} e^{(j\frac{2\pi}{N}(i-i')k)}, \tag{40}$$

which after straightforward manipulation can be shown to be

$$|\boldsymbol{a}_i^{\mathrm{H}}\boldsymbol{a}_{i'}| = \begin{cases} L+1, & (i=i') \\ \frac{1}{L+1}\left|\frac{sin(\pi(i-i')\frac{L+1}{N})}{sin(\pi(i-i')\frac{1}{N})}\right|, & (i \neq i') \end{cases} \tag{41}$$

This is a function of $(i-i')$ mod N. Thus, without loss of generality, we can set $i' = 1$ and plot this autocorrelation with respect to i. The autocorrelation decays with i as shown in Figure 2. We can use this observation in implementing our blind RLS algorithm. Specifically, note that the whole OFDM data is available to us and so we can visit the data subcarriers in any order we wish. The discussion above shows that the data subcarriers should be visited in the order i, $i+\Delta$, $i+2\Delta,\ldots$
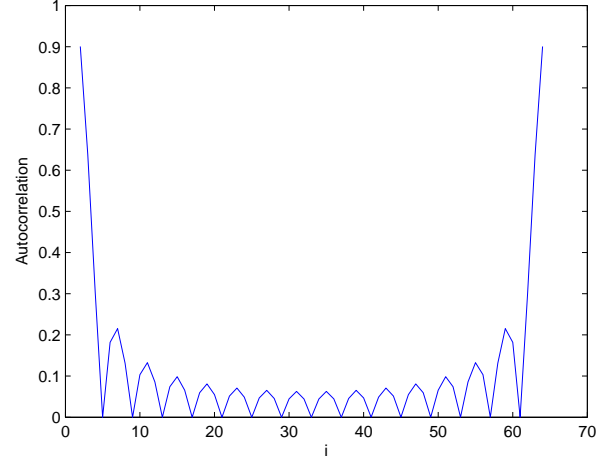

Fig. 2. Autocorrelation vs i for $N = 64$ and $L = 15$

TABLE III
ESTIMATED COMPUTATIONAL COST PER ITERATION OF THE RLS
ALGORITHM WITH CARRIER REORDERING

Term	$\times$	$+$	$\div$
$\sqrt{\rho}\,\hat{\mathcal{X}}(i)\boldsymbol{a}_i^{\mathrm{H}}\hat{\boldsymbol{h}}_{i-1}$	$2L+2$	L	
$\|\mathcal{Y}(i)-\sqrt{\rho}\,\hat{\mathcal{X}}(i)\boldsymbol{a}_i^{\mathrm{H}}\hat{\boldsymbol{h}}_{i-1}\|^2$	1	1	
$\rho\,\gamma(i)$	1		1
$M_{\hat{\mathcal{X}}_{(i)}}$	1	1	
$\hat{\boldsymbol{h}}_i$	$L+2$	$L+1$	1
$\gamma(i)$	3	1	1
Total per iteration	$4L+13$	$2L+4$	3

where Δ should be chosen as large as possible to make $\boldsymbol{a}_i$, $\boldsymbol{a}_{i+\Delta}$, $\boldsymbol{a}_{i+2\Delta},\ldots$ as orthogonal as possible, but small enough to avoid revisiting (or looping back to) a neighborhood too early. We found the choice $\Delta = \frac{N}{L+1}$ to be a good compromise. From Figure 2, which plots (41) for $N = 64$ and $L = 15$, columns $1, 5, 9, 13, 17, 21, \cdots, 61$ are orthogonal to each other and so are the columns $2, 6, 10, 14, 18, \cdots, 62$. So, if the vectors are visited in the following order $1, 5, 9, 13, 17, 21, \cdots, 61, 2, 6, 10, 14, 18, \cdots, 62, \cdots$, then we have a consecutive set of vectors that are orthogonal. The only exception is in going from column 61 to 2. These two columns are not really orthogonal but are nearly orthogonal (the correlation of columns 1 and 61 is zero, so the correlation of 61 with 2 should be very small since the correlation function is continuous as shown in Figure 2). In general, we chose $\Delta = \frac{N}{L+1}$ and visit the columns in the order $i+\Delta, i+2\Delta, \cdots, i+L\Delta, i = 1, \cdots, \Delta-1$.

Our simulation results show that the BER we get with exact calculation of $\boldsymbol{P}_i$ and that obtained when we set $\boldsymbol{P}_{-1} = \boldsymbol{I}$ with subcarrier reordering are almost the same. Table III gives the computational complexity incurred in the RLS calculation when subcarrier reordering is used (i.e., free from $\boldsymbol{P}_i$ calculation).

Note that with subcarrier reordering, the new version of the RLS runs without the need to use the power delay profile statistics, which relieves us from the need to provide this information.

V. Computational Complexity in the High SNR Regime

In the section, we study the other source of complexity (backtracking) and show that there is almost no backtracking[5] in the high SNR regime. To this end, consider the behavior of the algorithm when processing the ith subcarrier. There are $|\Omega|$ different alphabet possibilities to choose from at this subcarrier and a similar number of possibilities at the preceding $i-1$ subcarriers, creating a total of $|\Omega|^i - 1$ incorrect sequences $\bar{\mathcal{X}}_{(i)}$ and one correct sequence $\hat{\mathcal{X}}_{(i)}$. The best case scenario is to have only one sequence that satisfies $M_{\bar{\mathcal{X}}_{(i)}} \leq r$ in which case there would be only one node to visit. The worst case is having to visit the remaining $|\Omega|^i - 1$ wrong nodes before reaching the true sequence (visiting of nodes will happen through backtracking); this latter case is equivalent to the exhaustive search scenario (i.e., all possible sequences satisfy $M_{\bar{\mathcal{X}}_{(i)}} \leq r$). Thus, if we let C_i denote the expected number of nodes visited at the ith subcarrier, then from above we can write

$$C_i \leq 1 + (|\Omega|^i - 1)P_i \qquad (42)$$

where P_i is the maximum probability that an erroneous sequence of symbols $\bar{\mathcal{X}}_{(i)} \neq \hat{\mathcal{X}}_{(i)}$ has a cost less than r. We will show that this probability becomes negligibly small at high SNR values. Recall that

$$\mathcal{Y}_{(i)} = \sqrt{\rho}\,\mathrm{diag}(\hat{\mathcal{X}}_{(i)})A_{(i)}^{\mathrm{H}}h + \mathcal{N}_{(i)} \qquad (43)$$

where $\mathcal{N}_{(i)}$ denotes the first i symbols of $\mathcal{N}$. Note the (43) can be written as

$$\mathcal{Y}_{(i)} = \begin{bmatrix} \sqrt{\rho}\,\mathrm{diag}(\hat{\mathcal{X}}_{(i)})A_{(i)}^{\mathrm{H}} & I \end{bmatrix} \begin{bmatrix} h \\ \mathcal{N}_{(i)} \end{bmatrix} \qquad (44)$$

We first prove our claim for the least squares (LS) cost and then show how the MAP cost reduces to LS cost for high SNR.

A. LS cost

Suppose we have an erroneous sequence of symbols $\bar{\mathcal{X}}_{(i)} \neq \hat{\mathcal{X}}_{(i)}$. The LS estimate of h is found by minimizing the objective function

$$J_{LS} = \min_{h, \mathcal{X} \in \Omega^N} \left\{ \|\mathcal{Y}_{(i)} - \sqrt{\rho}\,\mathrm{diag}(\mathcal{X}_{(i)})A_{(i)}^{\mathrm{H}}h\|^2 \right\} \qquad (45)$$

and the solution of h is (see [38], Chapter 12, pp. 664)

$$\hat{h} = [A_{(i)}\mathrm{diag}(\hat{\mathcal{X}}_{(i)}^{\mathrm{H}})\mathrm{diag}(\hat{\mathcal{X}}_{(i)})A_{(i)}^{\mathrm{H}}]^{-1}\sqrt{\rho}\,A_{(i)}\mathrm{diag}(\hat{\mathcal{X}}_{(i)}^{\mathrm{H}})\mathcal{Y}_{(i)}. \qquad (46)$$

[5]The term "backtracking" refers to the case when the algorithm is currently at subcarrier i and it has to change the estimate of the data symbol at some subcarrier $j < i$. On the other hand, sweeping the constellation points at subcarrier to find the first one that satisfies $M_{\mathcal{X}_{(i)}} \leq r$ is not considered backtracking.

The cost associated with the LS solution is given by (see [38], Chapter 11, pp. 663)

$$
\begin{aligned}
M_{\bar{\mathcal{X}}_{(i)}} &= \mathcal{Y}_{(i)}^{\mathrm{H}}\Big(I - \sqrt{\rho}\,\mathrm{diag}(\bar{\mathcal{X}}_{(i)})A_{(i)}^{\mathrm{H}}\big[\sqrt{\rho}\,A_{(i)}\mathrm{diag}(\bar{\mathcal{X}}_{(i)})^{\mathrm{H}} \\
&\quad \sqrt{\rho}\,\mathrm{diag}(\bar{\mathcal{X}}_{(i)})A_{(i)}^{\mathrm{H}}\big]^{-1}\sqrt{\rho}A_{(i)}\mathrm{diag}(\bar{\mathcal{X}}_{(i)}^{\mathrm{H}})\Big)\mathcal{Y}_{(i)} \\
&= \mathcal{Y}_{(i)}^{\mathrm{H}}\Big(I - \rho\,\mathrm{diag}(\bar{\mathcal{X}}_{(i)})A_{(i)}^{\mathrm{H}}\big(\rho A_{(i)}|\mathrm{diag}(\bar{\mathcal{X}}_{(i)})|^2 A_{(i)}^{\mathrm{H}}\big)^{-1} \\
&\quad A_{(i)}\mathrm{diag}(\bar{\mathcal{X}}_{(i)}^{\mathrm{H}})\Big)\mathcal{Y}_{(i)} \\
&= \mathcal{Y}_{(i)}^{\mathrm{H}}\Big(I - \frac{\rho}{\rho}D\Big)\mathcal{Y}_{(i)} \\
M_{\bar{\mathcal{X}}_{(i)}} &= \mathcal{Y}_{(i)}^{\mathrm{H}}\Big(I - D\Big)\mathcal{Y}_{(i)}
\end{aligned} \qquad (47)
$$

where

$$D = \mathrm{diag}(\bar{\mathcal{X}}_{(i)})A_{(i)}^{\mathrm{H}}\Big(A_{(i)}|\mathrm{diag}(\bar{\mathcal{X}}_{(i)})|^2 A_{(i)}^{\mathrm{H}}\Big)^{-1}A_{(i)}\mathrm{diag}(\bar{\mathcal{X}}_{(i)}^{\mathrm{H}}). \qquad (48)$$

So the probability that the sequence $\bar{\mathcal{X}}_{(i)}$ satisfies $M_{\bar{\mathcal{X}}_{(i)}} \leq r$ reads

$$
\begin{aligned}
P_i &= \Pr(M_{\bar{\mathcal{X}}_{(i)}} \leq r) \\
P_i &= \Pr\Big(\mathcal{Y}_{(i)}^{\mathrm{H}}\big(I - D\big)\mathcal{Y}_{(i)} \leq r\Big)
\end{aligned} \qquad (49)
$$

In the strict sense of the word, backtracking means visiting Step 3 in our algorithm. Substituting (44) in (49) yields

$$P_i = \Pr\left(\left(\begin{bmatrix} h \\ \mathcal{N}_{(i)} \end{bmatrix}^{\mathrm{H}} G_{(i)} \begin{bmatrix} h \\ \mathcal{N}_{(i)} \end{bmatrix}\right) \leq r\right) \qquad (50)$$

where

$$G_{(i)} = \begin{bmatrix} \sqrt{\rho}\,A_{(i)}\mathrm{diag}(\hat{\mathcal{X}}_{(i)}^{\mathrm{H}}) \\ I \end{bmatrix}[I - D]\begin{bmatrix} \sqrt{\rho}\,\mathrm{diag}(\hat{\mathcal{X}}_{(i)})A_{(i)}^{\mathrm{H}} & I \end{bmatrix}. \qquad (51)$$

Let $B = \mathrm{diag}(\hat{\mathcal{X}}_{(i)})A_{(i)}^{\mathrm{H}}$, then $G_{(i)}$ can be written as

$$G_{(i)} = \begin{bmatrix} \rho\,B^{\mathrm{H}}[I - D]B & B^{\mathrm{H}}[I - D]I \\ I[I - D]B & I[I - D]I \end{bmatrix} \qquad (52)$$

which in compact form can be expressed as

$$G_{(i)} = \begin{bmatrix} \rho E & E_2 \\ E_2^{\mathrm{H}} & E_3 \end{bmatrix}. \qquad (53)$$

Using the Chernoff bound the right hand side of (50) can be bounded in the following way

$$P_i \leq e^{\mu r} E\left[\exp\left(-\mu \begin{bmatrix} h \\ \mathcal{N}_{(i)} \end{bmatrix}^{\mathrm{H}} G_{(i)} \begin{bmatrix} h \\ \mathcal{N}_{(i)} \end{bmatrix}\right)\right]. \qquad (54)$$

Noting that

$$\begin{bmatrix} h \\ \mathcal{N}_{(i)} \end{bmatrix} \sim \mathcal{N}(0, \Sigma_{(i)}) \qquad (55)$$

with

$$\Sigma_{(i)} = \begin{bmatrix} R_h & 0 \\ 0 & I_i \end{bmatrix}, \qquad (56)$$

we can solve the expression in (54) as

$$P_i \leq \frac{\int \exp\left(-\mu \begin{bmatrix} \boldsymbol{h} \\ \boldsymbol{\mathcal{N}}_{(i)} \end{bmatrix}^{\mathrm{H}} \boldsymbol{G}_{(i)} \begin{bmatrix} \boldsymbol{h} \\ \boldsymbol{\mathcal{N}}_{(i)} \end{bmatrix}\right)}{e^{-\mu r}\pi^{(L+i+1)}}$$
$$\exp\left(-\begin{bmatrix} \boldsymbol{h} \\ \boldsymbol{\mathcal{N}}_{(i)} \end{bmatrix}^{\mathrm{H}} \boldsymbol{\Sigma}_{(i)} \begin{bmatrix} \boldsymbol{h} \\ \boldsymbol{\mathcal{N}}_{(i)} \end{bmatrix}\right) d\boldsymbol{h} d\boldsymbol{\mathcal{N}}_{(i)}$$
$$= \frac{\int \exp\left(-\begin{bmatrix} \boldsymbol{h} \\ \boldsymbol{\mathcal{N}}_{(i)} \end{bmatrix}^{\mathrm{H}} (\boldsymbol{\Sigma}_{(i)} + \mu\boldsymbol{G}_{(i)}) \begin{bmatrix} \boldsymbol{h} \\ \boldsymbol{\mathcal{N}}_{(i)} \end{bmatrix}\right) d\boldsymbol{h} d\boldsymbol{\mathcal{N}}_{(i)}}{e^{-\mu r}\pi^{(L+i+1)}}$$
$$= \frac{\int \exp\left(-\left\|\begin{bmatrix} \boldsymbol{h} \\ \boldsymbol{\mathcal{N}}_{(i)} \end{bmatrix}\right\|^2_{(\boldsymbol{\Sigma}_{(i)} + \mu\boldsymbol{G}_{(i)})}\right) d\boldsymbol{h} d\boldsymbol{\mathcal{N}}_{(i)}}{e^{-\mu r}\pi^{(L+i+1)}}. \quad (57)$$

Note that the numerator in (57) is a multi-variate complex Gaussian integral. Recall that an n-dimensional complex Gaussian integral has the solution (see [19])

$$\int \exp\left(-\|\mathbf{x}\|^2_{\boldsymbol{W}}\right) d\mathbf{x} = \frac{\pi^n}{\det(\boldsymbol{W})}. \quad (58)$$

This allows us to simplify (57) as

$$P_i \leq \frac{e^{\mu r}}{\det(\boldsymbol{\Sigma}_{(i)} + \mu\boldsymbol{G}_{(i)})}. \quad (59)$$

Next, we show that the probability $P_i \to 0$ as $\rho \to \infty$. To show this, we just need to show that the largest eigenvalue of the term in the denominator goes to infinity as $\rho \to \infty$.

Lemma 1: Let $\boldsymbol{E} = \boldsymbol{A}_{(i)}\mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}^{\mathrm{H}}_{(i)})[\boldsymbol{I}-\boldsymbol{D}]\mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}^{\mathrm{H}}_{(i)}$ be a $(L+1) \times (L+1)$ matrix, then for any sequence $\hat{\boldsymbol{\mathcal{X}}}_{(i)}$, $\boldsymbol{E}$ has a positive maximum eigenvalue, $\lambda_{\max}$ and a corresponding unit-norm eigenvector $\mathbf{v}$ of size $(L+1) \times 1$.

Proof: Recall that

$$\boldsymbol{D} = \mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}^{\mathrm{H}}_{(i)} \left(\boldsymbol{A}_{(i)}\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}^{\mathrm{H}}_{(i)})\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}^{\mathrm{H}}_{(i)}\right)^{-1}$$
$$\boldsymbol{A}_{(i)}\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}^{\mathrm{H}}_{(i)}) \quad (60)$$

and let $\boldsymbol{F} = \mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}^{\mathrm{H}}_{(i)}$, then we can write the above equation as

$$\boldsymbol{D} = \boldsymbol{F}\left(\boldsymbol{F}^{\mathrm{H}}\boldsymbol{F}\right)^{-1}\boldsymbol{F}^{\mathrm{H}} = \boldsymbol{F}\boldsymbol{F}^{\dagger} \quad (61)$$

where $\boldsymbol{F}^{\dagger} = \left(\boldsymbol{F}^{\mathrm{H}}\boldsymbol{F}\right)^{-1}\boldsymbol{F}^{\mathrm{H}}$ is the Moore-Penrose pseudo-inverse[6] (see [41], Chapter 5, pp. 422). Therefore, $\boldsymbol{D}$ is an idempotent matrix with eigenvalues equal to either 0 or 1 [40] and hence, $[\boldsymbol{I}-\boldsymbol{D}]$ is also a positive semi-definite idempotent matrix. Note also that the matrix $\boldsymbol{E}$ in (53) can be written as

$$\boldsymbol{E} = \boldsymbol{A}_{(i)}\mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}^{\mathrm{H}}_{(i)})[\boldsymbol{I}-\boldsymbol{D}]\mathrm{diag}(\hat{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}^{\mathrm{H}}_{(i)}$$
$$= \boldsymbol{B}^{\mathrm{H}}[\boldsymbol{I}-\boldsymbol{D}]\boldsymbol{B} \quad (62)$$

and

$$\mathbf{z}^{\mathrm{H}}\boldsymbol{E}\mathbf{z} = \mathbf{z}^{\mathrm{H}}\boldsymbol{B}^{\mathrm{H}}[\boldsymbol{I}-\boldsymbol{D}]\boldsymbol{B}\mathbf{z} = (\boldsymbol{B}\mathbf{z})^{\mathrm{H}}[\boldsymbol{I}-\boldsymbol{D}](\boldsymbol{B}\mathbf{z}) \geq \mathbf{0} \quad (63)$$

and so $\boldsymbol{E}$ is Hermitian and positive semi-definite.

---

[6]the columns of $\boldsymbol{F}$ are linearly independent.

Let $\boldsymbol{U} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_{L+1}]$ be a $(L+1) \times (L+1)$ unitary matrix where $\mathbf{u}_i$ is the ith eigenvector. then, $\boldsymbol{E} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{H}}$ where $\boldsymbol{\Lambda}$ is a diagonal matrix containing ordered eigenvalues of $\boldsymbol{E}$ such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{L+1}$. Let $\mathbf{z} = \boldsymbol{U}^{\mathrm{H}}\mathbf{v}$, then the maximum eigenvalue of $\boldsymbol{E}$ is given as

$$\max_{\|\mathbf{v}\|_2=1} \mathbf{v}^{\mathrm{H}}\boldsymbol{E}\mathbf{v} = \max_{\|\mathbf{z}\|_2=1} \mathbf{z}^{\mathrm{H}}\boldsymbol{\Lambda}\mathbf{z} \quad (64)$$
$$= \max_{\|\mathbf{z}\|_2=1} \sum_{i=1}^{L+1} \lambda_i|z_i|^2 \quad (65)$$
$$\leq \max_{\|\mathbf{z}\|_2=1} \lambda_1 \sum_{i=1}^{L+1} |z_i|^2 \quad (66)$$
$$\leq \lambda_1 = \lambda_{\max} \quad (67)$$

The equality is attained when $\mathbf{v}$ is the eigenvector of $\lambda_{\max}$. ∎

Lemma 2: Given that $\boldsymbol{E}$ has a positive maximum eigenvalue $\lambda_{\max}$ with corresponding unit-norm vector $\mathbf{v}$ of size $(L+1) \times 1$, then the maximum eigenvalue of $\boldsymbol{G}_{(i)}$ in (52) is lower bounded by $\mathbf{w}^{\mathrm{H}}\boldsymbol{G}_{(i)}\mathbf{w} = \rho\,\lambda_{\max}$ where

$$\mathbf{w} = \begin{bmatrix} \mathbf{v}_{(L+1)\times 1} \\ \boldsymbol{0}_{i\times 1} \end{bmatrix} \quad (68)$$

Proof: From Lemma 1, the largest eigenvalue of $\boldsymbol{E}$ is $\lambda_{\max}$. It follows that the largest eigenvalue of $\rho\boldsymbol{E}$ is $\rho\lambda_{\max}$. Let $\lambda'_{\max}$ be the largest eigenvalue of $\boldsymbol{G}_{(i)}$. From (53), we can see that $\rho\boldsymbol{E}$ is a principal sub-matrix of $\boldsymbol{G}_{(i)}$ (see [41], Chapter 7, pp. 494) and thus

$$\lambda'_{\max} \geq \rho\lambda_{\max} \quad (69)$$

i.e., the largest eigenvalue of the principal sub-matrix $\rho\boldsymbol{E}$ is smaller than or equal to the largest eigenvalue of $\boldsymbol{G}_{(i)}$ (see [41], Chapter 7, pp. 551-552). Thus $\rho\lambda_{\max}$ is a lower bound on the largest eigenvalue of $\boldsymbol{G}_{(i)}$. ∎

Note that $\boldsymbol{\Sigma}_i$ is positive definite as it is a covariance matrix, hence it will have positive eigenvalues. From Lemma 2, the maximum eigenvalue of $\boldsymbol{G}_{(i)}$, $\lambda'_{\max} \to \infty$ as $\rho \to \infty$. Thus the denominator in (59) grows to infinity in the limit $\rho \to \infty$ and

$$\lim_{\rho\to\infty} P_i \to 0 \quad (70)$$

From (42) and (70), we have

$$\lim_{\rho\to\infty} C_i \leq 1 + (|\Omega|^i - 1)\lim_{\rho\to\infty} P_i \quad (71)$$
$$\lim_{\rho\to\infty} C_i \leq 1 \quad (72)$$

B. MAP cost

The cost associated with the MAP solution of an erroneous sequence of symbols $\bar{\boldsymbol{\mathcal{X}}}_{(i)} \neq \hat{\boldsymbol{\mathcal{X}}}_{(i)}$ is given as (see [38], Chapter 11, pp. 672)

$$M_{\bar{\boldsymbol{\mathcal{X}}}_{(i)}} = \boldsymbol{\mathcal{Y}}^{\mathrm{H}}_{(i)}\left(\boldsymbol{I} + \rho\,\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}^{\mathrm{H}}_{(i)}\boldsymbol{R}_h\boldsymbol{A}_{(i)}\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}^{\mathrm{H}}_{(i)})\right)^{-1}\boldsymbol{\mathcal{Y}}_{(i)} \quad (73)$$

Mathematically,

$$P_i = \Pr(M_{\bar{\mathcal{X}}_{(i)}} \leq r)$$

$$P_i =$$
$$\Pr\left(\boldsymbol{\mathcal{Y}}_{(i)}^{\mathrm{H}}\Big(\boldsymbol{I} + \rho\,\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}_{(i)}^{\mathrm{H}}\boldsymbol{R}_h\boldsymbol{A}_{(i)}\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)}^{\mathrm{H}})\Big)^{-1}\boldsymbol{\mathcal{Y}}_{(i)} \leq r\right).$$
(74)

By matrix inversion lemma

$$\Big(\boldsymbol{I} + \sqrt{\rho}\,\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}_{(i)}^{\mathrm{H}}\boldsymbol{R}_h\boldsymbol{A}_{(i)}\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)}^{\mathrm{H}})\Big)^{-1}$$
$$= \boldsymbol{I} - \rho\,\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}_{(i)}^{\mathrm{H}}\Big[\boldsymbol{R}_h^{-1} +$$
$$\rho\,\boldsymbol{A}_{(i)}\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)}^{\mathrm{H}})\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}_{(i)}^{\mathrm{H}}\Big]^{-1}\boldsymbol{A}_{(i)}\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)}^{\mathrm{H}})$$
(75)
$$= \boldsymbol{I} - \mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}_{(i)}^{\mathrm{H}}\Big[\frac{1}{\rho}\,\boldsymbol{R}_h^{-1} +$$
$$\boldsymbol{A}_{(i)}\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)}^{\mathrm{H}})\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}_{(i)}^{\mathrm{H}}\Big]^{-1}\boldsymbol{A}_{(i)}\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)}^{\mathrm{H}})$$
$$= \boldsymbol{I} - \boldsymbol{D}$$
(76)

where

$$\boldsymbol{D} =$$
$$\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}_{(i)}^{\mathrm{H}}\Big[\frac{1}{\rho}\,\boldsymbol{R}_h^{-1} + \boldsymbol{A}_{(i)}\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)}^{\mathrm{H}})\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)})\boldsymbol{A}_{(i)}^{\mathrm{H}}\Big]^{-1}$$
$$\boldsymbol{A}_{(i)}\mathrm{diag}(\bar{\boldsymbol{\mathcal{X}}}_{(i)}^{\mathrm{H}})$$
(77)

Thus (74) can be written as

$$P_i = \Pr\left(\boldsymbol{\mathcal{Y}}_{(i)}^{\mathrm{H}}\Big(\boldsymbol{I} - \boldsymbol{D}\Big)\boldsymbol{\mathcal{Y}}_{(i)} \leq r\right)$$
(78)

note that (78) is of the same form as (49). The only difference in the LS and MAP costs is the presence of the term $\frac{1}{\rho}\,\boldsymbol{R}_h^{-1}$ in (77). Also note that this term depends on the inverse of the SNR. For low SNR, the inverse term in (77) is always invertible due to the regularization term. At high SNR, the effect of regularization fades and inverse term in (77) is invertible. At high SNR, i.e., $\rho \to \infty$, $\frac{1}{\rho}\,\boldsymbol{R}_h^{-1} \to 0$ and $\boldsymbol{D}$ of (76) takes the same form as that of LS cost leading to (72).

Table IV lists the estimated computational cost for our blind algorithm in the high SNR regime. Since there is no backtracking, the total number of iterations is N, which explains our calculations in Table IV. It thus follows that the total number of operations needed for our algorithm is of the order $O(LN)$ in high SNR regime. The pilot based approach for channel estimation needs to invert an $(L+1) \times (L+1)$ matrix (assuming we need $L+1$ pilots to estimate a channel of length $L+1$) with a complexity of the order $O(L^2)$. Since the cyclic prefix is a fixed fraction of the OFDM symbol ($L = N/m$ with m typically set to $m = 4$ or 8) we see that the complexity of the two approaches become comparable in the high SNR regime.

VI. SIMULATION RESULTS

We consider an OFDM system with $N = 16$, or 64 subcarriers and a CP of length $L = \frac{N}{4}$. The uncoded data symbols are modulated using BPSK, 4-QAM, or 16-QAM. The constructed OFDM signal then passes through a channel

TABLE IV
TOTAL COMPUTATIONAL COST OF THE ML BLIND AND TRAINING BASED ALGORITHMS AT HIGH SNR

Algorithm	$\times$	$+$
Blind Algorithm	$(3L^2 + 11L + 17)N$	$(2L^2 + 5L + 4)N$
Blind algorithm with carrier reordering	$(4L + 13)N$	$(2L + 4)N$
Training based algorithm [39]	$4L^2 + 17L + 13$	$2L^2 + 6L + 4$

of length $L + 1$, which is assumed to be block fading (i.e., constant over one OFDM symbol but fades independently from one symbol to another) and whose taps follow an exponential decay profile ($E[|h(t)|^2] = e^{-0.2t}$).

A. Bench marking

We compare the performance of our algorithm against the following receivers

1) the subspace-based[7] blind receiver of [10],
2) the sphere decoding based receiver of [28],
3) a receiver that acquires the channel through training with $L + 1$ pilots and a priori channel correlation $\boldsymbol{R}_h$ [39],
4) the ML receiver that acquires data through exhaustive search.

The simulations are averaged over 500 Monte-Carlo runs.

Figure 3 compares the BER performance of our algorithm with the aforementioned algorithms for an OFDM system with $N = 16$ subcarriers and BPSK data symbols. Note in particular that our blind algorithm outperforms both the subspace and sphere decoding algorithms and almost matches the performance of the exhaustive search algorithm for low and high SNR, which confirms the ML nature of the algorithm.

Figure 4, which considers the 4-QAM case, shows the same trends observed for the BPSK case of Figure 3.

Figure 5 considers a more realistic OFDM symbol length ($N = 64$), drawn from a 4-QAM constellation and allows the SNR to grow to 45 dB. Our blind algorithm shows no error floor signs, which is characteristic of non-ML methods. Furthermore, the algorithm beats the training-based method and follows closely the performance of the perfect channel case. Figure 6 shows the results with $N = 64$ subcarriers and 16-QAM data symbols for SNR as large as 50 dB. Again, the proposed blind algorithm does not reach an error floor.

B. Low-Complexity Variations

In this subsection, we investigate the low-complexity variants of our algorithm. Specifically, we consider the performance of the blind algorithm with

1) $\boldsymbol{P}_i$ set to $\boldsymbol{I}$,
2) $\boldsymbol{P}_i$ set to $\boldsymbol{I}$ with subcarrier reordering

---

[7]The block fading assumption is maintained for all simulations. However, for the subspace blind receiver of [10] to work, the channel needs to stay constant over a sequence of OFDM symbols. For this particular receiver, the channel was kept fixed over 50 OFDM symbols.
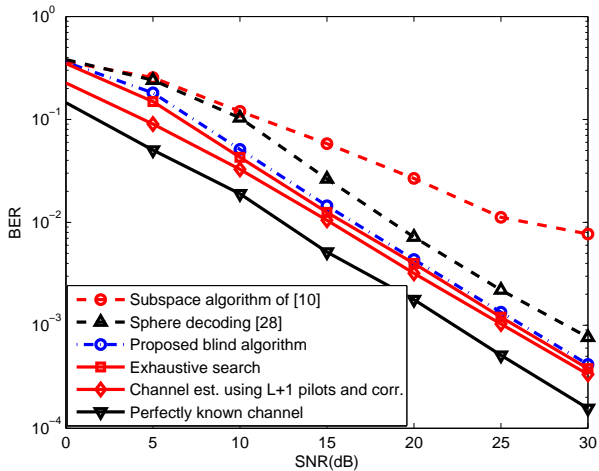
Fig. 3. BER vs SNR for BPSK OFDM over a Rayleigh channel with $N = 16$ and $L = 3$
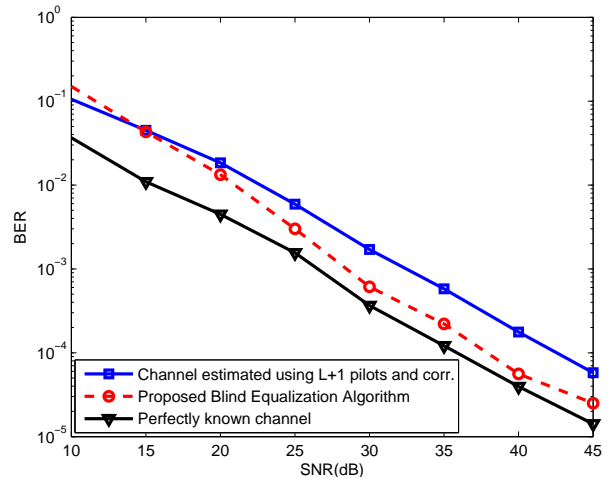

Fig. 5. BER vs SNR for 4-QAM OFDM over a Rayleigh channel with $N = 64$ and $L = 15$
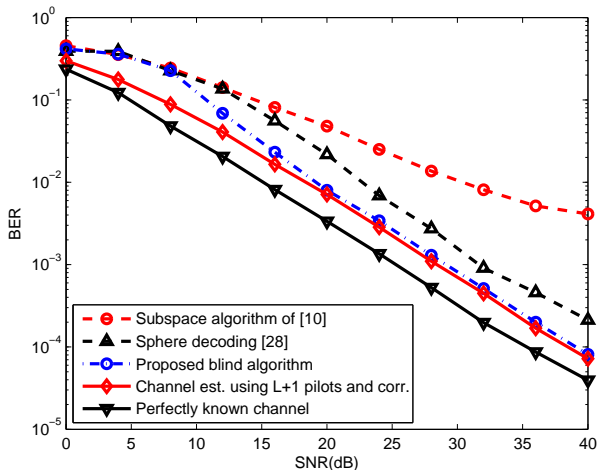

Fig. 4. BER vs SNR for 4-QAM OFDM over a Rayleigh channel with $N = 16$ and $L = 3$
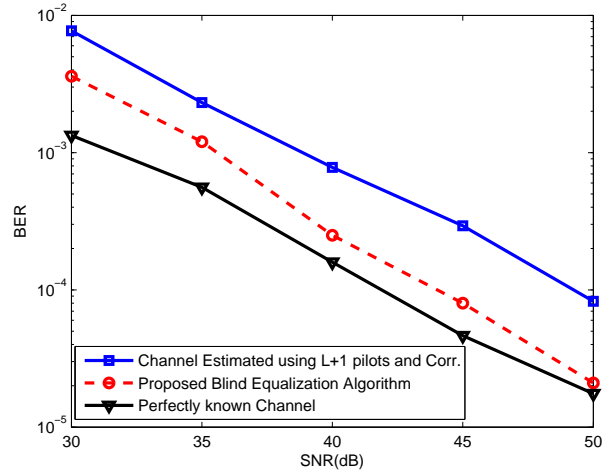

Fig. 6. BER vs SNR for 16-QAM OFDM over a Rayleigh channel with $N = 64$ and $L = 15$

Figure 7 exhibits the comparisons for the various algorithms for BPSK and $N = 16$. Note that with $\boldsymbol{P}_i$ set to $\boldsymbol{I}$ arbitrarily, the performance of the blind algorithm deteriorates and the BER reaches an error floor. Contrast this with the algorithm variant that uses subcarrier reordering as well, and note that the performance of this variant follows closely the performance of the exact blind algorithm. Also note that the BER of both of these algorithms beats that of the sphere decoding algorithm of [28]. The same trends are observed in Figure 8, which considers the 4-QAM case.

Figure 9 compares the average runtime of various algorithms as a function of the SNR. Note first that the extreme cases are the training-based receiver and the exhaustive search receiver, both of which are independent of the SNR. The runtime of the proposed algorithm decreases with the SNR and is sandwiched in-between the run time of the sphere decoding algorithm and that of the subspace algorithm for all values of

the SNR[8]. Note that in the high SNR regime our algorithm runs at the same speed as the subspace algorithm.

Figure 10 shows the average runtime of the proposed algorithm with $N = 16$ for various modulation schemes (BPSK, 4-QAM and 16-QAM). It is clear from the figure that the average runtime decreases considerably at higher SNR values.

VII. CONCLUSION

In this paper, we have proposed a low-complexity blind algorithm that is able to deal with channels that change on a symbol by symbol basis allowing it to deal with fast block fading channels. The algorithm works for general constellations and is able to recover the data from output observations only.

[8]The runtime of the subspace algorithm is adjusted to account for the fact that it requires the channel to be constant over a block of $L + 1$ OFDM symbols.
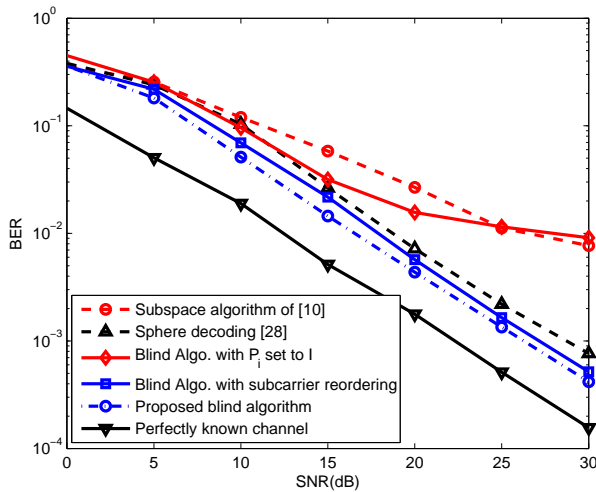
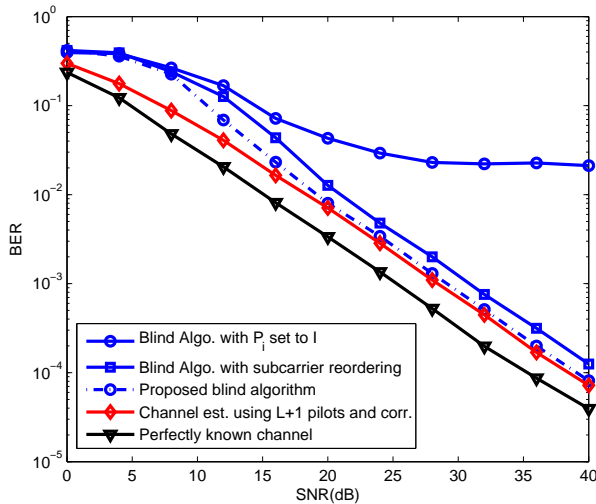Fig. 7. Comparison of low-complexity algorithms for BPSK OFDM with $N = 16$ and $L = 3$

Fig. 9. Average time comparison for BPSK data symbols with $N = 16$ and $L = 3$

Fig. 8. Comparison of low-complexity algorithms for 4-QAM OFDM with $N = 16$ and $L = 3$
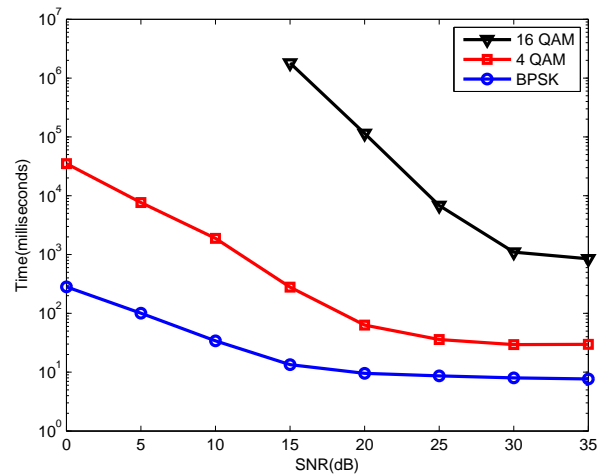

Fig. 10. Average Time Comparison for our Blind Algorithm for Different Modulation with $N = 16$ and $L = 3$

Simulation results demonstrate the favorable performance of the algorithm for general constellations and show that its performance matches the performance of the exhaustive search for small values of N.

We have also proposed an approximate blind equalization method (avoiding $\boldsymbol{P}_i$ with subcarrier reordering) to reduce the computational complexity. As evident from the simulation results, this approximate method performs quite close to the exact blind algorithm and can work properly without a priori knowledge of the channel statistics. Finally, we study the complexity of our blind algorithm and show that it becomes especially low in the high SNR regime.

ACKNOWLEDGMENT

REFERENCES

[1] C. Shin, R. W. Heath, Jr., and E. J. Powers, "Blind channel estimation for MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 670-685, Mar. 2007.
[2] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Process.*, vol. 43, no. 2, pp. 516-525, Feb. 1995.
[3] C.-C. Tu and B. Champagne, "Subspace blind MIMO-OFDM channel estimation with short averaging periods: Performance analysis," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Las Vegas, Neveda, April 2008, pp. 24-29.
[4] R. W. Heath Jr. and G. B. Giannakis, "Exploiting input cyclostationarity for blind channel identiffcation in OFDM systems," *IEEE Trans. Sig. Proc.*, vol. 47, no. 3, pp.848-856, Mar. 1999.
[5] B.Su and P.P. Vidyanathan, "Subspace-based blind channel identification for cyclic prefix systems using few received blocks", *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 4979-4993, Oct. 2007.

[6] C.-C. Tu and B. Champagne, "Subspace-based blind channel estimation for MIMO-OFDM systems with reduced time averaging," *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1539-1544, Mar. 2010.

[7] Y. Zeng and T. S. Ng, "A semi-blind channel estimation method for multiuser multiuser multiantenna OFDM systems," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1419-1429, May 2004.

[8] X. G. Doukopoulos and G.V. Moustakides, "Blind adaptive channel estimation in OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 5, no. 7, pp. 1716-1725, Jul. 2006.

[9] F. Gao, Y Zeng, A. Nallanathan, T.S. Ng, "Robust subspace blind channel estimation for cyclic prefixed MIMO OFDM systems", *IEEE J. Selet. Areas Commun.*, vol.26, no.2, pp. 378-388, Feb. 2008.

[10] B. Muquet, M. de Courville, and P. Duhamel, "Subspace-based blind and semi-blind channel estimation for OFDM systems," *IEEE Trans. Signal Process.*, vol. 50, no. 7, pp. 1699-1712, Jul. 2002.

[11] H. Bölcskei, R. W. Heath, Jr. and A. J. Paulraj, "Blind channel identification and equalization in OFDM-based multiantenna systems," *IEEE Trans. Signal Process.*, vol. 50, no. 1, pp. 96-109, Jan. 2002.

[12] F. Gao and A. Nallanathan, "Blind channel estimation for MIMO OFDM systems via nonredundant linear precoding"*IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 784-789, Feb. 2007.

[13] H. Muarkami, "Blind estimation of a fractionally sampled FIR channel for OFDM transmission using residue polynomials", *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 225-234, Jan, 2006.

[14] J. Choi and C.C. Lim, "Cholesky factorization based approach for blind FIR channel identification," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1730-1735, April 2008.

[15] S. A. Banani and R. G. Vaughan, "OFDM with iterative blind channel estimation", *IEEE Trans. on Veh. Technol.*, vol. 59, no. 9, Nov. 2010.

[16] C. Li and S. Roy, "Subspace-based blind channel estimation for OFDM by exploiting virtual carriers," *IEEE Trans. Wireless Commun.*, vol. 2, no. 1, pp. 141-150, Jan. 2003.

[17] F. Gao, A. Nallanathan and C. Tellambura, "Blind channel estimation for cyclic prefixed single-carrier systems by exploiting real symbol characteristics," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 2487-2498, Sep. 2007.

[18] A. Petropulu, R. Zhang and R. Lin, "Blind OFDM channel estimation through simple linear precoding," *IEEE Trans. Wireless Commun.*, vol. 3, no. 2, pp. 647-655, Mar. 2004.

[19] W. Xu, M. Stojnic and B. Hassibi, "Low-complexity blind maximum-likelihood detection for SIMO systems with general constellation," *IEEE Int. Conf. on Acoust., Speech and Signal Process. (ICASSP)*, Las Vegas, Neveda, Apr. 2008, vol. 1, pp. 2817-2820.

[20] N. Sarmadi, S. Shahbazpanahi and A. B. Greshman, "Blind channel estimation in orthogonally coded MIMO-OFDM systems", *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2354-2364, June 2009.

[21] W. Ma, B. Vo, T. Davidson, and P. Ching, "Blind ML detection of orthogonal space-time block codes: efficient high-performance implementations," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 738-751, 2006.

[22] E. Larsson, P. Stoica, and J. Li, "On maximum-likelihood detection and decoding for space-time coding systems," *IEEE Trans. Signal Process.*, vol. 50, no. 4, pp. 937-944, 2002.

[23] E. G. Larsson, P. Stoica, and J. Li, "Orthogonal space-time block codes: Maximum likelihood detection for unknown channels and unstructured interferences," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 362-372, 2003.

[24] P. Stoica and G. Ganesan, "Space-time block codes: Trained, blind, and semi-blind detection," *Digital Signal Processing*, vol. 13, pp. 93-105, 2003.

[25] W.-K. Ma, "Blind ML detection of orthogonal spacetime block codes: Identifiability and code construction," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3312-3324, Jul. 2007.

[26] T. Y. Al-Naffouri and A. A. Quadeer, "Cyclic prefix based enhanced data recovery in OFDM," *IEEE Trans. on Signal Process.*, vol. 58, no. 6, pp. 3406-3410, June, 2010.

[27] Y. Li, C. Georghiades, and G. Huang, "Iterative maximum likelihood sequence estimation for space-time coded systems," *IEEE Trans. Commun.*, vol. 49, no. 6, pp. 948-951, 2001.

[28] T. Cui and C.Tellambura, "Joint data detection and channel estimation for OFDM systems", *IEEE Trans. Commun.,*, vol. 54, no. 4, pp. 670-679, April 2006.

[29] A. Gallo, E. Chiavaccini, F. Muratori, and G. Vitetta, "BEM-based SISO detection of orthogonal space-time block codes over frequency flat-fading channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1885-1889, 2004.

[30] Y. Hua, "Fast maximum likelihood for blind identification of multiple FIR channels," *IEEE Trans. Signal Process.*, vol. 44, pp. 661-672, Mar. 1996.

[31] B. P. Paris, "Self-adaptive maximum-likelihood sequence estimation," in *Proc. IEEE GLOBECOM*, 1993.

[32] N. Seshadri, "Joint data and channel estimation using blind trellis search techniques," *IEEE Trans. Commun.*, vol. 42, pp. 1000-1011, Feb.-Apr. 1994.

[33] M. Ghosh and C. L. Weber, "Maximum-likelihood blind equalization," *Optical Engineering*, vol. 31, no. 6, pp. 1224-1228, June 1992.

[34] S. Talwar, M. Viberg, and A. Paulraj, "Blind estimation of multiple co-channel digital signals using an antenna array," *IEEE Signal Process. Lett.*, vol. 1, pp. 29-31, Feb. 1994.

[35] D. Yellin and B. Porat, "Blind identification of FIR systems excited by discrete-alphabet inputs," *IEEE Trans. Signal Process.*, vol. 41, pp. 1331-1339, Mar. 1993.

[36] M. C. Necker and G. L. Stüber, "Totally blind channel estimation for OFDM on fast varying mobile radio channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 15141525, Sep. 2004.

[37] T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Maximum-likelihood detection of orthogonal spacetime block coded OFDM in unknown block fading channels," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1637-1649, Apr. 2008.

[38] Ali H. Sayed, *Fundamentals of Adaptive Filtering.* John Wiley and Sons, Inc., 2003.

[39] T. Y. Al-Naffouri, A. Bahai, and A. Paulraj, "Semi-blind channel identification and equalization in OFDM: an expectation-maximization approach," in *Proc. IEEE Veh. Technol. Conf.*, Vancouver, Canada, Sep. 2002, vol. 1, pp. 13−17.

[40] R. A. Horn, C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1990.

[41] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, Philadelphia, PA: SIAM, 2000.