RESEARCH

Blind and semi-blind ML detection for space-time block-coded OFDM wireless systems

Alam Zaib[1*] and Tareq Y Al-Naffouri[1,2]

Abstract

This paper investigates the joint maximum likelihood (ML) data detection and channel estimation problem for Alamouti space-time block-coded (STBC) orthogonal frequency-division multiplexing (OFDM) wireless systems. The joint ML estimation and data detection is generally considered a hard combinatorial optimization problem. We propose an efficient low-complexity algorithm based on branch-estimate-bound strategy that renders exact joint ML solution. However, the computational complexity of blind algorithm becomes critical at low signal-to-noise ratio (SNR) as the number of OFDM carriers and constellation size are increased especially in multiple-antenna systems. To overcome this problem, a semi-blind algorithm based on a new framework for reducing the complexity is proposed by relying on subcarrier reordering and decoding the carriers with different levels of confidence using a suitable reliability criterion. In addition, it is shown that by utilizing the inherent structure of Alamouti coding, the estimation performance improvement or the complexity reduction can be achieved. The proposed algorithms can reliably track the wireless Rayleigh fading channel without requiring any channel statistics. Simulation results presented against the perfect coherent detection demonstrate the effectiveness of blind and semi-blind algorithms over frequency-selective channels with different fading characteristics.

Keywords: Blind ML detection; OFDM; Space-time block coding

1 Introduction

The increasing demand for higher data rates in recent years has called for transmissions over a broadband wireless channel which is frequency selective. The wireless channel is thus prone to inter-symbol interference (ISI) which severely degrades the system performance and requires complex equalization techniques at the receiver. The orthogonal frequency-division multiplexing (OFDM) has emerged as the most promising scheme to combat ISI and improve system performance. OFDM essentially transforms a broadband channel into a number of parallel narrowband channels using a cyclic prefix (CP) of appropriate length and renders simple one-tap channel equalizer for each OFDM subcarrier [1,2]. Due to these crucial advantages, OFDM is not only being used in many

existing standards such as digital subscriber line (DSL), WLAN standards (IEEE 802.11 a/b/g), LAN standard (IEEE 802.11n) [3], MAN standards (IEEE 802.16e), digital audio broadcast (DAB), and digital video broadcast (DVB) [4] but also adopted for future wireless standards such as LTE and 4G+ [5]. Besides OFDM, the spatial dimensions in wireless communications are often exploited to further enhance the system capacity and /or improve the transmission reliability by employing multiple antennas at the transmitter and/or receiver. This offers many advantages over single-antenna systems including multiplexing gain and diversity gain [6]. Of several diversity schemes available in the literature, the Alamouti scheme [7] with two transmit and one receive antenna is the optimum in both the capacity and the diversity. Alamouti coding achieves full spatial diversity at full transmission rate for any signal (real or complex) constellation and offers very simple receiver structures. However, to decouple the signals at the receiver side via simple decoding,

*Correspondence: alamzaib@kfupm.edu.sa
[1]Department of Electrical Engineering, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia
Full list of author information is available at the end of the article

the Alamouti scheme requires the channel between each transmit-receive antenna to be constant over two consecutive OFDM symbols. Moreover, when dealing with frequency-selective channel, the Alamouti scheme has to be implemented over the block level.

In many wireless communications studies, it is often assumed that channel state information (CSI) is available at the receiver side for coherent data detection. This assumption is certainly not realistic. The current standards use pilot symbols to estimate the channel, thus sacrificing bandwidth which otherwise would have been available for data transmission. In high-mobility wireless systems, the channels may even change so rapidly that this approach will become infeasible. Blind or semi-blind detection over time-varying wireless channels has shown to enhance the system performance considerably [8,9].

There are numerous blind estimation and equalization techniques available in the literature, namely, subspace-based methods [10,11], second-order statistics [12], Cholesky factorization [13] , and iterative methods [14]. These methods either suffer from slow convergence, higher computational costs or assume channel to be stationary over several OFDM symbols. These drawbacks make maximum likelihood (ML)-based approaches, e.g., [15,16] more attractive due to their fast convergence despite having the higher computational cost. Usually, suboptimal techniques are employed to reduce the computational cost by restricting the search space of exhaustive ML search. Some of the suboptimal techniques are applicable to specific constant modulus constellations [17,18]. Recently, in [19] and [20], the authors have proposed a low-complexity blind ML method for general constellations for single-input-multiple-output (SIMO) and single-input-single-output (SISO) systems, respectively, which form the basis of our paper.

1.1 Contributions
In this paper, we extend the previous algorithm developed for SISO OFDM systems [20] to multiple-input-single-output (MISO) OFDM systems with two transmit antennas and employing Alamouti coding. Parallelizing the results and discussions therein, we first derive the exact blind ML algorithm based on branch-estimate-and-bound strategy and then reduce its complexity using different methods. Then, a semi-blind version is proposed by assuming that few training symbols are available. The proposed algorithms offers low complexity, fast convergence, works for signals drawn from general modulation constellations, and do not require any channel statistics. The contributions are listed as:

1. Extension of blind algorithm for SISO to MISO OFDM systems and derivation of bound used to abort a certain search path in the blind search

   algorithm. The bound is calculated in a recursive manner offering low computations.
2. Using semi-orthogonality of the subcarriers by exploiting the structure of fast Fourier transform (FFT) matrix in order to reduce the computations for calculation of bound during the blind search.
3. Reducing the complexity at low signal-to-noise ratio (SNR) regime by employing the reliable carriers to accelerate the convergence of blind algorithm by reordering the subcarriers according to their reliability. Calculation of reliability, however, requires initial tentative channel estimates; therefore, a semi-blind mode is adopted.
4. Reducing the complexity and or improving the estimation performance of proposed algorithm by exploiting the orthogonal structure of Alamouti coding.

The framework of semi-blind algorithm is based on reliably decoding the carriers by computing the vector-wise likelihood ratio first suggested in [21]. By supplying the algorithm with reordered carriers according to their reliability index, the backtracking is minimized which is the major source of complexity especially at low SNR. At higher SNR regime, it has been shown that the probability of backtracking is almost zero [20], while the number of operations is of the order of constant times the total number of OFDM subcarriers.

1.2 Organization of the paper
Section 2 describes the Alamouti coded OFDM system model with frequency-selective time-variant channels, and Section 3 describes the proposed blind ML algorithm. In Section 4, a low complexity variant of blind ML algorithm is derived by exploiting the structure of the FFT matrix. To further reduce the complexity of the algorithm, we propose a semi-blind algorithm in Section 5, and in Section 6, we show how the structure of the Alamouti coding can be utilized to favour certain advantages. Simulation results are detailed in Section 7, and we conclude in Section 8.

1.3 Notation
We use lower case letters x to denote scalars, lower case boldface $\mathbf{x}$ to denote (column) vectors and $\mathbf{x}(i)$ to denote individual entries of a vector. Matrices are denoted by upper boldface letters $\mathbf{X}$, whereas the calligraphic notations $\mathcal{X}$ is reserved for vectors in frequency domain. We also use $\mathbf{x}_{(i)}$ to represent a partial vector consisting of first i elements of $\mathbf{x}$. $(.)^T$, $(.)^*$ and $(.)^H$ represent transpose, conjugate and conjugate transpose (hermitian) operations, respectively. $\langle \hat{\mathcal{X}}(k) \rangle$ will denote the hard decoding decision that maps $\hat{\mathcal{X}}(k)$ to $\mathcal{X}(k)$. The FFT and inverse FFT (IFFT) matrices are denoted by $\mathbf{Q}$ and $\mathbf{Q}^H$ respectively,

where we define $\mathbf{Q}$ as $q_{l,k} = e^{-j2\pi lk/N}$ with $l, k = 0, 1, 2, \cdots, N - 1$. The notation $\|\mathbf{a}\|_{\mathbf{B}}^2$ represents weighted norm defined as $\|\mathbf{a}\|_{\mathbf{B}}^2 = \mathbf{a}^H \mathbf{B} \mathbf{a}$.

2 System model

Consider a single-user OFDM system with two transmit and one receive antenna as shown in Figure 1. The frequency-selective channels from two transmit antennas to the receive antenna are modelled as finite impulse response (FIR) filters. We assume that both channels are independent Rayleigh-fading channels having maximum length L and that OFDM CP length is at least $L - 1$ to avoid ISI.

Let $\mathcal{X}$ represent information symbols and that OFDM system has N subcarriers so that after IFFT operation the OFDM symbol can be written as

$$\mathbf{x} = \sqrt{N}\mathbf{Q}^H \mathcal{X}, \tag{1}$$

where $\mathbf{Q}$ is FFT matrix with $[\mathbf{Q}]_{l,k} = e^{-j2\pi lk/N}$. Let the nth symbol of kth transmitted block from antenna i (= 1 or 2) be denoted by $x_i^{(k)}(n)$, with $n = 0, 1, \cdots, N - 1$. At times $k = 0, 2, 4, \cdots$ pair of blocks $x_1^{(k)}(n)$ and $x_2^{(k)}(n)$ are generated according to the following Alamouti coding rule [7,22]:

$$
\begin{aligned}
x_1^{(k+1)}(n) &= -x_2^{*(k)}((n)_N) \\
x_2^{(k+1)}(n) &= x_1^{*(k)}((n)_N),
\end{aligned}
\tag{2}
$$

where, $(.)_N$ is the modulo N operation. Each antenna transmits a data block of length N according to space-time block-coded (STBC) scheme after appending the CP. Adding CP eliminates inter-block interference and converts linear convolution into circular convolution. In the presence of additive white Gaussian noise (AWGN), the received data blocks over two consecutive time instants after discarding the CPs can be written as

$$\mathbf{y}^{(j)} = \sqrt{\rho}\,\mathbf{H}_1 \mathbf{x}_1^{(j)} + \sqrt{\rho}\,\mathbf{H}_2 \mathbf{x}_2^{(j)} + \mathbf{n}^{(j)}, j = k, k + 1, \tag{3}$$

where ρ is the SNR, $\mathbf{H}_1$ and $\mathbf{H}_2$ are circular channel matrices from transmit antenna-1 and transmit antenna-2 to receive antenna, respectively, and $\mathbf{n}$ is circular symmetric AWGN with *pdf:* $\mathbf{n} \sim CN(\mathbf{0}, \mathbf{I})$. In (3), we also assumed that channel is static over two consecutive OFDM blocks

at time instants k and $k + 1$. Specifically, the structure of two circular channel matrices is

$$
\mathbf{H}_i = \begin{pmatrix}
h_i(0) & 0 & \cdots & h_i(L-1) & \cdots & h_i(1) \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
h_i(L-2) & \cdots & h_i(0) & 0 & \cdots & h_i(L-1) \\
h_i(L-1) & h_i(L-2) & \cdots & h_i(0) & 0 & \vdots \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & h_i(L-1) & h_i(L-2) & \cdots & h_i(0)
\end{pmatrix}
$$

and where

$$\mathbf{h}_i = \begin{bmatrix} h_i(0) & h_i(1) & \cdots & h_i(L-1) \end{bmatrix}^T \tag{4}$$

represents the impulse response sequence of the ith channel matrix. At the receiver side, the frequency domain received symbols after FFT operations are obtained as

$$\mathcal{Y}^{(j)} = \sqrt{\rho}\,\Lambda_1 \mathcal{X}_1^{(j)} + \sqrt{\rho}\,\Lambda_2 \mathcal{X}_2^{(j)} + \mathcal{N}^{(j)}, j = k, k + 1, \tag{5}$$

where $\mathcal{X}_i^{(j)} = \frac{1}{\sqrt{N}}\mathbf{Q}\mathbf{x}_i^{(j)}$ and $\Lambda_i = \mathbf{Q}\mathbf{H}_i\mathbf{Q}^H$ are diagonal matrices whose entries are N-point discrete Fourier transform (DFT) of $\mathbf{h}_i$ after zero padding and $\mathcal{N}^{(j)} = \frac{1}{\sqrt{N}}\mathbf{Q}\mathbf{n}^{(j)}$. Expanding (5) and using DFT properties, we get:

$$
\begin{aligned}
\mathcal{Y}^{(k)} &= \sqrt{\rho}\,\Lambda_1 \mathcal{X}_1^{(k)} + \sqrt{\rho}\,\Lambda_2 \mathcal{X}_2^{(k)} + \mathcal{N}^{(k)}, \\
\mathcal{Y}^{(k+1)} &= \sqrt{\rho}\,\Lambda_1 \mathcal{X}_1^{(k+1)} + \sqrt{\rho}\,\Lambda_2 \mathcal{X}_2^{(k+1)} + \mathcal{N}^{(k+1)}
\end{aligned}
\tag{6}
$$

By stacking the received data symbols over consecutive intervals in one column and so as the DFT channel coefficients, (6) can be written in matrix-vector notation as

$$
\begin{bmatrix} \mathcal{Y}^{(k)} \\ \mathcal{Y}^{(k+1)} \end{bmatrix} = \sqrt{\rho} \begin{bmatrix} \text{diag}\left(\mathcal{X}_1^{(k)}\right) & \text{diag}\left(\mathcal{X}_2^{(k)}\right) \\ -\text{diag}\left(\mathcal{X}_2^{*(k)}\right) & \text{diag}\left(\mathcal{X}_1^{*(k)}\right) \end{bmatrix} \begin{bmatrix} \mathcal{H}_1 \\ \mathcal{H}_2 \end{bmatrix} + \begin{bmatrix} \mathcal{N}^{(k)} \\ \mathcal{N}^{(k+1)} \end{bmatrix}, \tag{7}
$$

where $\mathcal{H}_i = \text{diag}(\Lambda_i) = \mathbf{Q}\begin{bmatrix} \mathbf{h}_i \\ \mathbf{0} \end{bmatrix}$. Let $\mathbf{A}^H$ consists of first L columns of $\mathbf{Q}$, then

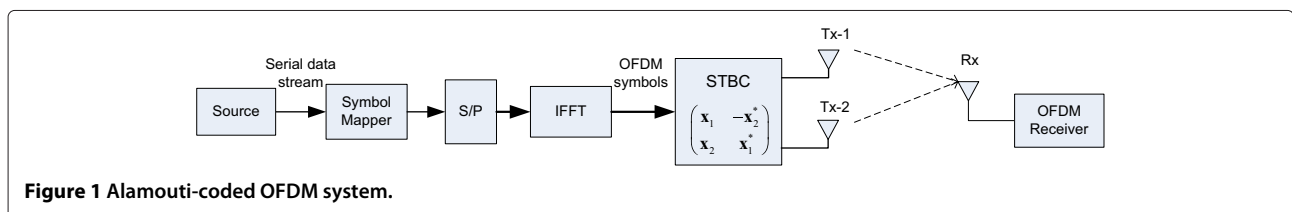$$\mathcal{H}_i = \mathbf{A}^H \mathbf{h}_i \text{ and } \mathbf{h}_i = \mathbf{A}\mathcal{H}_i \tag{8}$$

Figure 1 Alamouti-coded OFDM system.

This allows us to rewrite (7) as

$$
\underbrace{\begin{bmatrix} \boldsymbol{\mathcal{Y}}^{(k)} \\ \boldsymbol{\mathcal{Y}}^{(k+1)} \end{bmatrix}}_{\boldsymbol{\mathcal{Y}}} = \sqrt{\rho} \underbrace{\begin{bmatrix} \mathrm{diag}\left(\boldsymbol{\mathcal{X}}_1^{(k)}\right)\mathbf{A}^H & \mathrm{diag}\left(\boldsymbol{\mathcal{X}}_2^{(k)}\right)\mathbf{A}^H \\ -\mathrm{diag}\left(\boldsymbol{\mathcal{X}}_2^{*(k)}\right)\mathbf{A}^H & \mathrm{diag}\left(\boldsymbol{\mathcal{X}}_1^{*(k)}\right)\mathbf{A}^H \end{bmatrix}}_{\mathbf{X}_a}
$$

$$
\times \underbrace{\begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{bmatrix}}_{\mathbf{h}} + \underbrace{\begin{bmatrix} \boldsymbol{\mathcal{N}}^{(k)} \\ \boldsymbol{\mathcal{N}}^{(k+1)} \end{bmatrix}}_{\boldsymbol{\mathcal{N}}}
$$

$$(9)$$

or even more compactly as

$$
\boldsymbol{\mathcal{Y}} = \sqrt{\rho}\,\mathbf{X}_a\mathbf{h} + \boldsymbol{\mathcal{N}}, \tag{10}
$$

where $\boldsymbol{\mathcal{Y}}$ and $\boldsymbol{\mathcal{N}}$ are $2N \times 1$ observed data and noise vectors, respectively, $\mathbf{X}_a$ is a $2N \times 2L$ data matrix, which we shall refer to as *Alamouti matrix*, and $\mathbf{h}$ is a $2L \times 1$ composite channel vector. This model can be easily transformed to SISO-OFDM system of [20] by replacing $\mathbf{X}_a$ with $N \times N$ square matrix $\mathrm{diag}(\boldsymbol{\mathcal{X}})$ containing N data symbols on its diagonal. Specifically, the SISO model corresponding to (10) is given by

$$
\boldsymbol{\mathcal{Y}} = \sqrt{\rho}\,\mathrm{diag}(\boldsymbol{\mathcal{X}})\mathbf{h} + \boldsymbol{\mathcal{N}}, \tag{11}
$$

where $\boldsymbol{\mathcal{Y}}$ and $\boldsymbol{\mathcal{N}}$ are N-dimensional received OFDM symbol and noise vector, respectively, while $\mathbf{h}$ is the length L SISO channel vector. In either case, the task of receiver is to jointly estimate the channel $\mathbf{h}$ and the data vector $\boldsymbol{\mathcal{X}}$ given only the received data symbol $\boldsymbol{\mathcal{Y}}$.

3 Joint ML/MAP solution

Considering the data model in (10), the joint ML channel estimation and data detection problem reduces to minimizing the following objective function:

$$
J_{\mathrm{ML}} = \min_{\mathbf{h}, \boldsymbol{\mathcal{X}} \in \Omega^{2N}} \left\{ \|\boldsymbol{\mathcal{Y}} - \sqrt{\rho}\,\mathbf{X_a}\mathbf{h}\|^2 \right\}, \tag{12}
$$

where Ω^{2N} denotes all possible $2N$-dimensional signal vectors. As seen from (12), the joint ML problem is a combinatorial problem involving $|\Omega|^{2N}$ hypothesis tests, and it is almost impossible to solve it exactly for sufficiently large Ω and N.

To solve it efficiently, we propose the following strategy. We start by decomposing the original cost function as

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

$$(13)$$

and define

$$
M_{\mathcal{X}_{(i)}} = \|\boldsymbol{\mathcal{Y}}_{(i)} - \sqrt{\rho}\,\mathbf{X}_{a(i)}\mathbf{h}\|^2 \tag{14}
$$

as the partial joint ML metric up to the index i for $\boldsymbol{\mathcal{X}}$, where $\mathbf{X}_{a(i)} = \begin{bmatrix} \mathrm{diag}\left(\boldsymbol{\mathcal{X}}_{1(i)}^{(k)}\right)\mathbf{A}_{(i)}^H & \mathrm{diag}\left(\boldsymbol{\mathcal{X}}_{2(i)}^{(k)}\right)\mathbf{A}_{(i)}^H \\ -\mathrm{diag}\left(\boldsymbol{\mathcal{X}}_{2(i)}^{*(k)}\right)\mathbf{A}_{(i)}^H & \mathrm{diag}\left(\boldsymbol{\mathcal{X}}_{1(i)}^{*(k)}\right)\mathbf{A}_{(i)}^H \end{bmatrix}$ is a partial Alamouti matrix of dimension $2i \times 2L$, $\mathbf{X}_a(j)$, the $2 \times 2L$ matrix is the same as $\mathbf{X}_{a(j)}$ with all $\boldsymbol{\mathcal{X}}_{(j)}$ replaced by $\boldsymbol{\mathcal{X}}(j)$, $\boldsymbol{\mathcal{Y}}_{(i)} = \left[\left[\boldsymbol{\mathcal{Y}}_{(i)}^{(k)}\right]^T \left[\boldsymbol{\mathcal{Y}}_{(i)}^{(k+1)}\right]^T\right]^T$ is the partial data vector of dimension $2i \times 1$ and the partial matrix $\mathbf{A}_{(i)}^H$ consists of first i rows of $\mathbf{A}^H$. It should be noted that partial Alamouti matrix $\mathbf{X}_{a(i)}$ is the function of the first i data points, while $\mathbf{X}_a(i)$ is a function of the ith data point. Obviously, the solution that minimizes this partial joint ML metric is not the globally optimal. But we have the following lemma[a]:

Lemma 1. *Let R represent the optimal value of the objective function in (12). If $M_{\mathcal{X}_{(i)}} > R$, then $\boldsymbol{\mathcal{X}}_{(i)}$ cannot be the ML solution $\hat{\boldsymbol{\mathcal{X}}}_{(i)}^{ML}$ of (12). In other words, for any estimate $\hat{\boldsymbol{\mathcal{X}}}_{(i)}$ to correspond to the ML solution, we should have $M_{\mathcal{X}_{(i)}} < R$.*

From Lemma 1, if the optimal value R of objective function (12) can be estimated, then we can adopt the following tree search procedure for joint estimation and detection: At each subcarrier i, make a guess of new value of $\boldsymbol{\mathcal{X}}(i) = \begin{bmatrix} \mathcal{X}_1(i) & \mathcal{X}_2(i) \end{bmatrix}^T$ and use that along with previous estimates to construct $\hat{\boldsymbol{\mathcal{X}}}_{(i)}$ and $\hat{\mathbf{X}}_{a(i)}$. Then, estimate $\mathbf{h}$ to minimize the associated cost function:

$$
M_{\hat{\mathcal{X}}_{(i)}} = \min_{\mathbf{h}} \left\{ \|\boldsymbol{\mathcal{Y}}_{(i)} - \sqrt{\rho}\,\hat{\mathbf{X}}_{(i)}\mathbf{h}\|^2 \right\} \tag{15}
$$

and calculate the resulting metric $M_{\hat{\mathcal{X}}_{(i)}}$. If $M_{\hat{\mathcal{X}}_{(i)}} < R$, then proceed to the next subcarrier $i + 1$; otherwise, backtrack and change the guess of $\boldsymbol{\mathcal{X}}(j)$ for some $j \leq i$. We call this approach as the branch-estimate-and-bound strategy, which reduces the search space of exhaustive ML search to those (partial) sequences that satisfy the given constraint $M_{\hat{\mathcal{X}}_{(i)}} < R$. This approach however does not work for $i \leq L$ as $\mathbf{X}_{a(i)}$ will be full rank for any choice of $\boldsymbol{\mathcal{X}}_{(i)}$, and therefore, $\mathbf{h}$ with $2L$ degrees of freedom can always be chosen by least squares (LS) to yield the trivial (zero) value for $M_{\hat{\mathcal{X}}_{(i)}}$. To obtain a non-trivial value of $M_{\hat{\mathcal{X}}_{(i)}}$, we have to use L pilots, but it would defeat our original motive of *blind* estimation. To overcome this problem, we adopt weighted regularized LS and instead of minimizing the ML objective function, J_{ML}, we minimize the maximum *a posteriori* (MAP) objective function:

$$
J_{\mathrm{MAP}} = \min_{\mathbf{h}, \boldsymbol{\mathcal{X}} \in \Omega^{2N}} \left\{ \|\mathbf{h}\|_{\mathbf{R}_h^{-1}}^2 + \|\boldsymbol{\mathcal{Y}} - \sqrt{\rho}\,\mathbf{X}_a\mathbf{h}\|^2 \right\}, \tag{16}
$$

where $\mathbf{R}_h$ is the block diagonal autocorrelation matrix of the composite channel vector $\mathbf{h}$, i.e., $\mathbf{R}_h = E\{\mathbf{h}\mathbf{h}^H\}$. The objective function in (16) can also be decomposed as

$$J_{\text{MAP}} = \min_{\mathbf{h}\mathcal{X}\in\Omega^{2N}} \left\{ \underbrace{\|\mathbf{h}\|^2_{\mathbf{R}_h^{-1}} + \|\mathcal{Y}_{(i)} - \sqrt{\rho}\,\mathbf{X}_{a(i)}\mathbf{h}\|^2}_{M_{\mathcal{X}_{(i)}}} \right. $$
$$\left. + \sum_{j=i+1}^{N} \|\mathcal{Y}(j) - \sqrt{\rho}\,\mathbf{X}_a(j)\mathbf{h}\|^2 \right\} \tag{17}$$

So, if we have a guess of $\hat{\mathcal{X}}_{(i-1)}$, the partial metric for $\mathcal{X}$ up to index $i-1$ can be written as

$$M_{\hat{\mathcal{X}}_{(i-1)}} = \min_{\mathbf{h}}\left\{\|\mathbf{h}\|^2_{\mathbf{R}_h^{-1}} + \|\mathcal{Y}_{(i-1)} - \sqrt{\rho}\,\hat{\mathbf{X}}_{a(i-1)}\mathbf{h}\|^2\right\} \tag{18}$$

whose optimum value $\hat{\mathbf{h}}$ and the cost can be computed [23].

3.1 Recursive derivation of bound

For our blind search strategy, the calculation of the metric or bound $M_{\hat{\mathcal{X}}_{(i)}}$ is needed at each tree node for comparison with the optimal value of objective function, R. This bound can be derived recursively by simply expressing $M_{\hat{\mathcal{X}}_{(i)}}$ in terms of new observation and an additional regressor $\hat{\mathbf{X}}_a(i)$ as follows:

$$M_{\hat{\mathcal{X}}_{(i)}} = \min_{\mathbf{h}}\left\{\|\mathbf{h}\|^2_{\mathbf{R}_h^{-1}} + \|\mathcal{Y}_{(i)} - \sqrt{\rho}\,\hat{\mathbf{X}}_{a(i)}\mathbf{h}\|^2\right\}$$
$$= \min_{\mathbf{h}}\left\{\|\mathbf{h}\|^2_{\mathbf{R}_h^{-1}} + \left\|\begin{bmatrix}\mathcal{Y}_{(i-1)} \\ \mathcal{Y}_{(i)}\end{bmatrix} - \sqrt{\rho}\begin{bmatrix}\hat{\mathbf{X}}_{a(i-1)} \\ \hat{\mathbf{X}}_a(i)\end{bmatrix}\mathbf{h}\right\|^2\right\} \tag{19}$$

By invoking the block version of recursive least squares (RLS) algorithm to the cost function in (19) with the data vector of size 2×1 and the regressor matrix of dimension $2\times 2L$, we get [23]

$$M_{\hat{\mathcal{X}}_{(i)}} = M_{\hat{\mathcal{X}}_{(i-1)}} + \mathbf{e}_i^H \mathbf{\Gamma}_i \mathbf{e}_i \tag{20}$$

$$\hat{\mathbf{h}}_i = \hat{\mathbf{h}}_{i-1} + \mathbf{G}_i \mathbf{e}_i, \tag{21}$$

where

$$\mathbf{e}_i = \mathcal{Y}(i) - \sqrt{\rho}\,\hat{\mathbf{X}}_a(i)\hat{\mathbf{h}}_{i-1} \tag{22}$$

$$\mathbf{\Gamma}_i = \left[\mathbf{I}_2 + \rho\hat{\mathbf{X}}_a(i)\mathbf{P}_{i-1}\hat{\mathbf{X}}_a(i)^H\right]^{-1} \tag{23}$$

$$\mathbf{G}_i = \sqrt{\rho}\,\mathbf{P}_{i-1}\hat{\mathbf{X}}_a(i)^H\mathbf{\Gamma}_i \tag{24}$$

$$\mathbf{P}_i = \mathbf{P}_{i-1} - \mathbf{G}_i\mathbf{\Gamma}_i^{-1}\mathbf{G}_i^H \tag{25}$$

The RLS recursions are initialized by

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

Before introducing our algorithm, we first number the $|\Omega|^2$ combinations of the constellation points from two antennas by $1, 2, \ldots, |\Omega|^2$ and treat them as a big constellation set Ψ, where the kth $\left(1 \leq k \leq |\Omega|^2\right)$ vector constellation point is denoted by $\Psi(k)$. We then perform the depth-first search of signal tree for joint ML solution.

Algorithm 1 Blind MAP algorithm

Parameters: Initial search radius r, ρ and channel covariance matrix $\mathbf{R}_h$.
Inputs: $\mathcal{Y}$, modulation constellation Ω and the $1 \times N$ carrier index vector I.
Outputs: Estimated channel $\hat{\mathbf{h}}$ and data vector $\hat{\mathcal{X}}$.

1. *(Initialize) Set $i = 1$,$I(i) = 1$, $\hat{\mathcal{X}}(i) = \Psi\big(I(i)\big)$ and construct the Alamouti matrix $\hat{\mathbf{X}}_a(i)$.*
2. *(Compare with lower 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. *(Backtracking) Find the largest $1 \leq j \leq i$ such that $I(j) < |\Omega|^2$. If there exists such j, set $i = j$ and go to 5; else go to 6.*
4. *(Increment subcarrier) If $i < N$, set $i = i + 1, I(i) = 1, \hat{\mathcal{X}}(i) = \Psi\left(I(i)\right)$ and go to 2; else store the 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 $\hat{\mathcal{X}}(i) = \Psi\left(I(i)\right)$. 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 algorithm essentially reduces the search space of exhaustive ML search by performing a trimmed search over the signal tree of N layers, where each tree node at the ith layer corresponds to a specific partial sequence $\mathcal{X}_{(i)}$ and each tree node at the intermediate layer has $|\Omega|^2$ offsprings to the next layer.

The parameter ρ can be easily determined by estimating the noise variance, whereas for $\mathbf{R}_h$, our simulation results indicate that we can replace it with an identity matrix with almost no effect on the performance via carrier reordering (see next section). To obtain the initial guess of search radius, we can use the strategy described in [20] to determine r that would guarantee a MAP solution with very high probability. Nevertheless, the algorithm itself takes care of the value of r, in that if it is too small such that the algorithm is not able to backtrack, then it doubles the value of r; if it is too large such that the algorithm reaches the last subcarrier too quickly, then it reduces r to the most recent value of objective function (see steps 4 and 6). Therefore, any choice of r would guarantee the MAP solution.

The complexity of the algorithm is mainly attributed to the calculation of the bound $M_{\hat{\mathcal{X}}_{(i)}}$ (step 2) and the backtracking (step 3). The rest of the steps are simple additions and subtractions. From the RLS recursions, it can be seen that the calculation of the bound depends heavily on computation of $2L \times 2L$ matrix $\mathbf{P}_i$ in (25). In Section 4, we shall see how computation of $\mathbf{P}_i$ can be avoided by exploiting the structure of the FFT matrix, while in Section 5 we shall deal with the issue of backtracking.

4 Low-complexity blind algorithm
As mentioned earlier, the complexity of computing the metric $M_{\hat{\mathcal{X}}_{(i)}}$ depends heavily on computation of matrix $\mathbf{P}_i$. We show how we can completely avoid computing $\mathbf{P}_i$ and hence simply discard (25) from RLS recursions. This means that RLS algorithm will reduce to least mean square (LMS) in terms of complexity.

4.1 Reducing complexity by avoiding $\mathbf{P}_i$
Let us assume that $\mathbf{P}_1 = \mathbf{I}$ and that $\mathbf{a}_i$ are orthogonal for $i = 0, 1, 2, \cdots, N - 1$, i.e. $\mathbf{a}_i^H \mathbf{a}_j = 0$ for $i \neq j$ [b]. First, we merge (24) and (25) to get

$$\mathbf{P}_i = \mathbf{P}_{i-1} - \rho \mathbf{P}_{i-1} \hat{\mathbf{X}}_a^H(i) \Gamma_i^H \hat{\mathbf{X}}_a(i) \mathbf{P}_{i-1} \qquad (26)$$

Now using (26) and our assumptions, it follows by induction that $\mathbf{P}_i \hat{\mathbf{X}}_a^H(i+1) = \hat{\mathbf{X}}_a^H(i+1)$, $\mathbf{P}_i \hat{\mathbf{X}}_a^H(i+2) = \hat{\mathbf{X}}_a^H(i+2)$ and $\mathbf{P}_{i+1} \hat{\mathbf{X}}_a^H(i+2) = \hat{\mathbf{X}}_a^H(i+2)$. This means that if the successive regressors are orthogonal, we can simply replace $\mathbf{P}_i$ with an identity matrix and hence discard Equation 25. Moreover, from the orthogonality assumptions, it also follows that

$$\hat{\mathbf{X}}_a(i)\hat{\mathbf{X}}_a^H(j) = \begin{cases} 0 & \text{if } i \neq j \\ L\left(\|\hat{\mathbf{X}}_1(0)\|^2 + \|\hat{\mathbf{X}}_2(0)\|^2\right)\mathbf{I}_2 & \text{if } i = j \end{cases}$$
$$(27)$$

where $\mathbf{I}_n$ represents an $n \times n$ identity matrix. Incorporating these results into the RLS recursions, the matrices Γ_i and $\mathbf{G}_i$ become independent of $\mathbf{P_i}$ and are given by

$$\Gamma_i = \frac{1}{1 + \rho L \left(\|\hat{\mathcal{X}}_1(i)\|^2 + \|\hat{\mathcal{X}}_2(i)\|^2\right)} \mathbf{I}_2 \qquad (28)$$

$$\mathbf{G}_i = \sqrt{\rho}\,\hat{\mathbf{X}}_a(i)^H \Gamma_i \qquad (29)$$

The resulting low-complexity blind algorithm based on (20) to (22), (28) and (29) for metric computation, requires no matrix inversion or computation of $\mathbf{P}_i$.

### 4.2 Reducing complexity by carrier reordering
In the above approximation, we assumed that $\mathbf{P}_1 = \mathbf{I}$ and $\mathbf{a}_i$ are orthogonal which allows us to use (27). However, $\mathbf{a}_i$ are columns of the partial FFT matrix $\mathbf{A}_i$, so strictly speaking, they are not orthogonal. Hence, the successive regressor matrices would not be orthogonal, too. We will show that we can make them orthogonal or semi-orthogonal by carrier reordering. To understand the idea, we compute and plot the magnitude of autocorrelation of these partial vectors given by

$$|\mathbf{a}_i^H \mathbf{a}_l| = \begin{cases} L & \text{if } i = l \\ \frac{1}{L}\left|\frac{\sin(\pi(i-1)L/N)}{\sin(\pi(i-1)/N)}\right| & \text{if } i \neq l \end{cases} \qquad (30)$$

in Figure 2 for $N = 16$ and $L = 4$ and where we set $l = 1$. It can be seen that columns 1, 5, 9, and 13 are orthogonal to each other and so are the columns 2, 6, 10, 14 and so on. If we visit the subcarriers in the order $1, 5, 9, 13, 2, 6, 10, 14, \cdots, 4, 8, 12, 16$, we find that consecutive vectors will be orthogonal or approximately orthogonal. In general, it is found that with $\triangle = N/L$, the vectors $\mathbf{a}_i, \mathbf{a}_{i+\triangle}, \mathbf{a}_{i+2\triangle}, \forall i$ are approximately orthogonal. Therefore, by simple reordering the carriers, we can achieve orthogonality among different subcarriers and use that fact to reduce the complexity of our algorithm as done in section 4.1.

5 Complexity reduction by using reliable carriers
Despite avoiding $\mathbf{P}_i$ and the carrier reordering approach, it has been observed that complexity of algorithm becomes very large in multiple-antenna systems as compared to single-antenna systems especially in the low SNR regime. The major source of complexity is attributed to backtracking of the proposed algorithm. Backtracking occurs whenever the constraint $M_{\hat{\mathcal{X}}_{(i)}} < r$ is not satisfied. The algorithm then goes back either to the nearest subcarrier or to the current subcarrier whose alphabet is not exhausted and increments the alphabet (step 3 then step 5). This issue is rigorously analysed in [20] where it is shown that the probability of backtracking is almost zero at higher SNR; however, no solution is proposed to deal with it in the lower SNR case. The fact is that backtracking is
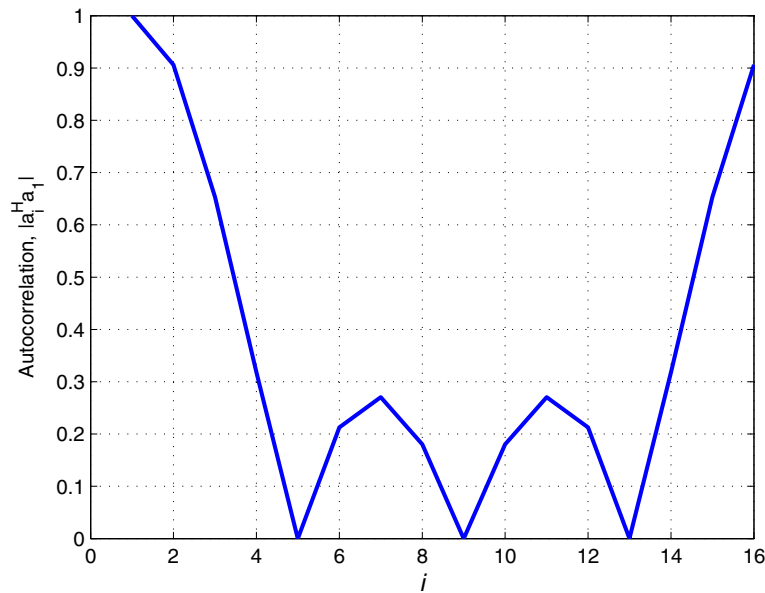
Figure 2 Autocorrelation as a function of *i* for *N* = 16 and *L* = 4.

inevitable in all blind search algorithms. Although it cannot be avoided, we propose to minimize its effect by using the concept of reliable carriers. The previous methods of reducing the complexity have the following drawbacks:

1. The proposed solutions are still very complex as they do not take into account the issue of backtracking which is a major source of complexity. They can be considered to have low complexity only in the high SNR regime, where we essentially get rid of backtracking.
2. The proposed solutions do not work in low SNR regime and becomes infeasible for multiple-antenna systems. This is due to the fact that the search space at each node grows as $|\Omega|^{N_t}$ as compared to $|\Omega|$ for the SISO system, where $|\Omega|$ is the alphabet size and N_t is the number of transmit antennas. Thus, the complexity of the proposed algorithms due to backtracking ultimately dominates the complexity induced by computing the matrix $\mathbf{P}_i$ (or its inverse) and becomes the real bottleneck.
3. The proposed methods do not make use of the fact that pilots are usually present in real systems to aid in channel estimation and that the channel is usually slowly varying.

We can make use of the third point to our advantage. Specifically, the presence of some pilots and slow variation in the channel allows us to get a tentative estimate of the data. If we are able to arrange the data according to its reliability, starting with the most reliable data, then there would be a less chance that we need to backtrack. Since

earlier data is reliable, there is no need to backtrack for this part. The later data might not be reliable, but by the time we start processing this data, the algorithm would have been already converged. However, measuring data reliability requires tentative channel/data estimates. Thus, the blind algorithm can be turned into a semi-blind algorithm to reliably track the channel along with the data detection. The semi-blind algorithm would require a short training sequence of L symbols only at the start of transmission to get a tentative estimate of the data and its reliability and no further pilots or channel statistics would be required.

5.1 Measuring the reliability
For measuring the reliability of data carriers, we borrow the idea presented in [21] by the author of the current paper, where it was used in the context of non-linear distortion mitigation in OFDM. To minimize backtracking, the algorithm must devise a procedure to identify the reliable subcarriers from the tentative estimates of channel and the data. With receiver having an estimate of channel, the decoding process can be accomplished by rewriting (10) into the form as shown below:

$$\widetilde{\mathcal{Y}} = \sqrt{\rho}\, \mathbf{H}_a \mathcal{X} + \widetilde{\mathcal{N}}, \tag{31}$$

where $\widetilde{\mathcal{Y}} = \left[\left[\mathcal{Y}^{(k)}\right]^T \left[\mathcal{Y}^{*(k+1)}\right]^T\right]^T, \widetilde{\mathcal{N}} = \left[\left[\mathcal{N}^{(k)}\right]^T \left[\mathcal{N}^{*(k+1)}\right]^T\right]^T$ and $\mathbf{H}_a$ is an *Alamouti-like* matrix defined as

$$\mathbf{H}_a \triangleq \begin{bmatrix} \Lambda_1 & \Lambda_2 \\ \Lambda_2^* & -\Lambda_1^* \end{bmatrix} \tag{32}$$

By left multiplying both sides of (31) with $\frac{1}{\sqrt{\rho}}\mathbf{H}_a^{-1}$ and rearranging the terms, we get:

$$\hat{\boldsymbol{\mathcal{X}}} = \boldsymbol{\mathcal{X}} + \mathbf{D} \tag{33}$$

with the difference vector; $\mathbf{D} \triangleq \frac{1}{\sqrt{\rho}}\mathbf{H}_a^{-1}\widetilde{\mathcal{N}}$. The imperfect knowledge of the channel results in an estimation error $\Delta\mathbf{H}_a$ and consequently the vector $\mathbf{D}$ represents the distortion due to channel estimation error and the effect of additive noise.

To assess the reliability, consider a data carrier $\hat{\boldsymbol{\mathcal{X}}}(k)$ (in scalar case) and its nearest constellation point $\langle\hat{\boldsymbol{\mathcal{X}}}(k)\rangle$. Treating channel estimation error as noise, ML-based decoding would yield $\boldsymbol{\mathcal{X}}(k)$ by mapping $\hat{\boldsymbol{\mathcal{X}}}(k)$ to the nearest constellation point $\langle\hat{\boldsymbol{\mathcal{X}}}(k)\rangle$. Such a scheme would be very efficient at higher SNR if distortion was only due to AWGN. However, in our case, we have an additional perturbation due to channel estimation error that is independent of SNR, and therefore, we expect that part of data samples would be severely effected by the channel distortion and fall outside their actual decision regions. Clearly, there is a need to assess and identify these unreliable data coefficients for our algorithm to reduce the backtracking.

Authors in [21] have developed a rigorous method for assessing the reliability of estimated data coefficients. Intuitively, for the data carrier $\hat{\boldsymbol{\mathcal{X}}}(k)$, we can measure its reliability based on relative posterior probability that the difference term $D(k)$ equals $\hat{\boldsymbol{\mathcal{X}}}(k) - \langle\hat{\boldsymbol{\mathcal{X}}}(k)\rangle$ to the probability that it equals some other vector $\hat{\boldsymbol{\mathcal{X}}}(k) - \Omega_m | \Omega_m \neq \langle\hat{\boldsymbol{\mathcal{X}}}(k)\rangle$. For example,

$$
\begin{aligned}
\mathfrak{R}(k) &= \log \frac{Pr\left(\langle\hat{\boldsymbol{\mathcal{X}}}(k)\rangle = \boldsymbol{\mathcal{X}}(k) \mid \hat{\boldsymbol{\mathcal{X}}}(k)\right)}{Pr\left(\langle\hat{\boldsymbol{\mathcal{X}}}(k)\rangle = \boldsymbol{\mathcal{X}}_{NN}(k) \mid \hat{\boldsymbol{\mathcal{X}}}(k)\right)} \\
&= \log \frac{Pr\left(D(k) = \hat{\boldsymbol{\mathcal{X}}}(k) - \langle\hat{\boldsymbol{\mathcal{X}}}(k)\rangle\right)}{Pr\left(D(k) = \hat{\boldsymbol{\mathcal{X}}}(k) - \boldsymbol{\mathcal{X}}_{NN}(k)\right)}
\end{aligned} \tag{34}
$$

defines the reliability in decoding $\hat{\boldsymbol{\mathcal{X}}}(k)$ to the closest constellation point relative to decoding to the nearest neighbour $\boldsymbol{\mathcal{X}}_{NN}(k)$. Figure 3 illustrates this concept, such that for instance even though $\hat{\boldsymbol{\mathcal{X}}}(1)$ and $\hat{\boldsymbol{\mathcal{X}}}(2)$ have the same distance from $\langle\hat{\boldsymbol{\mathcal{X}}}(1)\rangle = \langle\hat{\boldsymbol{\mathcal{X}}}(2)\rangle$, $\hat{\boldsymbol{\mathcal{X}}}(2)$ has a higher reliability than $\hat{\boldsymbol{\mathcal{X}}}(1)$ as it is farther from nearest neighbour. This suggests the dependence of phase $\theta_{\hat{\boldsymbol{\mathcal{X}}}(k) - \langle\hat{\boldsymbol{\mathcal{X}}}(k)\rangle}$ on the reliability in addition to the magnitude $|\hat{\boldsymbol{\mathcal{X}}}(k) - \langle\hat{\boldsymbol{\mathcal{X}}}(k)\rangle|$. In essence, the reliability of a measurement at each tone is a function that maps a triple (magnitude, phase and channel gain) into $\mathbb{R}_{[1,\infty]}$. The exact expression for reliability
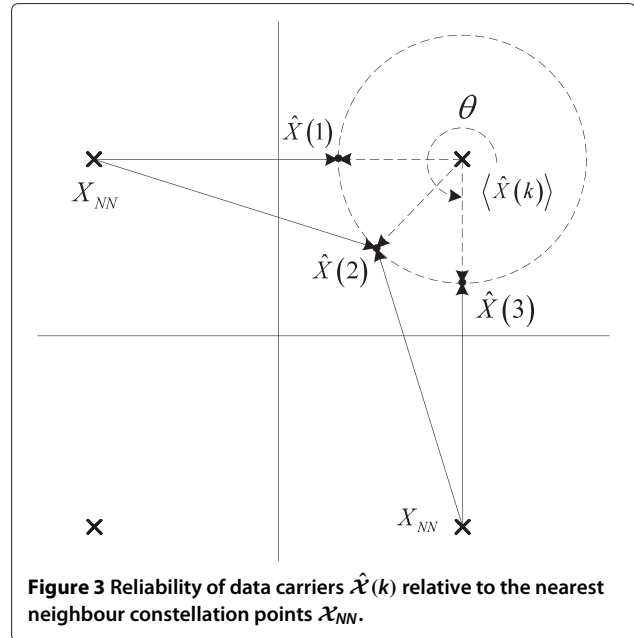
Figure 3 Reliability of data carriers $\hat{\boldsymbol{\mathcal{X}}}(k)$ relative to the nearest neighbour constellation points $\boldsymbol{\mathcal{X}}_{NN}$.

is the generalization of (34) to the vector-wise likelihood ratio defined as

$$\mathfrak{R}^{\text{exact}} = \log \frac{f_D\left(\hat{\boldsymbol{\mathcal{X}}} - \langle\hat{\boldsymbol{\mathcal{X}}}\rangle\right)}{\sum_{m=0,\Omega_m \neq \langle\hat{\boldsymbol{\mathcal{X}}}\rangle}^{|\Omega|-1} f_D\left(\hat{\boldsymbol{\mathcal{X}}} - \Omega_m\right)}, \tag{35}$$

where $f_D(.)$ is the *pdf* of the distortion vector $\mathbf{D}$, which by definition, can be easily seen to be Gaussian circularly symmetric with variance $\sigma_D^2 = \frac{1}{\rho}\left(\mathbf{Ha}^{-1}\right)\left(\mathbf{Ha}^{-1}\right)^H$. The above computation for exact reliability is however inefficient since it would require $O(N|\Omega|)$ evaluations of $f_D(.)$ which grows with constellation size $|\Omega|$. Alternately, the geometric-based approximations for assessing reliability as derived in [21] may be employed with marginal loss in the performance. Ultimately, once the vector $\mathfrak{R}$ is computed, we can proceed to select the most reliable data tones. These reliable data tones can then be supplied to our algorithm for initial search of the ML solution. Based on the above developments, we now introduce the semi-blind algorithm.

5.2 The semi-blind algorithm

1. Obtain an initial estimate of the channel vector $\mathbf{h}$ from L training/pilot symbols at start of transmission, then repeat the following steps over two consecutive time instants.
2. Predict and decode the carriers $\hat{\boldsymbol{\mathcal{X}}}$ from previous channel estimate $\hat{\mathbf{h}}$ and observation vector $\widetilde{\boldsymbol{\mathcal{Y}}}$.
3. Use (35) to compute reliability of data carriers, $\mathfrak{R}(k)_{k=1}^P$, with $1 \leq P \leq N$ and rearrange them in decreasing order of their reliability. The parameter P

represents the total number of reordered carriers used by the algorithm.

4. Run the blind algorithm proposed previously, starting with most reliable data first, to obtain the exact ML estimates of the channel and the data.

Remarks: The first three steps of the semi-blind algorithm serve as pre-processing steps tailored to minimizing the backtracking of blind algorithm in step 4. One can easily obtain the channel estimates from pilots in step 1. The prediction step 2 is trivial and would suffer only little distortion as the channel does not change much in slow fading. To initiate the RLS recursions of blind algorithm in step 4, we initialize the channel vector with its previous estimate and set $\mathbf{P}_i = \mathbf{I}$; thus, no channel statistics are required *a priori*. The blind algorithm is supplied with reordered carriers with most reliable data to start over the search (step 3), so that the there will be almost no backtracking during the convergence of the algorithm. However, we observe that the carrier reordering based on reliability measures does not ensure orthogonality of successive regressors; therefore, the low complexity variants of RLS introduced earlier cannot be employed.

6 Exploiting the structure of Alamouti coding

In our previous approaches to reducing the complexity of blind algorithm, we did not use the Alamouti coding structure, which can simplify the receiver design by linearly processing the received data samples. We will show that by exploiting the orthogonal structure of Alamouti coding, we can either achieve (a) The complexity reduction of blind algorithm for MISO OFDM system to that of SISO OFDM system of [20] with little effect on estimation performance, or (b) improve the estimation performance of our algorithm with full complexity of the MISO OFDM. To achieve the first goal, we start by multiplying (31) by $\mathbf{H}_a^H$ to get

$$\begin{aligned}\mathbf{z} = \mathbf{H}_a^H \widetilde{\mathcal{Y}} &= \sqrt{\rho}\, \mathbf{H}_a^H \mathbf{H}_a \mathcal{X} + \mathbf{H}_a^H \widetilde{\mathcal{N}} \\ &= \sqrt{\rho}\, \mathbf{B}\mathcal{X} + \mathbf{w},\end{aligned} \tag{36}$$

where $\mathbf{z}$ is the new observation vector, $\mathbf{w}$ is the noise vector, and the block diagonal matrix $\mathbf{B}$ is defined as

$$\mathbf{B} = \mathbf{I}_2 \otimes \left(\Lambda_1^* \Lambda_1 + \Lambda_2^* \Lambda_2\right), \tag{37}$$

where $\otimes$ represents the Kronecker product. Thus, owing to the orthogonal structure of Alamouti, the system in (36) is decoupled into two SISO OFDM systems as

$$\mathbf{z}_1 = \sqrt{\rho}\, \widetilde{\mathbf{B}}\mathcal{X}_1^{(k)} + \mathbf{w}_1 \tag{38}$$

$$\mathbf{z}_2 = \sqrt{\rho}\, \widetilde{\mathbf{B}}\mathcal{X}_2^{(k)} + \mathbf{w}_2, \tag{39}$$

where $\widetilde{\mathbf{B}} = \left(\Lambda_1^* \Lambda_1 + \Lambda_2^* \Lambda_2\right)$. The joint ML solutions for SISO systems (42) would involve only $2|\Omega|^N$ hypoth-

esis tests as opposed to $|\Omega|^{2N}$ for the original system, which implies huge reduction in complexity. However, this comes at the cost of estimation performance as number of unknown channel parameters in (42) are increased to $2L$ as opposed to L is the SISO system. In a similar fashion, the estimation performance of the blind algorithm can be improved by multiplying (10) by $\mathbf{X}_a^H$ as follows:

$$\begin{aligned}\bar{\mathbf{z}} = \mathbf{X}_a^H \mathcal{Y} &= \sqrt{\rho}\, \mathbf{X}_a^H \mathbf{X}_a \mathbf{h} + \mathbf{X}_a^H \mathcal{N} \\ &= \sqrt{\rho}\, \overline{\mathbf{B}}\mathbf{h} + \overline{\mathbf{w}},\end{aligned} \tag{40}$$

where $\bar{\mathbf{z}}$ is the new observation vector, $\overline{\mathbf{w}}$ is the noise vector, and the block diagonal matrix $\overline{\mathbf{B}}$ is

$$\overline{\mathbf{B}} = \mathbf{I}_2 \otimes \left(\mathbf{A}\mathrm{diag}\left(|\mathcal{X}_1^{(k)}|^2 + |\mathcal{X}_2^{(k)}|^2\right)\mathbf{A}^H\right) \tag{41}$$

We can see that the system in (40) can also be decoupled as

$$\bar{\mathbf{z}}_1 = \sqrt{\rho}\, \widetilde{\overline{\mathbf{B}}}\mathbf{h}_1 + \overline{\mathbf{w}}_1 \tag{42}$$

$$\bar{\mathbf{z}}_2 = \sqrt{\rho}\, \widetilde{\overline{\mathbf{B}}}\mathbf{h}_2 + \overline{\mathbf{w}}_2, \tag{43}$$

where $\widetilde{\overline{\mathbf{B}}} = \left(\mathbf{A}\mathrm{diag}\left(|\mathcal{X}_1^{(k)}|^2 + |\mathcal{X}_2^{(k)}|^2\right)\mathbf{A}^H\right)$. It can be seen that the above equations correspond to the SISO systems for which the joint ML solution can be derived easily. It is worth mentioning that the search space for blind ML is not reduced due to dependence of observations and the matrix $\overline{\mathbf{B}}$ on $\mathcal{X}_1^{(k)}$ and $\mathcal{X}_2^{(k)}$. However, the number of channel parameters is reduced to half, which results in an improved estimation performance. In short, the Alamouti structure enables us to either improve upon the estimation performance or the complexity reduction but not the both.

7 Simulation results

For simulations, we first consider OFDM system with $N = 16$ subcarriers and channel length $L = 4$ for each transmit-receive channel and CP length of at least $L-1$. For blind algorithm, both channels are independent Rayleigh fading, assumed stationary over two consecutive OFDM blocks and each having an exponential power decay profile, i.e. $E\left\{|h_i(\tau)|^2\right\} = e^{-0.2\tau}$. Information symbols are modulated using binary phase shift keying (BPSK) or quadrature amplitude modulation (QAM).

In Figure 4, we plot the results for $N = 16$, BPSK data symbols using perfectly known channel and our exact blind algorithm, together with low-complexity variants, i.e. blind algorithm with (a) $\mathbf{P}_i = \mathbf{I}$ and (b) $\mathbf{P}_i = \mathbf{I}$ with subcarrier reordering. In the first case, the performance degrades and bit error rate (BER) reaches an error floor. However, with subcarrier reordering approach, we almost get the same performance as that of exact blind algorithm
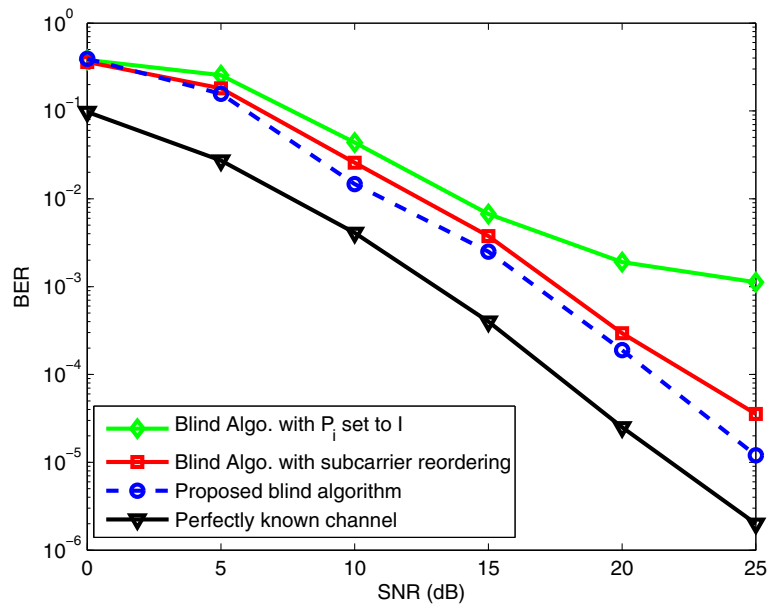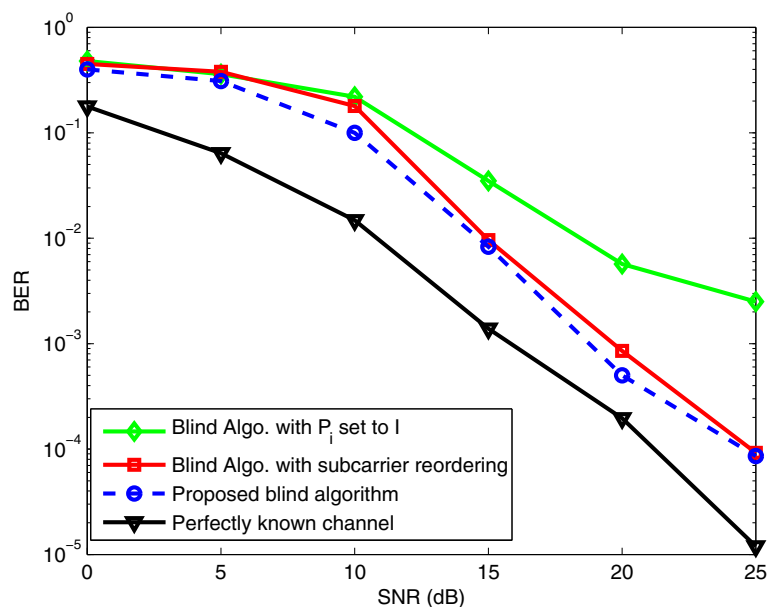
Figure 4 BER vs SNR for Alamouti-coded BPSK OFDM over Rayleigh fading channel with *N* = 16 and *L* = 4.

without requiring the channel statistics. A similar trend is observed in Figure 5, when 4-QAM signal modulation is considered.

For semi-blind algorithm, we adopt the AR(1) process to model the slow rayleigh fading channels, where the channel weight vector varies as [23]: $\mathbf{h}(n) = \alpha\mathbf{h}(n-1) + \mathbf{q}(n)$ and where $\alpha = J_0(2\pi f_d T_s)$ and $\mathbf{q}$ is a complex normal vector with covariance matrix $(1 - \alpha^2)\mathbf{I}$. The product of

maximum Doppler frequency f_d and sampling time T_s referred to as normalized Doppler frequency F_D, controls the amount of variations in the channel statistics. Two different values of normalized Doppler frequency, i.e., 0.1 and 0.001, are considered in the simulations. Results for semi-blind algorithm are presented against perfect coherent detection in Figures 6 and 7 for BPSK and 4−QAM modulations respectively which show favourable perfor-

Figure 5 BER vs SNR for Alamouti-coded 4-QAM OFDM over Rayleigh fading channel with *N* = 16 and *L* = 4.
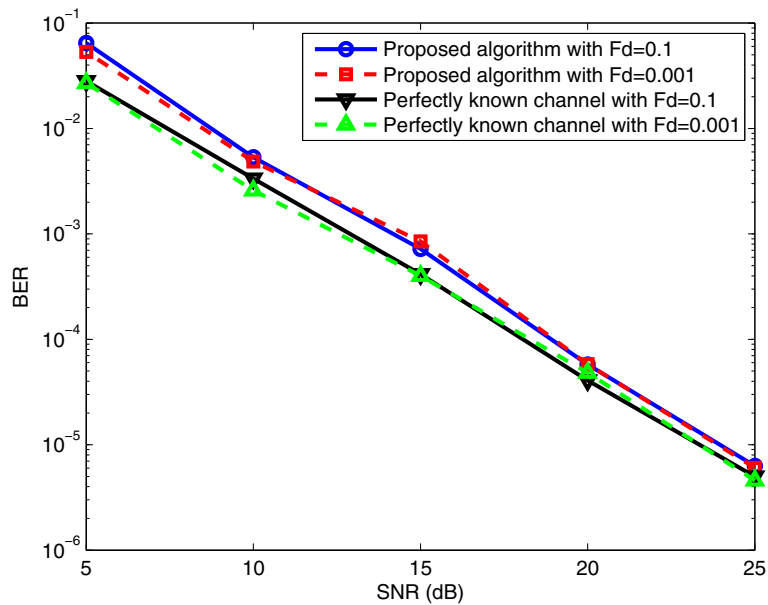
Figure 6 BER vs SNR for MISO-OFDM with BPSK modulation over Rayleigh fading channel with $N = 32$ and $L = 4$.

mance of algorithm under different fading conditions. For higher modulations and number of subcarriers, we use the SISO OFDM systems obtained after decoupling by Alamouti structure and results are presented in Figure 8 for $16-$QAM constellation with $N = 64$.

To assess the computational complexity of proposed algorithm, we compare the average number of nodes visited by the algorithm with various reliability measures

in Figure 9. It is clearly observed that proposed reliability scheme offers significantly lower complexity at lower SNR values. At higher SNR, the complexity is constant, confirming the fact that there is almost no backtracking. Figure 10 shows that the performance for various degrees of reliability measures is almost identical which means that computational advantages are attained without degrading the performance. Through simulations, it
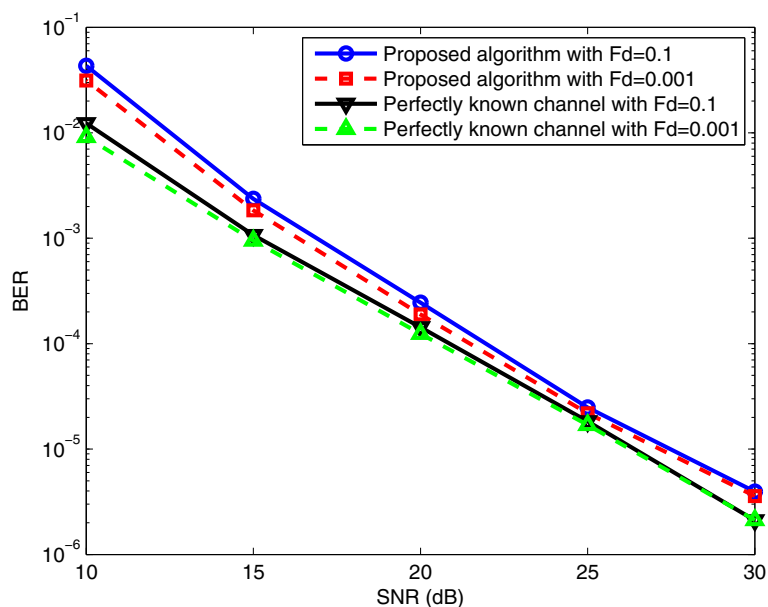
Figure 7 BER vs SNR for MISO-OFDM with 4-QAM modulation over Rayleigh fading channel with $N = 32$ and $L = 4$.
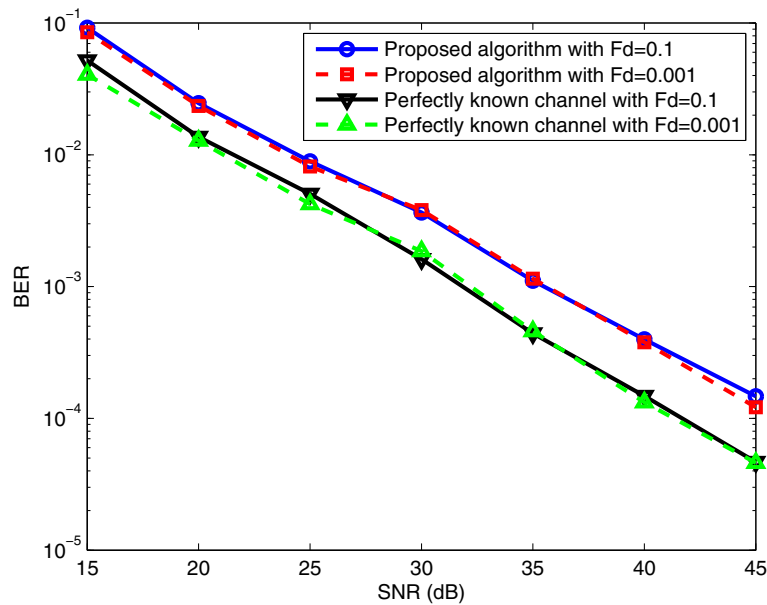
Figure 8 BER vs SNR for SISO-OFDM with 16-QAM modulation over Rayleigh fading channel with $N = 64$ **and** $L = 4$.

has also been observed that the reliability of around 50% to 60% is enough for a good performance, although more importantly, the algorithm does not disfavour the usage of more reliable carriers.

Finally, in Figure 11, the complexity of comparison for different modulation schemes such as BPSK, 4-QAM and 16-QAM is presented which clearly indicates the computational advantages of proposed reliability-based method.

8 Conclusions

In this paper, we presented a blind ML algorithm for joint channel estimation and data detection in OFDM wireless systems using Alamouti STBC coding. The simulation results show favourable performance of algorithm. As evident from simulations, our low-complexity blind algorithm performs equally well as exact blind algorithm. Moreover, the new algorithm does not need any prior
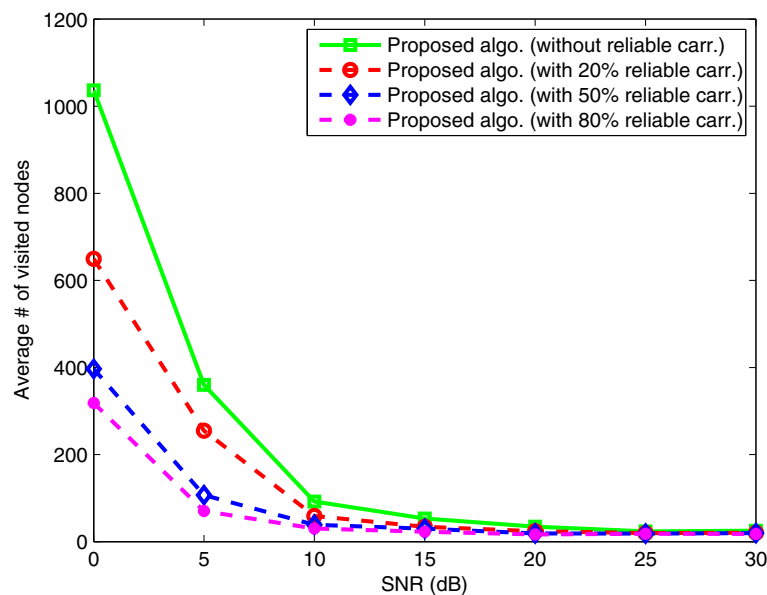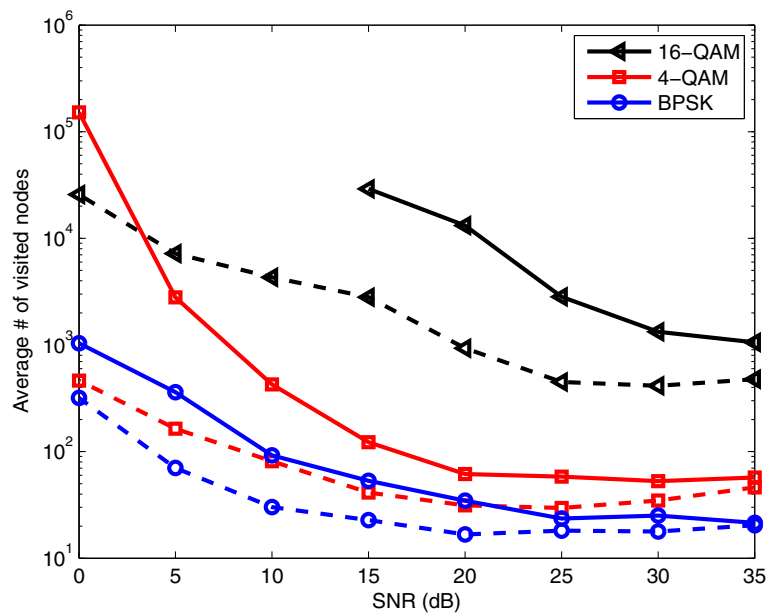

Figure 9 Complexity of proposed semi-blind algorithm with various degrees of reliability measurements. BPSK modulation with $N = 16$ and $L = 4$.
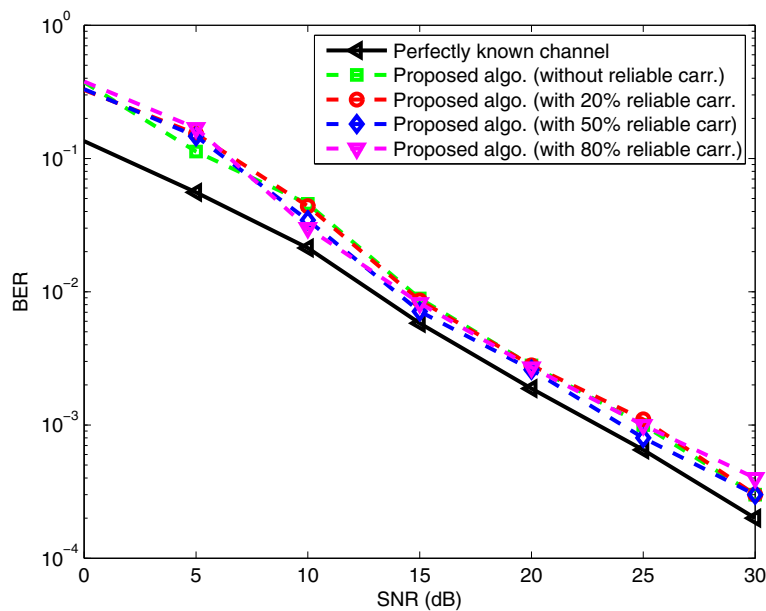
Figure 10 Performance of the proposed algorithm with various degrees of reliability measurements. BPSK modulation with $N = 16$ and $L = 4$.

information about channel statistics as it avoids calculating matrix $\mathbf{P}_i$ with subcarrier reordering. Another major source of complexity in blind algorithm is the issue of backtracking which becomes more prominent in the low-SNR regime. This issue was not tackled in the previous studies. We proposed a semi-blind algorithm which minimizes the probability of backtracking by supplying the

blind algorithm with reordered subcarriers based on their reliability computations using a sophisticated reliability criterion. By minimizing the backtracking, significant improvement is achieved in terms of complexity without compromising the performance. Moreover, the orthogonal structure of Alamouti coding was also exploited to yield complexity reduction or improved estimation

Figure 11 Complexity of the proposed semi-blind algorithm for different modulations with $N = 16$ and $L = 4$. Proposed algorithm without (solid lines) and with reliable carriers (dashed lines).

performance. The reduced complexity variant turns out to be very handy when large number of subcarriers and/or large size constellation modulations are employed.

Endnotes

[a]This lemma was proved in [20] for SISO case; we simply extend it here to the multi-antenna case.

[b]In fact, a weaker condition that three consecutive vectors $\mathbf{a}_i$, $\mathbf{a}_{i+1}$ and $\mathbf{a}_{i+2}$ are orthogonal would suffice.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Electrical Engineering, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia. [2]Department of Electrical Engineering, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia.

References

1. LJ Cimini, Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing. IEEE Trans. Commun. **33**(7), 665–675 (1985)
2. H Bölcskei, *Principles of MIMO-OFDM Wireless Systems,* Chapter in *CRC Handbook on Signal Processing for Mobile Communications, M. Ibnkahla, Ed*, (Ch. 12 2004)
3. BP Crow, I Widjaja, JG Kim, PT Sakai, IEEE 802.11 wireless local area networks. IEEE Commun. Mag. **35**(9), 116–126 (1997)
4. I Koffman, V Roman, Broadband wireless access solutions based on OFDM access in IEEE 802.16. IEEE Commun. Mag. **40**(4), 96–103 (2002)
5. 3GPP TR 36.211, Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation (ver. 10.5.0, 2012)
6. E Telatar, Capacity of multi-antenna gaussian channels. Eur. Trans. Telecomm. **10**(6), 585–595 (1999)
7. S Alamouti, A simple transmit diversity technique for wireless communications. IEEE J. Selected Areas Commun. **16**(8), 1451–1458 (1998)
8. P Stoica, G Ganesan, Space-time block codes: trained, blind and semi-blind detection. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 2 (IEEE Orlando, FL, USA, 13–17 May 2002). pp. 1609–1612
9. AL Swindlehurst, G Leus, Blind and semi-blind equalization for generalized space-time block codes. IEEE Trans. Signal Process. **50**(10), 2489–2498 (2002)
10. B Muquet, M De Courville, P Duhamel, Subspace-based blind and semi-blind channel estimation for OFDM systems. IEEE Trans. Signal Process. **50**(7), 1699–1712 (2002)
11. C-C Tu, B Champagne, Subspace-based blind channel estimation for MIMO-OFDM systems with reduced time averaging. IEEE Trans. Veh. Tech. **59**(3), 1539–1544 (2010)
12. H Bolcskei, RW Heath, AJ Paulraj, Blind equalization in OFDM-based multi-antenna systems. *The IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (AS-SPCC)* (IEEE, Lake Louise, Alta, 01–04 Oct 2000), pp. 58–63
13. J Choi, C-C Lim, A Cholesky factorization based approach for blind FIR channel identification. IEEE Trans. Signal Process. **56**(4), 1730–1735 (2008)
14. SA Banani, RG Vaughan, OFDM with iterative blind channel estimation. IEEE Trans. Veh. Tech. **59**(9), 4298–4308 (2010)
15. T Cui, C Tellambura, Joint data detection and channel estimation for OFDM systems. IEEE Trans. Commun. **54**(4), 670–679 (2006)
16. EG Larsson, P Stoica, J Li, On maximum-likelihood detection and decoding for space-time coding systems. IEEE Trans. Signal Process. **50**(4), 937–944 (2002)
17. T-H Chang, W-K Ma, C-Y Chi, Maximum-likelihood detection of orthogonal space-time block coded OFDM in unknown block fading channels. IEEE Trans. Signal Process. **56**(4), 1637–1649 (2008)
18. TY Al-Naffouri, AA Quadeer, Cyclic prefix based enhanced data recovery in OFDM. IEEE Trans. Signal Process. **58**(6), 3406–3410 (2010)
19. W Xu, M Stojnic, B Hassibi, Low-complexity blind maximum-likelihood detection for SIMO systems with general constellations. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, Las Vegas, NV, USA, 31 Mar–4 Apr 2008), pp. 2817–2820
20. TY Al-Naffouri, AA Dahman, MS Sohail, W Xu, B Hassibi, Low-complexity blind equalization for OFDM systems with general constellations. IEEE Trans. Signal Process. **60**(12), 6395–6407 (2012)
21. EB Al-Safadi, TY Al-Naffouri, Pilotless recovery of clipped OFDM signals by compressive sensing over reliable data carriers, in *IEEE 13th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)* (IEEE Cesme, 17–20 June 2012), pp. 580–584
22. N Al-Dhahir, Single-carrier frequency-domain equalization for space-time block-coded transmissions over frequency-selective fading channels. IEEE Commun. Lett. **5**(7), 304–306 (2001)
23. AH Sayed, *Fundamentals of Adaptive Filtering*. (Wiley, New York, 2003)