CLOSED-FORM BLIND DECODING OF ORTHOGONAL SPACE-TIME BLOCK CODES

Shahram Shahbazpanahi[*] *Alex B. Gershman*[*] *Jonathan H. Manton*[†]

[*]Dept. Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada
[†]Dept. of Electrical and Electronic Engineering, University of Melbourne, Australia

ABSTRACT

A new computationally simple approach to blind decoding of orthogonal space-time block codes (STBCs) is proposed. Our approach estimates the channel matrix in a closed form and uses this estimate in the maximum likelihood (ML) receiver to decode the symbols. It exploits specific properties of the orthogonal STBCs and is free of major drawbacks of other blind space-time decoding schemes.

1. INTRODUCTION

Orthogonal STBCs (OSTBCs) [1]-[2] represent an attractive class of space-time codes because they enjoy full diversity gain and low decoding complexity. If the channel state information (CSI) is available at the receiver, the optimal ML decoder for this class of codes is a simple linear receiver followed by the symbol-by-symbol detector. Although training approaches can be used to obtain the CSI at the receiver, the price that has to be paid is the reduced bandwidth efficiency and (potentially) inaccurate channel estimates which may result in a severe degradation of the decoding performance. This has been a strong motivation to develop blind space-time decoders [3]-[7].

In this paper, we present a novel computationally efficient approach to blind decoding of OSTBCs. Our approach is based on estimating the channel matrix in a closed form and using this estimate in the ML decoder. It exploits specific properties of OSTBCs and is free of major drawbacks of other blind receivers. In particular, there is no SNR penalty as in the algorithms of [3]-[6] provided that the coherence time of the channel is sufficiently large. Furthermore, unlike the blind approach of [7], our technique is applicable to the majority of full-rate OSTBCs. Moreover, it finds the channel estimate in a closed form and, therefore, it does not have global convergence problems that iterative techniques may have.

2. BACKGROUND

Let us assume a MIMO system with N transmit and M receive antennas and flat block-fading channel. Then, we can use the familiar model [2]

$$\mathbf{Y} = \mathbf{X}\mathbf{H} + \mathbf{V} \tag{1}$$

where $\mathbf{Y} \triangleq [\mathbf{y}^T(1)\ \cdots\ \mathbf{y}^T(T)]^T$, $\mathbf{X} \triangleq [\mathbf{x}^T(1)\ \cdots\ \mathbf{x}^T(T)]^T$, and $\mathbf{V} \triangleq [\mathbf{v}^T(1)\ \cdots\ \mathbf{v}^T(T)]^T$ are the matrices of the received signals, transmitted signals, and noise, respectively, $\mathbf{H}$ is the $N \times M$ complex channel matrix, T is the block length, and $(\cdot)^T$ denotes the transpose. Here, $\mathbf{y}(t) = [y_1(t)\ \cdots\ y_M(t)]$, $\mathbf{x}(t) = [x_1(t)\ \cdots\ x_N(t)]$, and $\mathbf{v}(t) = [v_1(t)\ \cdots\ v_M(t)]$ are the complex row vectors of the received signal, transmitted signal, and noise, respectively. We assume that the noise is both spatially and temporally white with the variance of σ^2 per complex dimension. The slow fading channel case is considered, i.e., the channel coherence time is assumed to be substantially larger than the data block length T.

We denote the complex information symbols prior to space-time encoding as $s_1, \ldots, s_K$ and assume that they are zero-mean random variables drawn from (possibly different) constellations $\mathcal{U}_k$, $k = 1, 2, \ldots, K$. Let us introduce the vector $\mathbf{s} \triangleq [s_1\ \cdots\ s_K]^T$. Note that $\mathbf{s} \in \mathcal{S}$ where $\mathcal{S} = \{\mathbf{s}^{(1)}, \cdots, \mathbf{s}^{(L)}\}$ is the set of all possible symbol vectors and L is the cardinality of this set. The $T \times N$ matrix $\mathbf{X}(\mathbf{s})$ is called an OSTBC if [2]: (i) all elements of $\mathbf{X}(\mathbf{s})$ are linear functions of the K complex variables $s_1, s_2, \ldots, s_K$ and their complex conjugates; (ii) for any $\mathbf{s}$, this matrix satisfies $\mathbf{X}^H(\mathbf{s})\mathbf{X}(\mathbf{s}) = \|\mathbf{s}\|^2 \mathbf{I}_N$ where $\mathbf{I}_N$ is the $N \times N$ identity matrix and $\|\cdot\|$ denotes the Euclidean norm.

It can be readily verified [8] that the matrix $\mathbf{X}(\mathbf{s})$ can be written as

$$\mathbf{X}(\mathbf{s}) = \sum_{k=1}^{K} \left(\mathbf{C}_k \mathrm{Re}\{s_k\} + \mathbf{D}_k \mathrm{Im}\{s_k\}\right) \tag{2}$$

where the matrices $\mathbf{C}_k$ and $\mathbf{D}_k$ are defined as $\mathbf{C}_k \triangleq \mathbf{X}(\mathbf{e}_k)$ and $\mathbf{D}_k \triangleq \mathbf{X}(j\mathbf{e}_k)$, respectively. Here, $j = \sqrt{-1}$ and $\mathbf{e}_k$ is a $K \times 1$ vector having one in the kth position and zeros elsewhere. Using (2), one can rewrite the model (1) as [8]

$$\underline{\mathbf{Y}} = \mathbf{A}(\mathbf{H})\underline{\mathbf{s}} + \underline{\mathbf{V}} \tag{3}$$

where the "underline" operator for any matrix $\mathbf{P}$ is defined as

$$\underline{\mathbf{P}} = [(\mathrm{vec}\{\mathrm{Re}(\mathbf{P})\})^T,\ (\mathrm{vec}\{\mathrm{Im}(\mathbf{P})\})^T]^T \tag{4}$$

and $\mathrm{vec}\{\cdot\}$ is the vectorization operator stacking all columns of a matrix on top of each other. The $2MT \times 2K$ real matrix $\mathbf{A}(\mathbf{H})$ in (3) is given by $\mathbf{A}(\mathbf{H}) = [\underline{\mathbf{C}_1\mathbf{H}} \cdots \underline{\mathbf{C}_K\mathbf{H}}\ \underline{\mathbf{D}_1\mathbf{H}} \cdots \underline{\mathbf{D}_K\mathbf{H}}]$. It is easy to verify that, regardless of the value of the channel matrix $\mathbf{H}$, the columns of $\mathbf{A}(\mathbf{H})$ have the same norms and are orthogonal to each other [8]:

$$\mathbf{A}^T(\mathbf{H})\mathbf{A}(\mathbf{H}) = \|\mathbf{H}\|_F^2\, \mathbf{I}_{2K} \tag{5}$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

If $\mathbf{H}$ is known at the receiver, the optimal (ML) decoder uses this knowledge to obtain $l_{\mathrm{opt}} = \arg\min_{l \in \{1,\ldots,L\}} \|\mathbf{Y} - \mathbf{Y}^{(l)}\|_F$ and then exploits l_{opt} to decode the data symbols [2]. Here, $\mathbf{Y}^{(l)}$ is the noise-free received data matrix that corresponds to the vector of information symbols $\mathbf{s}^{(l)}$.

The ML decoder can also be viewed as the MF receiver whose output SNR is maximized [9]. In the OSTBC case, the following equivalent MF interpretation of the ML decoder has been derived in [8]. First of all, the estimate

$$\hat{\underline{\mathbf{s}}} = \frac{1}{\|\mathbf{H}\|_F^2}\, \mathbf{A}^T(\mathbf{H})\underline{\mathbf{Y}} \tag{6}$$

of $\underline{\mathbf{s}}$ is obtained and then the estimate $\hat{\mathbf{s}} = [\mathbf{I}_K \ j\mathbf{I}_K]\,\hat{\underline{\mathbf{s}}}$ of the vector $\mathbf{s}$ is computed. The kth element of $\hat{\mathbf{s}}$ is compared with all points in $\mathcal{U}_k$ and the closest point in this constellation is introduced as an estimate of the kth entry of $\mathbf{s}$. This procedure is repeated for all $k = 1, 2, \ldots, K$ (i.e., the decoding is done symbol-by-symbol).

3. BLIND SPACE-TIME DECODING

Let us introduce the $2MN \times 1$ *channel vector* $\mathbf{h} \triangleq \underline{\mathbf{H}}$ and, with a small abuse of notation, let us use hereafter $\mathbf{A}(\mathbf{h})$ instead of $\mathbf{A}(\mathbf{H})$. Then, (5) can be rewritten as

$$\mathbf{A}^T(\mathbf{h})\mathbf{A}(\mathbf{h}) = \|\mathbf{h}\|^2\mathbf{I}_{2K} \tag{7}$$

As $\mathbf{A}(\mathbf{h})$ is linear in $\mathbf{h}$, there exists a unique $4KMT \times 2MN$ matrix $\mathbf{\Phi}$ such that

$$\mathrm{vec}\{\mathbf{A}(\mathbf{h})\} = \mathbf{\Phi}\mathbf{h} \tag{8}$$

The kth column of $\mathbf{\Phi}$ can be written as $[\mathbf{\Phi}]_k = \mathrm{vec}\{\mathbf{A}(\mathbf{e}_k)\}$ where the dimension of $\mathbf{e}_k$ is now $2MN \times 1$. Using (3), we have that the covariance matrix of the real-valued vectorized data $\underline{\mathbf{Y}}$ is given by

$$\mathbf{R} \triangleq \mathrm{E}\{\underline{\mathbf{Y}}\,\underline{\mathbf{Y}}^T\} = \mathbf{A}(\mathbf{h})\mathbf{\Lambda}_s\mathbf{A}^T(\mathbf{h}) + (\sigma^2/2)\mathbf{I}_{2MT} \tag{9}$$

where $\mathbf{\Lambda}_s \triangleq \mathrm{E}\{\underline{\mathbf{s}}\,\underline{\mathbf{s}}^T\}$ is the covariance matrix of the real vector $\underline{\mathbf{s}}$. Since the elements of $\underline{\mathbf{s}}$ can be assumed uncorrelated with each other, $\mathbf{\Lambda}_s$ is a diagonal matrix. Each diagonal element of $\mathbf{\Lambda}_s$ represents the average power of the real or imaginary part of the corresponding data symbol and depends only on the shape of the constellation of that particular symbol. Hence, the matrix $\mathbf{\Lambda}_s$ is known at the receiver.

Multiplying (9) from the right by $\mathbf{A}(\mathbf{h})/\|\mathbf{h}\|$ and using (7), we obtain that

$$\mathbf{R}\mathbf{A}(\mathbf{h})/\|\mathbf{h}\| = (\mathbf{A}(\mathbf{h})/\|\mathbf{h}\|)\left[\|\mathbf{h}\|^2\mathbf{\Lambda}_s + (\sigma^2/2)\mathbf{I}_{2K}\right] \tag{10}$$

Since $\mathbf{A}(\mathbf{h})/\|\mathbf{h}\|$ has orthonormal columns and $\|\mathbf{h}\|^2\mathbf{\Lambda}_s + \frac{\sigma^2}{2}\mathbf{I}_{2K}$ is a diagonal matrix, (10) can be viewed as the characteristic equation for the data covariance matrix $\mathbf{R}$. This means that the diagonal elements of the matrix

$$\mathbf{\Lambda} \triangleq \|\mathbf{h}\|^2\mathbf{\Lambda}_s + (\sigma^2/2)\mathbf{I}_{2K} \tag{11}$$

are the $2K$ largest eigenvalues of the matrix $\mathbf{R}$ while the columns of $\mathbf{A}(\mathbf{h})/\|\mathbf{h}\|$ are the corresponding (normalized) eigenvectors.

Lemma 1: Let $\mathbf{Q}$ be an $m \times q$ real matrix where $q < m$. Then, for any $m \times m$ real symmetric matrix $\mathbf{P}$ the solution to the following optimization problem

$$\max_{\mathbf{Q}} \mathrm{tr}\{\mathbf{Q}^T\mathbf{P}\mathbf{Q}\} \quad \text{subject to} \quad \mathbf{Q}^T\mathbf{Q} = \mathbf{I}_q \tag{12}$$

is given by any matrix $\mathbf{Q}_*$ whose column space is the same as the subspace spanned by the q principal eigenvectors[1] of $\mathbf{P}$ and, for any such $\mathbf{Q}_*$,

$$\mathrm{tr}\{\mathbf{Q}_*^T\mathbf{P}\mathbf{Q}_*\} = \sum_{i=1}^{q}\nu_q \tag{13}$$

where ν_i, $i = 1, \ldots, q$ are the q largest eigenvalues of $\mathbf{P}$.

Proof: See [10]. □

[1] i.e., the eigenvectors that correspond to the q largest eigenvalues.

Replacing $\mathbf{P}$ by $\mathbf{R}$ and setting $q = 2K$ in (12), we obtain from Lemma 1 that the solution to the following optimization problem

$$\max_{\mathbf{Q}} \mathrm{tr}\{\mathbf{Q}^T\mathbf{R}\mathbf{Q}\} \quad \text{subject to} \quad \mathbf{Q}^T\mathbf{Q} = \mathbf{I}_{2K} \tag{14}$$

will be given by any matrix $\mathbf{Q}_*$ which satisfies

$$\mathrm{range}\{\mathbf{Q}_*\} = \mathrm{range}\{\mathbf{A}(\mathbf{h})\} \tag{15}$$

Moreover, from Lemma 1 and the established properties of $\mathbf{R}$, we have

$$\mathrm{tr}\{\mathbf{Q}_*^T\mathbf{R}\mathbf{Q}_*\} = \mathrm{tr}\{\mathbf{\Lambda}\} \tag{16}$$

Let us replace $\mathbf{Q}$ in (14) by $\mathbf{A}(\tilde{\mathbf{h}})/\|\tilde{\mathbf{h}}\|$ where $\tilde{\mathbf{h}}$ is a vector of optimization variables. In this case, the constraint in (14) becomes $\mathbf{A}^T(\tilde{\mathbf{h}})\mathbf{A}(\tilde{\mathbf{h}})/\|\tilde{\mathbf{h}}\|^2 = \mathbf{I}_{2K}$. According to (7), this constraint is satisfied for any $\tilde{\mathbf{h}}$ and, therefore, is redundant. Omitting it, we obtain the following unconstrained optimization problem:

$$\max_{\tilde{\mathbf{h}}} \mathrm{tr}\{\mathbf{A}^T(\tilde{\mathbf{h}})\mathbf{R}\mathbf{A}(\tilde{\mathbf{h}})\}/\|\tilde{\mathbf{h}}\|^2 \tag{17}$$

It is important that the problems (14) and (17) are not equivalent to each other because the matrix $\mathbf{A}(\tilde{\mathbf{h}})$ in (17) has a particular structure while the matrix $\mathbf{Q}$ in (14) is unstructured. Therefore, the sets of solutions to (14) and (17) for the matrices $\mathbf{Q}$ and $\mathbf{A}(\tilde{\mathbf{h}})$, respectively, may be different. Moreover, this means that the maximum value of the objective function in (17) cannot exceed the maximum value of the objective function in (14).

Inserting (9) into the objective function (17) and using (7), we obtain that for $\tilde{\mathbf{h}} = \mathbf{h}$,

$$\mathrm{tr}\{\mathbf{A}^T(\tilde{\mathbf{h}})\mathbf{R}\mathbf{A}(\tilde{\mathbf{h}})\}/\|\tilde{\mathbf{h}}\|^2\big|_{\tilde{\mathbf{h}}=\mathbf{h}} = \mathrm{tr}\{\mathbf{\Lambda}\} \tag{18}$$

Comparing (16) and (18), we obtain that the maxima of the objective functions in (14) and (17) coincide and, therefore, in terms of $\mathbf{Q}$ and $\mathbf{A}(\tilde{\mathbf{h}})$ the set of all possible solutions to (17) represents a *subset* of the set of all possible solutions to (14). Using this property and (15), we obtain that for any $\tilde{\mathbf{h}}$ achieving the maximum in (17),

$$\mathrm{range}\{\mathbf{A}(\tilde{\mathbf{h}})\} = \mathrm{range}\{\mathbf{A}(\mathbf{h})\} \tag{19}$$

At this point, we make a conjecture[2] that for most OSTBCs, (19) holds true if and only if $\tilde{\mathbf{h}} = \gamma\mathbf{h}$ where γ is a real value.

Let our conjecture hold true. In this case, if any value of $\tilde{\mathbf{h}}$ is obtained for which $\mathbf{A}(\tilde{\mathbf{h}})$ is a solution to (17) then the channel vector $\mathbf{h}$ can be obtained from this value of $\tilde{\mathbf{h}}$ up to a real scalar.

Now, let us find a closed-form solution to (17). To simplify the cost function in (17), we can express its numerator as

$$\mathrm{tr}\{\mathbf{A}^T(\tilde{\mathbf{h}})\mathbf{R}\mathbf{A}(\tilde{\mathbf{h}})\} = (\mathrm{vec}\{\mathbf{A}(\tilde{\mathbf{h}})\})^T(\mathbf{I}_{2K} \otimes \mathbf{R})\mathrm{vec}\{\mathbf{A}(\tilde{\mathbf{h}})\} \tag{20}$$

Using (8), we can write

$$\mathrm{vec}\{\mathbf{A}(\tilde{\mathbf{h}})\} = \mathbf{\Phi}\tilde{\mathbf{h}} \tag{21}$$

Inserting (21) into (20), we can rewrite (20) as

$$\mathrm{tr}\{\mathbf{A}^T(\tilde{\mathbf{h}})\mathbf{R}\mathbf{A}(\tilde{\mathbf{h}})\} = \tilde{\mathbf{h}}^T\mathbf{\Phi}^T(\mathbf{I}_{2K} \otimes \mathbf{R})\mathbf{\Phi}\tilde{\mathbf{h}} \tag{22}$$

and the optimization problem (17) can be expressed as

$$\max_{\tilde{\mathbf{h}}} \tilde{\mathbf{h}}^T\mathbf{\Phi}^T(\mathbf{I}_{2K} \otimes \mathbf{R})\mathbf{\Phi}\tilde{\mathbf{h}}/\|\tilde{\mathbf{h}}\|^2 \tag{23}$$

[2] We will verify this conjecture numerically and demonstrate that it holds true for most of OSTBCs with a few exceptions.

Note that all solutions to (23) belong to the subspace spanned by the n principal eigenvectors of the matrix $\mathbf{\Phi}^T(\mathbf{I}_{2K}\otimes\mathbf{R})\mathbf{\Phi}$ where n is the multiplicity order of its largest eigenvalue. However, assuming that our conjecture holds true, we have that this subspace is one-dimensional and, therefore, the maximal eigenvalue of $\mathbf{\Phi}^T(\mathbf{I}_{2K}\otimes\mathbf{R})\mathbf{\Phi}$ has the multiplicity order equal to one. Hence, ignoring the scaling ambiguity, the normalized solution to (23) can be written as

$$\tilde{\mathbf{h}}_{\mathrm{opt}} = \mathcal{P}\{\mathbf{\Phi}^T(\mathbf{I}_{2K}\otimes\mathbf{R})\mathbf{\Phi}\} \tag{24}$$

where $\mathcal{P}\{\cdot\}$ denotes the normalized principal eigenvector of a matrix ($\|\mathcal{P}\{\cdot\}\| = 1$). Therefore, the true channel vector $\mathbf{h}$ can be written as

$$\mathbf{h} = \|\mathbf{h}\|\mathcal{P}\{\mathbf{\Phi}^T(\mathbf{I}_{2K}\otimes\mathbf{R})\mathbf{\Phi}\} \tag{25}$$

In what follows, we propose a simple method to determine $\|\mathbf{h}\|$. Using Lemma 1 along with (11), we have that

$$\max_{\tilde{\mathbf{h}}} \mathrm{tr}\{\mathbf{A}^T(\tilde{\mathbf{h}})\mathbf{R}\mathbf{A}(\tilde{\mathbf{h}})\}/\|\tilde{\mathbf{h}}\|^2 = \|\mathbf{h}\|^2\mathrm{tr}\{\mathbf{\Lambda}_s\} + K\sigma^2 \tag{26}$$

Using (24) and (26), the value of $\|\mathbf{h}\|$ can be obtained as

$$\|\mathbf{h}\| = \sqrt{(\mathrm{tr}\{\mathbf{A}^T(\tilde{\mathbf{h}}_{\mathrm{opt}})\mathbf{R}\mathbf{A}(\tilde{\mathbf{h}}_{\mathrm{opt}})\} - K\sigma^2)/\mathrm{tr}\{\mathbf{\Lambda}_s\}} \tag{27}$$

Using (25) along with (27), we obtain the true channel vector $\mathbf{h}$ in a closed form as

$$\mathbf{h} = \sqrt{(\mathrm{tr}\{\mathbf{A}^T(\tilde{\mathbf{h}}_{\mathrm{opt}})\mathbf{R}\mathbf{A}(\tilde{\mathbf{h}}_{\mathrm{opt}})\} - K\sigma^2)/\mathrm{tr}\{\mathbf{\Lambda}_s\}}\ \tilde{\mathbf{h}}_{\mathrm{opt}} \tag{28}$$

where $\tilde{\mathbf{h}}_{\mathrm{opt}}$ is given by (24).

In practice, the true covariance matrix $\mathbf{R}$ is unavailable. Instead of it, the sample covariance matrix

$$\hat{\mathbf{R}} = \frac{1}{J}\sum_{i=1}^{J}\underline{\mathbf{Y}_i}\,\underline{\mathbf{Y}_i}^T = \frac{1}{J}\mathbb{Y}\mathbb{Y}^T \tag{29}$$

can be used. Here, $\mathbb{Y} \triangleq [\underline{\mathbf{Y}_1}\ \cdots\ \underline{\mathbf{Y}_J}]$ is the matrix of the received data, $\mathbf{Y}_i$ is the ith data block, and J is the number of data blocks used.

Replacing the true covariance matrix $\mathbf{R}$ by the sample covariance matrix $\hat{\mathbf{R}}$ in (28), we obtain the following consistent closed-form blind channel estimate:

$$\hat{\mathbf{h}} = \sqrt{(\mathrm{tr}\{\mathbf{A}^T(\hat{\tilde{\mathbf{h}}}_{\mathrm{opt}})\hat{\mathbf{R}}\mathbf{A}(\hat{\tilde{\mathbf{h}}}_{\mathrm{opt}})\} - K\sigma^2)/\mathrm{tr}\{\mathbf{\Lambda}_s\}}\ \hat{\tilde{\mathbf{h}}}_{\mathrm{opt}} \tag{30}$$

where the estimate

$$\hat{\tilde{\mathbf{h}}}_{\mathrm{opt}} = \mathcal{P}\{\mathbf{\Phi}^T(\mathbf{I}_{2K}\otimes\hat{\mathbf{R}})\mathbf{\Phi}\} \tag{31}$$

is obtained by replacing $\mathbf{R}$ by $\hat{\mathbf{R}}$ in (24).

It is important to stress that if *constant modulus constellations* are used for each of the encoded symbols, the symbol-by-symbol detector is not sensitive to multiplication of the channel estimate by any real constant. This fact follows from the structure of the linear receiver (6) and linearity of $\mathbf{A}(\mathbf{h})$ in $\mathbf{h}$. Therefore, in the constant modulus case, the scalar square root factor in (30) can be ignored, i.e., the value of $\hat{\tilde{\mathbf{h}}}_{\mathrm{opt}}$ can be directly used as the channel estimate.

Our joint blind channel estimation and symbol detection algorithm can be summarized as follows:

1. Obtain $\hat{\mathbf{R}}$ from J consecutive received data vectors using (29). If the symbol constellations are not constant modulus, compute the estimate $\hat{\mathbf{h}}$ of the channel vector $\mathbf{h}$ using (30) and (31). Otherwise, use $\hat{\mathbf{h}} = \hat{\tilde{\mathbf{h}}}_{\mathrm{opt}}$ of (31) as the channel estimate.
2. Use the so-obtained $\hat{\mathbf{h}}$ in the MF receiver to obtain $\hat{\underline{\hat{\mathbf{s}}}} = \frac{1}{\|\hat{\mathbf{h}}\|_F^2}\mathbf{A}^T(\hat{\mathbf{h}})\underline{\mathbf{Y}}$ and $\hat{\hat{\mathbf{s}}} = [\mathbf{I}_K\ j\mathbf{I}_K]\hat{\underline{\hat{\mathbf{s}}}}$ where the "double-hat" symbol stresses that the signal detection is blind.
3. Decode the kth symbol ($k = 1, 2, \ldots, K$) as a point of the constellation $\mathcal{U}_k$ which is closest to the kth entry of $\hat{\hat{\mathbf{s}}}$.

If constant modulus constellations are used for each of the encoded symbols, the knowledge of σ^2 is not required in our algorithm. However, this knowledge is required if the symbol constellations are not constant modulus. In practice, the noise power can be estimated by averaging the $2MT - 2K$ smallest eigenvalues of the sample covariance matrix $\hat{\mathbf{R}}$.

The algorithm proposed is suitable for batch implementation. However, its computationally efficient on-line implementation can be also derived using subspace tracking techniques.

4. RELATIONSHIP TO THE RELAXED BLIND ML ESTIMATOR

Let the ith ($i = 1, 2, \cdots, J$) received data vector $\underline{\mathbf{Y}_i}$ be generated according to the real data model (3), so that

$$\underline{\mathbf{Y}_i} = \mathbf{A}(\mathbf{h})\underline{\mathbf{s}_i} + \underline{\mathbf{V}_i} \tag{32}$$

In the blind (non-coherent) ML detector, the channel vector $\mathbf{h}$ and the transmitted symbols $\underline{\mathbf{s}_i}$ ($i = 1, 2, \ldots, J$) should be treated as unknown deterministic parameters and their ML estimates can be obtained by maximizing the log-likelihood (LL) function as

$$\{\hat{\mathbf{h}}_{\mathrm{ML}}, \hat{\mathbf{S}}_{\mathrm{ML}}\} = \arg\max_{\mathbf{h}}\max_{\mathbf{S}\in\Omega}\log f(\underline{\mathbf{Y}_1}, \underline{\mathbf{Y}_2}, \ldots, \underline{\mathbf{Y}_J}\,|\,\mathbf{S}, \mathbf{h}) \tag{33}$$

where $\mathbf{S} \triangleq [\underline{\mathbf{s}_1}\ \underline{\mathbf{s}_2}\ \cdots\ \underline{\mathbf{s}_J}]$ and Ω is the set of all possible values of $\mathbf{S}$. The problem (33) has an exponential complexity but can be simplified by relaxing the finite alphabet constraint $\mathbf{S}\in\Omega$. Assuming that the noise vectors $\underline{\mathbf{V}_i}$ ($i = 1, 2, \ldots, J$) are zero-mean independent Gaussian random vectors with the covariance matrix $(\sigma^2/2)\mathbf{I}$, we have that the conditional pdf of $\underline{\mathbf{Y}_i}$ is Gaussian, that is, $f(\underline{\mathbf{Y}_i}\,|\,\underline{\mathbf{s}_i}, \mathbf{h}) = (\pi\sigma^2)^{-MT}e^{-\|\underline{\mathbf{Y}_i}-\mathbf{A}(\mathbf{h})\underline{\mathbf{s}_i}\|^2/\sigma^2}$. Since all $\underline{\mathbf{Y}_i}$ ($i = 1, \ldots, J$) are independent random vectors, we have $f(\underline{\mathbf{Y}_1}, \underline{\mathbf{Y}_2}, \ldots, \underline{\mathbf{Y}_J}\,|\,\mathbf{S}, \mathbf{h}) = \prod_{i=1}^{J} f(\underline{\mathbf{Y}_i}\,|\,\underline{\mathbf{s}_i}, \mathbf{h})$. Hence, we can write the relaxed blind ML decoder as

$$\{\hat{\mathbf{h}}_{\mathrm{RML}}, \hat{\mathbf{S}}_{\mathrm{RML}}\} = \arg\min_{\mathbf{h}}\min_{\mathbf{S}}\sum_{i=1}^{J}\|\underline{\mathbf{Y}_i} - \mathbf{A}(\mathbf{h})\underline{\mathbf{s}_i}\|^2 \tag{34}$$

In (34), the ith term of the summation is minimized for $\underline{\hat{\mathbf{s}}_i}_{\mathrm{RML}} = \mathbf{A}^T(\mathbf{h})\underline{\mathbf{Y}_i}/\|\mathbf{h}\|^2$. Using this fact, (34) can be concentrated with respect to $\mathbf{S}$ and after straightforward manipulations, we obtain

$$\begin{aligned}
\hat{\mathbf{h}}_{\mathrm{RML}} &= \arg\min_{\mathbf{h}}\sum_{i=1}^{J}\left\|\underline{\mathbf{Y}_i} - \mathbf{A}(\mathbf{h})\mathbf{A}^T(\mathbf{h})\underline{\mathbf{Y}_i}/\|\mathbf{h}\|^2\right\|^2 \\
&= \arg\min_{\mathbf{h}}\left\|\mathbb{Y} - \mathbf{A}(\mathbf{h})\mathbf{A}^T(\mathbf{h})\mathbb{Y}/\|\mathbf{h}\|^2\right\|_F^2 \\
&= \arg\max_{\mathbf{h}}\mathrm{tr}\{\mathbf{A}^T(\mathbf{h})\mathbb{Y}\mathbb{Y}^T\mathbf{A}(\mathbf{h})\}/\|\mathbf{h}\|^2 \\
&= \arg\max_{\mathbf{h}}\mathrm{tr}\{\mathbf{A}^T(\mathbf{h})\hat{\mathbf{R}}\mathbf{A}(\mathbf{h})\}/\|\mathbf{h}\|^2
\end{aligned} \tag{35}$$

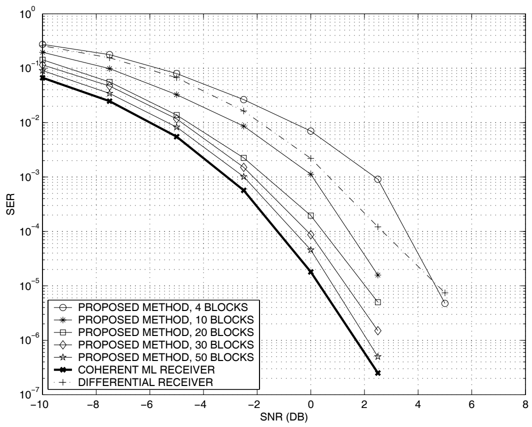

Figure 1: SERs versus SNR. First example.

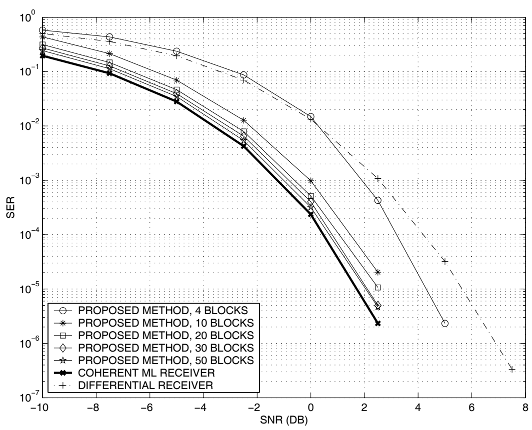

Figure 2: SERs versus SNR. Second example.

Comparing (35) with (17), we see that these two problems are identical. Therefore, in the Gaussian case our blind channel estimator (30) can be viewed as the blind ML estimator in which the finite alphabet constraint $\mathbf{S} \in \Omega$ is relaxed.

5. SIMULATION RESULTS

Through simulations, we tested numerous generalized orthogonal design and amicable design-based STBCs with different parameters and rates. We found out that if $M > 1$ then our conjecture is valid for all codes tested except the generalized orthogonal design based STBCs with $N = K$ (including Alamouti code).

We illustrate the performance of our blind decoder with two examples. In the first example, we have chosen the full-rate OSTBC with $N = 3$, $M = 4$, and $K = T = 4$ (this code is given by eqn. (27) of [2]). The BPSK signals are used in this example.

In the second example, we have tested the $3/4$ OSTBC with $N = M = T = 4$ and $K = 3$ (this code is given by eqn. (7.4.10) of [5]). The QPSK signals are used in this example.

Figures 1 and 2 display the symbol error rates (SERs) versus the SNR for the first and second examples, respectively. The SNR is defined in the way similar to [7]. In both figures, our blind space-time decoder with different numbers of blocks J is compared with the differential modulation approach and the coherent ML decoder. Note that, unlike the proposed and differential decoding schemes tested, the coherent ML decoder exploits the exact knowledge of the CSI at the receiver and is included in our figures only to illustrate performance losses of the blind techniques with respect to the ideal (informed receiver) case. Two different types of differential approaches have been used: in the first example, the technique of [4] has been chosen, while in the second example, the technique of [5] has been employed.

From the figures, it follows that in both examples our blind decoder performs substantially better than the differential approach if $J \geq 10$. With $J \geq 30$, the performance of the proposed decoder is very close to that of the coherent ML decoder.

6. CONCLUSIONS

A new computationally efficient closed-form approach to blind decoding of orthogonal space-time block codes has been proposed. It is applicable to most of the OSTBCs and is free of common shortcomings (such as substantial performance penalties, high computational costs, scalar ambiguities, need of pilot symbols, etc.) of the other known blind space-time decoding schemes.

7. REFERENCES

[1] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas in Communications,* vol. 45, pp. 1451-1458, Oct. 1998.

[2] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory,* vol. 45, pp. 1456-1467, July 1999.

[3] B. M. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. Communications,* vol. 48, pp. 2041-2052, Dec. 2000.

[4] H. Jafarkhani and V. Tarokh, "Multiple transmit antenna differential detection from generalized orthogonal design," *IEEE Trans. Inf. Theory,* vol. 47, pp. 2626-2631, Sept. 2001.

[5] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications,* Cambridge University Press, Cambridge, UK, 2003.

[6] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inf. Theory,* vol. 46, pp. 543-564, March 2000.

[7] A. L. Swindlehurst and G. Leus, "Blind and semi-blind equalization for generalized space-time block codes," *IEEE Trans. Signal Processing,* vol. 50, pp. 2489-2498, Oct. 2002.

[8] M. Gharavi-Alkhansari and A. B. Gershman, "Constellation space invariance of space-time block codes with application to optimal ML decoding," in *Proc. IEEE VTC'03,* Orlando, FL, Oct. 2003.

[9] G. Ganesan and P. Stoica, "Space-time block codes: a maximum SNR approach," *IEEE Trans. Inf. Theory,* vol. 47, pp. 1650-1656, May 2001.

[10] J. H. Manton, "Optimization algorithms exploiting unitary constraints," *IEEE Trans. Signal Processing*, vol. 50, pp. 635-650, March 2002.