TECHNICAL UNIVERSITY OF CRETE
DEPARTMENT OF ELECTRONIC AND COMPUTER ENGINEERING

Maximum-Likelihood Noncoherent MPSK OSTBC Detection With Polynomial Complexity

Submitted in partial fulfilment of the requirements
for the Electronic & Computer Engineering diploma degree

Author:
PANAGIOTIS P. MARKOPOULOS

Supervisor:
ASSISTANT PROFESSOR
GEORGE N. KARYSTINOS

OCTOBER, 2010

I dedicate my thesis to my beloved parents, Christine and Praxitelis, who have always been an inexhaustible source of encouragement and inspiration to me.

Acknowledgements

First and foremost, I offer my sincere gratitude to my supervisor, Assistant Professor George Karystinos, who gave me the opportunity to work on such an intriguing and innovative subject. Had it not been for his guidance and patient support, from the initial to the final level, this project would never have been accomplished.
In addition, I am grateful to Professor Athanasios Liavas and to Assistant Professor Aggelos Bletsas for all the knowledge and inspiration they have imparted to me.
Moreover, I would like to take this opportunity to thank Dimitris Papailiopoulos and Anastasios Kyrillidis for their overall assistance as well as for setting up a basis on which I could work and develop new ideas.
Finally, I want to thank all my friends and my family for keeping my spirits up throughout this project.

Contents

List of Figures

I. Introduction

A. Multiple Input-Multiple Output (MIMO) Systems & STBCs

High data rate achievement and reliable communications assurance have been principal targets of modern wireless standards. These features can be afforded by multiple antenna systems that are proven to attain higher channel capacity than single antenna setups while lowering the error probability. Elaborate informationtheoretic results tailored to Rayleigh fading prove, that channel capacity actually grows linearly when the number of receive and transmit antennas increases. However, it is beyond question that antenna arrays are costly and space demanding, thus being a more plausible setup at base stations rather than remote terminals. Consequently, primary focus has been set on transmit diversity techniques, with the first work coming from Siavash Alamouti that delivered the first full-diversity, full-rate space-time block code (STBC) for two transmit antennas.

B. Orthogonal Space-Time Block Coding

Tarokh, Jafarkhani, and Calderbank generalized the design to more than two transmit antennas introducing a paradigm for the construction of space-time block codes based on orthogonal designs. The so-called orthogonal STBCs (OSTBCs) are proven to achieve full antenna diversity gain with linear-complexity single-symbol maximum-likelihood (ML) coherent detection. Moreover, OSTBCs outperform non-orthogonal designs in terms of error rate. Rate-one full-diversity OSTBCs error-rate provides a lower bound on the one of quasi-orthogonal STBCs due to lack of intersymbol interference (ISI). Such an error rate is attainable with linear complexity, if the channel state information (CSI) is available at the receiver.

However, the nature of wireless channels suggests rapidly varying channel conditions that render channel estimation inadequate and inefficient. Even when the fading channel coefficients are not fast varying, channel estimation requires transmission of long pilot symbol sequences, especially for the cases where large antenna arrays are used, with the direct implication of reduced transmission rate. Interestingly, the ergodic capacity promised by multiple antenna systems is attained even when CSI is not available to either transmitter or receiver. The work of Zheng and Tse shows that when CSI is not available the capacity of multiantenna systems with full CSI knowledge at the receiver under Rayleigh fading is approached at the high signal-to-noise ratio (SNR) regime, if one transmits equal-energy symbols and utilizes space-time codes that are mutually orthogonal during each coherence time interval. Certainly, when OSTBCs are used and the receiver has no CSI, ML noncoherent sequence detection has to be performed on the entire coherence interval for optimal performance. However, if sequence detection is performed through exhaustive search among all possible data sequences then exponential computational complexity is required.

C. ML Noncoherent Detection with Exponential Complexity

The problem of ML noncoherent OSTBC detection under static independent and identically distributed (i.i.d.) Rayleigh fading was originally expressed as a trace maximization and later proven to also take the form of a binary quadratic form maximization problem that in the general case is NP-hard. In [17] it was shown that the ML noncoherent OSTBC detection problem can be solved optimally by the sphere decoder, certainly an exponential expected complexity approach for any fixed

SNR. To avoid the exponential complexity of the optimal receiver many suboptimal schemes have been proposed in the literature. Moreover, the use of pilot symbols has already been proposed to combat the exponential cost of the optimal noncoherent receiver, at the expense of information rate. Lately, D. S. Papailiopoulos and G. N. Karystinos in [31] proved that when the transmitted symbols belong to a BPSK constellation, polynomial-time ML detection complexity is always achievable for static Rayleigh correlated (in general) channels.

D. Maximum-Likelihood Noncoherent Detection in Polynomial Time

In this work, we consider the case of static Rayleigh channels, consisted of independent and identically distributed coefficients, prove that ML noncoherent MPSK OSTBC detection can be performed in polynomial time and define the constraints under which this is achievable. We note that the polynomial complexity order is completely determined by the product of the number of antennas used at the transmitter and the receiver.

II. System Model and Problem Statement

We consider a MIMO system with M_t transmit and M_r receive antennas that employs orthogonal space-time coded transmission of size $M_t \times T$ and rate $R = \frac{N}{T}, N \leq T$. We assume transmission of M-ary phase-shift keying (MPSK) data symbols, each one selected from an M-ary alphabet $\mathsf{A}_M \triangleq \{e^{j\frac{\pi}{M}(2m+1)} | m = 0, 1, \ldots, M-1\}$, that are split into vectors of N elements. Each vector forms a corresponding space-time block matrix of size $M_t \times T$. The $M_t \times T$ space time block $\mathbf{C}(\mathbf{s}) \in \mathbb{C}^{M_t \times T}$ that corresponds to the $N \times 1$ data vector $\mathbf{s} \in {\mathsf{A}_M}^N$ is given by

$$\mathbf{C}(\mathbf{s}) = \sum_{n=1}^{N} (\mathbf{A}_n \Re\{s_n\} + j\mathbf{B}_n \Im\{s_n\}) \tag{1}$$

where s_n denotes the nth element of $\mathbf{s}$, $n = 1, 2, \ldots, N$ and $\{\mathbf{A}_n, \mathbf{B}_n\}$is an amicable orthogonal design of fixed (in general complex-valued) code matrices of dimension $M_t \times T$ that satisfies the property

$$\mathbf{C}(\mathbf{s})\mathbf{C}^H(\mathbf{s}) = \|\mathbf{s}\|^2 \mathbf{I}_{M_t} = E_s T \mathbf{I}_{M_t}, \tag{2}$$

for any $\mathbf{s} \in {\mathsf{A}_M}^N$, where E_s is he energy of the A_M constellation used. Equality (2) denotes orthogonality and according to [9] leads to maximum spatial diversity gain.
The properties that hold for $\{\mathbf{A}_n, \mathbf{B}_n\}$ are :

$$\begin{aligned}
\mathbf{A}_n, \mathbf{B}_n &\in \mathbf{C}^{M_t \times T} \\
\mathbf{A}_n \mathbf{A}_n^H &= \mathbf{B}_n \mathbf{B}_n^H = I \\
\mathbf{A}_n \mathbf{A}_p H &= -\mathbf{A}_p \mathbf{A}_n^H \quad , \mathbf{B}_n \mathbf{B}_p^H = -\mathbf{B}_p \mathbf{B}_n^H \quad n \neq p \\
\mathbf{A}_n \mathbf{B}_p^H &= \mathbf{B}_p \mathbf{A}_n^H .
\end{aligned}$$

Let $\mathbf{s}^{(i)}$ denote the data vector contained in the i-th transmitted code block, $i = 1, 2, \ldots, P$, where P is the total number of transmissions. The form of the corresponding OSTBC code matrix is

$$\mathbf{C}(\mathbf{s}^{(i)}) = \sum_{n=1}^{N} (\mathbf{A}_n \Re\{s_n^{(i)}\} + j\mathbf{B}_n \Im\{s_n^{(i)}\}). \tag{3}$$

We observe that (3) can be rewritten as

$$\mathbf{C}(\mathbf{s}^{(i)}) = \sum_{n=1}^{N} (\check{\mathbf{A}}_n s_n^{(i)} + \check{\mathbf{B}}_n s_n^{(i)^*}), \tag{4}$$

for code matrices $\{\check{\mathbf{A}}, \check{\mathbf{B}}\}$ that satisfy

$$\begin{aligned} \check{\mathbf{A}}_n &= \frac{\mathbf{A}_n + \mathbf{B}_n}{2} \\ \check{\mathbf{B}}_n &= \frac{\mathbf{A}_n - \mathbf{B}_n}{2}. \end{aligned}$$

Based on the algorithm proposed in [34], we can create and use square (i.e. $M_t = T$) complex orthogonal designs, which have either "conjugate" or "non-conjugate" rows[1]. This means, that the code matrices, for all the transmissions, have the same rows "conjugate" or "non-conjugate".

According to the consideration above, for a certain row of $\mathbf{C}(\mathbf{s})$ – let it be the xth – either the xth row of all $\check{\mathbf{A}}_n$'s or the xth row of all $\check{\mathbf{B}}_n$'s is non-zero. In our developments, we regard that all the transmitted code matrices belong to this set of complex orthogonal designs and we propose the following lemma.

Lemma 1: Having fixed $x \in \{1, \ldots, M_t\}$, either $\begin{bmatrix} \check{\mathbf{A}}_1 & \ldots & \check{\mathbf{A}}_{NP} \end{bmatrix}_{x,:}$ or $\begin{bmatrix} \check{\mathbf{B}}_1 & \ldots & \check{\mathbf{B}}_{NP} \end{bmatrix}_{x,:}$ can be non-zero.

$\square$

Moreover, from equation (4) we obtain

$$\begin{aligned} \mathbf{C}(\mathbf{s}^{(i)}) &= \sum_{n=1}^{N} \begin{bmatrix} \check{\mathbf{A}}_n & \check{\mathbf{B}}_n \end{bmatrix} \left(\begin{bmatrix} s_n^{(i)} \\ s_n^{(i)^*} \end{bmatrix} \otimes \mathbf{I}_T \right) \\ &\triangleq \sum_{n=1}^{N} \mathbf{X}_n (\tilde{\mathbf{s}}_n^{(i)} \otimes \mathbf{I}_T) \\ &= \begin{bmatrix} \mathbf{X}_1 & \cdots & \mathbf{X}_N \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{s}_1}^{(i)} \otimes \mathbf{I}_T \\ \vdots \\ \tilde{\mathbf{s}_N}^{(i)} \otimes \mathbf{I}_T \end{bmatrix} \\ &\triangleq \mathbf{X}\left(\begin{bmatrix} \tilde{\mathbf{s}_1}^{(i)} \\ \vdots \\ \tilde{\mathbf{s}_N}^{(i)} \end{bmatrix} \otimes \mathbf{I}_T \right) \\ &\triangleq \mathbf{X}(\tilde{\mathbf{s}}^{(i)} \otimes \mathbf{I}_T) \end{aligned} \tag{5}$$

where $\tilde{\mathbf{s}}_n^{(i)} = \begin{bmatrix} s_n^{(i)} \\ s_n^{(i)^*} \end{bmatrix}$, $\tilde{\mathbf{s}}^{(i)} = \begin{bmatrix} \tilde{\mathbf{s}_1}^{(i)} \\ \vdots \\ \tilde{\mathbf{s}_N}^{(i)} \end{bmatrix}$, $\forall n \in \{1, \ldots, N\}$ and $\forall i \in \{1, \ldots, P\}$, and $\mathbf{X} \triangleq \begin{bmatrix} \mathbf{X}_1 & \ldots & \mathbf{X}_N \end{bmatrix} \in \mathbb{C}^{M_t \times 2TN}$.

[1]For convenience of explanation, let us say a row in an orthogonal design is conjugate (non-conjugate) if all symbols except zeros in this row have (do not have) complex conjugate.

The downconverted and pulse-matched equivalent ith received block of size $M_r \times T$ is

$$\mathbf{Y}^{(i)} = \mathbf{H}^{(i)}\mathbf{C}(\mathbf{s}^{(i)}) + \mathbf{V}^{(i)}. \tag{6}$$

In (6), $\mathbf{H}^{(i)} \in \mathbb{C}^{M_r \times M_t}$ represents the Rayleigh channel matrix between the M_t transmit and M_r receive antennas for the ith transmission. Note that in the following derivations we consider the channel to remain stable for a fixed number of transmissions. In the sequel we assume that $\mathbf{H}^{(i)}$ consists of independent and identically distributed (i.i.d.) coefficients that are modeled as circular complex gaussian random variables and account for flat fading. In addition $\mathbf{V}^{(i)} \in \mathbb{C}^{M_r \times T}$ denotes zero mean additive spatially and temporally white circular complex Gaussian noise with variance σ_v^2. The channel and noise matrices $\mathbf{H}^{(i)}$ and $\mathbf{V}^{(i)}$, respectively, $i = 1, 2, \ldots, P$, are independent of each other.

If the receiver has knowledge of the channel matrix, then coherent ML detection simplifies to one-shot block decisions according to

$$\hat{\mathbf{s}}_{opt}^{(i)} \quad = \quad \arg \min_{\mathbf{s}^{(i)} \in \mathsf{A}_M{}^N} \|\mathbf{Y}^{(i)} - \mathbf{H}^{(i)}\mathbf{C}(\mathbf{s}^{(i)})\|_F^2, \tag{7}$$

for $i = 1, 2, \ldots, P$. In this work we consider the channel matrices $\mathbf{H}^{(i)}$, $i = 1, 2, \ldots, P$, to be unavailable to the receiver. Hence, coherent detection in (7) cannot be utilized and the ML receiver takes the form of a sequence detector. We consider a sequence of P space-time blocks consecutively transmitted by the source and collected by the receiver, say $\mathbf{Y}^{(1)}, \mathbf{Y}^{(2)}, \ldots, \mathbf{Y}^{(P)}$, and form the $M_r \times TP$ observation matrix

$$\mathbf{Y} \triangleq \begin{bmatrix} \mathbf{Y}^{(1)} & \ldots & \mathbf{Y}^{(P)} \end{bmatrix} = \begin{bmatrix} \mathbf{H}^{(1)}\mathbf{C}(\mathbf{s}^{(1)}) & \ldots & \mathbf{H}^{(P)}\mathbf{C}(\mathbf{s}^{(P)}) \end{bmatrix} + \begin{bmatrix} \mathbf{V}^{(1)} & \ldots & \mathbf{V}^{(P)} \end{bmatrix}. \tag{8}$$

Based on the convention that the channel remains stable for P consecutive transmissions, (8) can be written as

$$\mathbf{Y} \quad \triangleq \quad \mathbf{H} \begin{bmatrix} \mathbf{C}(\mathbf{s}^{(1)}) & \ldots & \mathbf{C}(\mathbf{s}^{(P)}) \end{bmatrix} + \begin{bmatrix} \mathbf{V}^{(1)} & \ldots & \mathbf{V}^{(P)} \end{bmatrix}. \tag{9}$$

In the following section, we present ML noncoherent detection developments.

III. Maximum Likelihood Noncoherent Detection

We consider a time-invariant Rayleigh fading MIMO channel and prove that the complexity of the ML detector at the receiver can be polynomial in the sequence length P if the rank of the channel covariance is not a function of the sequence length. Interestingly, the order of the polynomial complexity depends strictly on the rank of the channel covariance matrix and therefore on the number of antennas used at the transmitter and the receiver.

Due to Rayleigh fading, the vectorized single-transmission channel matrix $\mathbf{h}$ is a zero-mean circular complex Gaussian vector of length M_rM_t with covariance matrix $\mathbf{C}_h = E\{\mathbf{h}\mathbf{h}^H\} = a\mathbf{I}_{M_rM_t}$, where $a \triangleq {\sigma_h}^2$ is the variance of the channel. The covariance matrix is such, due to the independence and identical distribution of the channel coefficients, and has rank $D = M_rM_t$. Given the $M_r \times TP$ observation matrix $\mathbf{Y}$, the ML detector for the symbol sequence $\mathbf{s} = \begin{bmatrix} (\mathbf{s}^{(1)})^T & \ldots & (\mathbf{s}^{(P)})^T \end{bmatrix}^T \in \mathsf{A}_M{}^{NP}$

maximizes the conditional probability density function of $\mathbf{Y}$ given $\mathbf{s}$. Thus, the optimal decision is given by

$$\hat{\mathbf{s}}_{opt} = \arg\max_{\mathbf{s}\in \mathsf{A}_M{}^{NP}} f(\mathbf{Y}|\mathbf{s}) = \arg\max_{\mathbf{s}\in \mathsf{A}_M{}^{NP}} f(\mathrm{vec}(\mathbf{Y})|\mathbf{s}) = \arg\max_{\mathbf{s}\in \mathsf{A}_M{}^{NP}} f(\mathbf{y}|\mathbf{s}) \tag{10}$$

where $\mathbf{y} \stackrel{\triangle}{=} \mathrm{vec}(\mathbf{Y}) \in \mathbb{C}^{M_r M_t P}$ and $f(\cdot|\cdot)$ represents the pertinent matrix/vector probability density function of the channel output conditioned on a symbol sequence.
We define the concatenated matrix of the transmitted $\mathbf{G} = \begin{bmatrix} \mathbf{C}(\mathbf{s}^{(1)}) & \ldots & \mathbf{C}(\mathbf{s}^{(P)}) \end{bmatrix} \in \mathbb{C}^{M_t \times TP}$ and note that it satisfies the orthogonality property since

$$\begin{aligned} \mathbf{G}(\mathbf{s})\mathbf{G}^H(\mathbf{s}) &= (\begin{bmatrix} \mathbf{C}(\mathbf{s}^{(1)}) & \ldots & \mathbf{C}(\mathbf{s}^{(P)}) \end{bmatrix})(\begin{bmatrix} \mathbf{C}(\mathbf{s}^{(1)}) & \ldots & \mathbf{C}(\mathbf{s}^{(P)}) \end{bmatrix})^H \\ &= \sum_{i=1}^{P} \mathbf{C}(\mathbf{s}^{(i)})\mathbf{C}^H(\mathbf{s}^{(i)}) \\ &= E_s TP\mathbf{I}_{M_t}. \end{aligned} \tag{11}$$

Then, the received matrix in (9) becomes

$$\mathbf{Y} \stackrel{\triangle}{=} \mathbf{HG}(\mathbf{s}) + \mathbf{V} \in \mathbb{C}^{M_r \times TP} \tag{12}$$

As stated in [25]

$$\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{B}). \tag{13}$$

Due to (13) we obtain

$$\mathbf{y} = \mathrm{vec}(\mathbf{HG}(\mathbf{s}) + \mathbf{V}) = \mathrm{vec}(\mathbf{I}_{M_r}\mathbf{HG}(\mathbf{s})) + \mathrm{vec}(\mathbf{V}) = (\mathbf{G}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r})\mathbf{h} + \mathbf{v} \tag{14}$$

where $\mathbf{v} = \mathrm{vec}(\mathbf{V}) \in \mathbb{C}^{M_r TP}$ and operator $\otimes$ denotes the Kronecker tensor product. Then, it can be proven that $\mathbf{y}$ given $\mathbf{s}$ is a complex Gaussian vector with mean $E\{\mathbf{y}|\mathbf{s}\} = E\{(\mathbf{G}^T(\mathbf{s}) \otimes I_{M_r})\mathbf{h} + \mathbf{v}|\mathbf{s}\} = (\mathbf{G}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r})E\{\mathbf{h}\} + E\{\mathbf{v}\} = \mathbf{0}_{M_r TP}$ and covariance matrix

$$\begin{aligned} \mathbf{C}_y(\mathbf{s}) &= E\{((\mathbf{G}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r})\mathbf{h} + \mathbf{v})((\mathbf{G}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r})\mathbf{h} + \mathbf{v})^H|\mathbf{s}\} \\ &= E\{(\mathbf{G}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r})\mathbf{hh}^H(\mathbf{G}^*(\mathbf{s}) \otimes \mathbf{I}_{M_r})|\mathbf{s}\} + E\{\mathbf{vv}^H|\mathbf{s}\} \\ &= (\mathbf{G}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r})\mathbf{C}_h(\mathbf{G}^*(\mathbf{s}) \otimes \mathbf{I}_{M_r}) + \sigma_v^2\mathbf{I}_{M_r TP} \\ &= a(\mathbf{G}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r})(\mathbf{G}^*(\mathbf{s}) \otimes \mathbf{I}_{M_r}) + \sigma_v^2\mathbf{I}_{M_r TP}. \end{aligned} \tag{15}$$

Therefore, the optimization problem in (10) is rewritten as

$$\begin{aligned} \hat{\mathbf{s}}_{opt} &= \arg\max_{\mathbf{s}\in \mathsf{A}_M{}^{NP}} \frac{1}{\pi^{M_r TP}|\mathbf{C}_y(\mathbf{s})|} \exp\{-(\mathbf{y} - (\mathbf{G}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r})\boldsymbol{\mu})^H \mathbf{C}_y^{-1}(\mathbf{s})(\mathbf{y} - (\mathbf{G}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r})\boldsymbol{\mu})\} \\ &= \arg\max_{\mathbf{s}\in \mathsf{A}_M{}^{NP}} \frac{1}{\pi^{M_r TP}|\mathbf{C}_y(\mathbf{s})|} \exp\{-\mathbf{y}^H\mathbf{C}_y^{-1}\mathbf{y}\}. \end{aligned} \tag{16}$$

A natural approach to (16) would be an exhaustive search among all M^{NP} data sequences $\mathbf{s} \in \mathsf{A}_M{}^{NP}$, but such a receiver is impractical even for moderate values of P, since its complexity grows exponentially with P. In the sequel, we prove that we can appliy an efficient algorithm that performs the maximization of (16) with $\mathsf{O}((\frac{MNP}{2})^{2D})$ calculations.

Using Sylvester's determinant theorem and Sherman-Morrison-Woodbury formula for the inverse of a rank-deficient update [26], we compute

$$\begin{aligned}
|\mathbf{C}_y(\mathbf{s})| &= |\sigma_v^2\mathbf{I}_{M_rTP}||\mathbf{I}_D+\frac{a}{\sigma_v^2}(\mathbf{G}^*(\mathbf{s})\otimes\mathbf{I}_{M_r})(\mathbf{G}^T(\mathbf{s})\otimes\mathbf{I}_{M_r})| \\
&= |\sigma_v^2\mathbf{I}_{M_rTP}||\mathbf{I}_D+\frac{a}{\sigma_v^2}(\mathbf{G}^*(\mathbf{s})\mathbf{G}^T(\mathbf{s})\otimes\mathbf{I}_{M_r})| \overset{(11)}{=} \sigma_v^{2M_rTP}|\mathbf{I}_D+\frac{aE_sTP}{\sigma_v^2}(\mathbf{I}_{M_t}\otimes\mathbf{I}_{M_r})| \\
&= \sigma_v^{2M_rTP}|\mathbf{I}_D+\frac{aE_sT}{\sigma_v^2}\mathbf{I}_D| \\
&= \sigma_v^{2M_rTP}|\frac{aE_sTP+1}{\sigma_v^2}\mathbf{I}_D| \qquad (17)
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{C}_y^{-1}(\mathbf{s}) &= \frac{1}{\sigma_v^2}\mathbf{I}_{M_rTP}-\frac{a}{\sigma_v^2}(\mathbf{G}^T(\mathbf{s})\otimes\mathbf{I}_{M_r})((\mathbf{I}_D+\frac{a}{\sigma_v^2}(\mathbf{G}^*(\mathbf{s})\otimes\mathbf{I}_{M_r})(\mathbf{G}^T(\mathbf{s})\otimes\mathbf{I}_{M_r}))^{-1}(\mathbf{G}^*(\mathbf{s})\otimes\mathbf{I}_{M_r})\frac{1}{\sigma_v^2} \\
&\overset{(11)}{=} \frac{1}{\sigma_v^2}\mathbf{I}_{M_rTP}-\frac{a}{\sigma_v^2}(\mathbf{G}^T(\mathbf{s})\otimes\mathbf{I}_{M_r})(\mathbf{I}_D+\frac{aE_sTP}{\sigma_v^2}\mathbf{I}_{M_t}\otimes\mathbf{I}_{M_r})^{-1}(\mathbf{G}^*(\mathbf{s})\otimes\mathbf{I}_{M_r})\frac{1}{\sigma_v^2} \\
&= \frac{1}{\sigma_v^2}\mathbf{I}_{M_rTP}-\frac{a}{\sigma_v^4}(\mathbf{G}^T(\mathbf{s})\otimes\mathbf{I}_{M_r})(\frac{aE_sTP+1}{\sigma_v^2}\mathbf{I}_D)^{-1}(\mathbf{G}^*(\mathbf{s})\otimes\mathbf{I}_{M_r}) \\
&= \frac{1}{\sigma_v^2}\mathbf{I}_{M_rTP}-\frac{a}{(aE_sTP+1)\sigma_v^2}(\mathbf{G}^T(\mathbf{s})\otimes\mathbf{I}_{M_r})(\mathbf{G}^*(\mathbf{s})\otimes\mathbf{I}_{M_r}) \\
&= \frac{1}{\sigma_v^2}\mathbf{I}_{M_rTP}-b(\mathbf{G}^T(\mathbf{s})\otimes\mathbf{I}_{M_r})(\mathbf{G}^*(\mathbf{s})\otimes\mathbf{I}_{M_r}) \qquad (18)
\end{aligned}$$

where $b\overset{\triangle}{=}\frac{a}{(aE_sTP+1)\sigma_v^2}$. We observe that $|\mathbf{C}_y(\mathbf{s})|$ is independent of the transmitted sequence $\mathbf{s}$, drop it from the maximization in (16), and substitute (18) in (16) to obtain

$$\begin{aligned}
\hat{\mathbf{s}}_{opt} &= \arg\max_{\mathbf{s}\in\mathsf{A}_M{}^{NP}}\{-\mathbf{y}^H(\frac{1}{\sigma_v^2}\mathbf{I}_{M_rTP}-b(\mathbf{G}^T(\mathbf{s})\otimes\mathbf{I}_{M_r})(\mathbf{G}^*(\mathbf{s})\otimes\mathbf{I}_{M_r}))\mathbf{y}\} \\
&= \arg\max_{\mathbf{s}\in\mathsf{A}_M{}^{NP}}\{-\frac{1}{\sigma_v^2}\mathbf{y}^H\mathbf{y}+b\mathbf{y}^H(\mathbf{G}^T(\mathbf{s})\otimes\mathbf{I}_{M_r})(\mathbf{G}^*(\mathbf{s})\otimes\mathbf{I}_{M_r})\mathbf{y}\} \\
&= \arg\max_{\mathbf{s}\in\mathsf{A}_M{}^{NP}}\{\mathbf{y}^H(\mathbf{G}^T(\mathbf{s})\otimes\mathbf{I}_{M_r})(\mathbf{G}^*(\mathbf{s})\otimes\mathbf{I}_{M_r})\mathbf{y}\} \\
&\overset{\triangle}{=} \arg\max_{\mathbf{s}\in\mathsf{A}_M{}^{NP}}\{\mathbf{z}^H\mathbf{z}\}. \qquad (19)
\end{aligned}$$

We continue our algorithmic developments by defining the matrices $\mathbf{S}=\left[\begin{array}{ccc}\mathbf{s}^{(1)} & \cdots & \mathbf{s}^{(P)}\end{array}\right]\in\mathsf{A}_M{}^{N\times P}$, $\tilde{\mathbf{S}}=\left[\begin{array}{ccc}\tilde{\mathbf{s}}^{(1)} & \cdots & \tilde{\mathbf{s}}^{(P)}\end{array}\right]\in\mathsf{A}_M{}^{2N\times P}$.
Note that

$$\tilde{\mathbf{S}}=\left[\begin{array}{ccc}\mathbf{s}^{(1)} & \cdots & \mathbf{s}^{(P)}\end{array}\right]\otimes\begin{bmatrix}1\\0\end{bmatrix}+\left[\begin{array}{ccc}\mathbf{s}^{(1)} & \cdots & \mathbf{s}^{(P)}\end{array}\right]^*\otimes\begin{bmatrix}0\\1\end{bmatrix}=\mathbf{S}\otimes\begin{bmatrix}1\\0\end{bmatrix}+\mathbf{S}^*\otimes\begin{bmatrix}0\\1\end{bmatrix}$$

and

$$
\begin{aligned}
\mathbf{G}(\mathbf{s}) &= \left[\begin{array}{ccc} \mathbf{C}(\mathbf{s}^{(1)}) & \cdots & \mathbf{C}(\mathbf{s}^{(P)}) \end{array}\right] \\
&= \left[\begin{array}{ccc} \mathbf{X}(\tilde{\mathbf{s}}^{(1)} \otimes \mathbf{I}_T) & \cdots & \mathbf{X}(\tilde{\mathbf{s}}^{(P)} \otimes \mathbf{I}_T) \end{array}\right] \\
&= \mathbf{X}(\left[\begin{array}{ccc} \tilde{\mathbf{s}}^{(1)} & \cdots & \tilde{\mathbf{s}}^{(P)} \end{array}\right] \otimes \mathbf{I}_T) \\
&\triangleq \mathbf{X}(\tilde{\mathbf{S}} \otimes \mathbf{I}_T). \qquad (20)
\end{aligned}
$$

We also observe that

$$
\mathbf{s} \triangleq \mathrm{vec}(\mathbf{S}) \in \mathsf{A}_M{}^{NP}
$$

and

$$
\begin{aligned}
\tilde{\mathbf{s}} \triangleq \mathrm{vec}(\tilde{\mathbf{S}}) &= \mathrm{vec}(\mathbf{S}) \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \mathrm{vec}(\mathbf{S}^*) \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
&\triangleq \mathbf{s} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \mathbf{s}^* \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix}.
\end{aligned}
$$

In addition, vector $\mathbf{z}$ that appears in the maximization problem in (19) is re-expressed as

$$
\mathbf{z} = (\mathbf{G}^*(\mathbf{s}) \otimes \mathbf{I}_{M_r})\mathbf{y} = \mathrm{vec}(\mathbf{Y}\mathbf{G}^H(\mathbf{s})\mathbf{I}_{M_t}) = (\mathbf{I}_{M_t} \otimes \mathbf{Y})\mathrm{vec}((\tilde{\mathbf{S}}^H \otimes \mathbf{I}_T)\mathbf{X}^H). \qquad (21)
$$

We denote by $\tilde{\mathbf{X}}_m$ the matrix that contains the mth rows of all N space-time matrices, that

$$
\tilde{\mathbf{X}}_m \triangleq \begin{bmatrix} \left[\begin{array}{c} \mathbf{X}_1 \end{array}\right]_{m,:} \\ \vdots \\ \left[\begin{array}{c} \mathbf{X}_N \end{array}\right]_{m,:} \end{bmatrix} \in \mathbb{C}^{N \times 2T} \qquad (22)
$$

and observe that

$$
\mathbf{X}^H = \left[\begin{array}{ccc} \mathbf{X}_1^H & \cdots & \mathbf{X}_N^H \end{array}\right]^T = \left[\begin{array}{ccc} \mathrm{vec}(\tilde{\mathbf{X}}_1^H) & \cdots & \mathrm{vec}(\tilde{\mathbf{X}}_{M_t}^H) \end{array}\right]
$$

and

$$
\begin{aligned}
\mathrm{vec}(\mathbf{X}^H) &= \mathrm{vec}(\left[\begin{array}{ccc} \mathbf{X}_1^H & \cdots & \mathbf{X}_N^H \end{array}\right]^T) = \mathrm{vec}(\left[\begin{array}{ccc} \mathrm{vec}(\tilde{\mathbf{X}}_1^H) & \cdots & \mathrm{vec}(\tilde{\mathbf{X}}_{M_t}^H) \end{array}\right]) \\
&= \mathrm{vec}(\left[\begin{array}{ccc} \tilde{\mathbf{X}}_1^H & \cdots & \tilde{\mathbf{X}}_{M_t}^H \end{array}\right]).
\end{aligned}
$$

Then,

$$
\begin{aligned}
\mathrm{vec}((\tilde{\mathbf{S}}^H \otimes \mathbf{I}_T)\mathbf{X}^H) &= \underbrace{\begin{bmatrix} \mathbf{I}_P \otimes (([1\ 0] \otimes \mathbf{I}_T)\tilde{\mathbf{X}}_1^H) \\ \vdots \\ \mathbf{I}_P \otimes (([1\ 0] \otimes \mathbf{I}_T)\tilde{\mathbf{X}}_{M_t}^H) \end{bmatrix}}_{\mathbf{Z}_A^H} \mathbf{s}^* + \underbrace{\begin{bmatrix} \mathbf{I}_P \otimes (([0\ 1] \otimes \mathbf{I}_T)\tilde{\mathbf{X}}_1^H) \\ \vdots \\ \mathbf{I}_P \otimes (([0\ 1] \otimes \mathbf{I}_T)\tilde{\mathbf{X}}_{M_t}^H) \end{bmatrix}}_{\mathbf{Z}_B^H} \mathbf{s} \\
&\triangleq \mathbf{Z}_A^H \mathbf{s}^* + \mathbf{Z}_B^H \mathbf{s} = [\mathbf{Z}_A^H\ \ \mathbf{Z}_B^H] \begin{bmatrix} \mathbf{s}^* \\ \mathbf{s} \end{bmatrix} \triangleq \check{\mathbf{Z}}^H \dot{\mathbf{s}}. \qquad (23)
\end{aligned}
$$

A proof of (23) is provided in the Appendix.
Substituting (23) in (21) , we obtain

$$\mathbf{z} = (\mathbf{I}_{M_t} \otimes \mathbf{Y})\check{\mathbf{Z}}^H \dot{\mathbf{s}}. \tag{24}$$

Due to (24), the maximization argument in (19) becomes

$$\dot{\mathbf{s}}^H \check{\mathbf{Z}}(\mathbf{I}_{M_t} \otimes \mathbf{Y}^H)(\mathbf{I}_{M_t} \otimes \mathbf{Y})\check{\mathbf{Z}}^H \dot{\mathbf{s}}. \tag{25}$$

and by setting $\mathbf{V} \overset{\triangle}{=} \check{\mathbf{Z}}(\mathbf{I}_{M_t} \otimes \mathbf{Y}^H)$, maximization in (19) can be written as

$$\hat{\mathbf{s}}_{opt} \quad = \quad \arg \max_{\mathbf{s} \in \mathsf{A}_M{}^{NP}} \{\dot{\mathbf{s}}^H \mathbf{V}\mathbf{V}^H \dot{\mathbf{s}}\}. \tag{26}$$

In this point, we introduce the $2D-1$ spherical coordinates [2]

$$\phi_1 \in (-\pi, \pi], \phi_2, \ldots, \phi_{2D-1} \in (-\frac{\pi}{2}, \frac{\pi}{2}]$$

and define the spherical coordinate vector

$$\boldsymbol{\phi} \overset{\triangle}{=} \begin{bmatrix} \phi_1 & \phi_2 & \ldots & \phi_{2D-1} \end{bmatrix}$$

as well as the $2D \times 1$ hyperpolar vector

$$\tilde{\mathbf{c}}(\boldsymbol{\phi}_{1:2D-1}) = \begin{bmatrix} \sin \phi_1 \\ \cos \phi_1 \sin \phi_2 \\ \cos \phi_1 \cos \phi_2 \sin \phi_3 \\ \vdots \\ \prod_{i=1}^{2D-2} \cos \phi_i \sin \phi_{2D-1} \\ \prod_{i=1}^{2D-1} \cos \phi_i \end{bmatrix}_{2D \times 1} .$$

We also define the *complex* vector of hyperspherical coordinates

$$\begin{aligned} \mathbf{c}'(\phi_{1:2D-1}) &= \tilde{\mathbf{c}}_{2:2:2D}(\phi_{1:2D-1}) + j\, \tilde{\mathbf{c}}_{1:2:2D}(\phi_{1:2D-1}) \\ &= \begin{bmatrix} \cos \phi_1 \sin \phi_2 + j \ \sin \phi_1 \\ \prod_{i=1}^{3} \cos \phi_i \sin \phi_4 + j \ \cos \phi_1 \cos \phi_2 \sin \phi_3 \\ \vdots \\ \prod_{i=1}^{2D-1} \cos \phi_i + j \ \prod_{i=1}^{2D-2} \cos \phi_i \sin \phi_{2D-1} \end{bmatrix}_{D \times 1} \\ &= \begin{bmatrix} c'_1(\phi_{1:2D-1}) \\ c'_2(\phi_{1:2D-1}) \\ \vdots \\ c'_D(\phi_{1:2D-1}) \end{bmatrix} \end{aligned}$$

[2]We recall that $D = M_r M_t$.

and separate it in M_t subvectors

$$\mathbf{c}'(\boldsymbol{\phi}) = \begin{bmatrix} \mathbf{c}'^{(1)}(\boldsymbol{\phi}) \\ \vdots \\ \mathbf{c}'^{(M_t)}(\boldsymbol{\phi}) \end{bmatrix}. \tag{27}$$

Subsequently, we map each one of the above subvectors $\mathbf{c}'^{(x)}, \forall x \in \{1, \ldots, M_t\}$, to a new vector $\mathbf{c}^{(x)}(\boldsymbol{\phi})$ according to the following rule:
$\forall x \in \{1, \ldots, M_t\}$

- if $x \in \mathbb{X}$, then $\mathbf{c}^{(x)}(\boldsymbol{\phi}) = \mathbf{c}'^{(x)}(\boldsymbol{\phi})$
- if $x \in \mathbb{Y}$, then $\mathbf{c}^{(x)}(\boldsymbol{\phi}) = \mathbf{c}'^{(x)*}(\boldsymbol{\phi})$

where $\mathbb{X}$ is the set of the "non-conjugate" rows of the COD and $\mathbb{Y}$ is the set of "conjugate" rows of the COD. Note that $\mathbb{X} \equiv \bar{\mathbb{Y}}$.

So, the maximization problem of (26) can be written as

$$\hat{\mathbf{s}}_{opt} \quad = \quad \arg \max_{\mathbf{s} \in \mathsf{A}_M{}^{NP}} \max_{\phi_1 \in (-\pi, \pi]} \max_{\phi_{2:2D-1} \in (-\frac{\pi}{2}, \frac{\pi}{2}]^{2D-2}} |\dot{\mathbf{s}}^H \mathbf{V} \mathbf{c}(\boldsymbol{\phi})|. \tag{28}$$

Due to Cauchy-Schwartz Inequality which states that, for any $\mathbf{v} \in \mathbb{R}^{2D}$,

$$\mathbf{v}^H \mathbf{c}(\boldsymbol{\phi}_{1:2D-1}) \leq |\mathbf{v}^H \mathbf{c}(\boldsymbol{\phi}_{1:2D-1})| \leq \|\mathbf{v}\| \|\mathbf{c}(\boldsymbol{\phi}_{1:2D-1})\|.$$

Since,

$$\begin{aligned} \|\mathbf{c}'(\boldsymbol{\phi}_{1:2D-1})\| &= \|\mathbf{c}(\boldsymbol{\phi}_{1:2D-1})\| \\ &= \sqrt{|c_1(\phi_{1:2D-1})|^2 + \cdots + |c_{2D}(\phi_{1:2D-1})|^2} \\ &= \sqrt{\cos^2 \phi_1 sin^2 \phi_2 + \sin^2 \phi_1 + \prod_{i=1}^{3} \cos^2 \phi_i sin \phi_4 + \cos^2 \phi_1 cos^2 \phi_2 sin^2 \phi_3 + \ldots} \\ &= \sqrt{\sin^2 \phi_1 + \cos^2 \phi_1 (\sin^2 \phi_2 + \cos^2 \phi_2 (\sin^2 \phi_3 + \ldots)))} \\ &= \sqrt{1} = 1 \end{aligned}$$

the above Cauchy-Schwartz Inequality can be written as

$$|\mathbf{v}^H \mathbf{c}(\boldsymbol{\phi})| \leq \|\mathbf{v}\|. \tag{29}$$

The equality of (29) is achieved if and only if $\phi_1, \phi_2, \ldots, \phi_{2D-1}$ are hyperspherical coordinates of $\mathbf{v}$. So, (28) becomes

$$\hat{\mathbf{s}}_{opt} \quad = \quad \arg \max_{\mathbf{s} \in \mathsf{A}_M{}^{NP}} \max_{\phi_1 \in (-\pi, \pi]} \max_{\phi_{2:2D-1} \in (-\frac{\pi}{2}, \frac{\pi}{2}]^{2D-2}} \Re\{\dot{\mathbf{s}}^H \mathbf{V} \mathbf{c}(\boldsymbol{\phi})\}. \tag{30}$$

Equivalently,

$$\begin{aligned}\hat{\mathbf{s}}_{opt} &= \arg \max_{\mathbf{s}\in \mathsf{A}_M{}^{NP}} \max_{\phi_1\in(-\pi,\pi]} \max_{\phi_{2:2D-1}\in(-\frac{\pi}{2},\frac{\pi}{2}]^{2D-2}} \sum_{n=1}^{2NP} \Re\{\dot{s}_n^* \mathbf{V}_{n,1:D}\mathbf{c}(\boldsymbol{\phi})\} \\ &= \arg \max_{\mathbf{s}\in \mathsf{A}_M{}^{NP}} \max_{\phi_{1:2D-1}} \sum_{n=1}^{2NP} \Re\{\dot{s}_n^* \mathbf{V}_{n,1:D}\mathbf{c}(\boldsymbol{\phi})\}. \end{aligned} \tag{31}$$

By interchanging the maximizations in (31) we obtain

$$\begin{aligned}\hat{\mathbf{s}}_{opt} &= \arg\max_{\phi_{1:2D-1}} \sum_{n=1}^{2NP} \max_{s_n\in \mathsf{A}_M} \Re\{\dot{s}_n^* \mathbf{V}_{n,1:D}\mathbf{c}(\boldsymbol{\phi})\} \\ &= \arg\max_{\phi_{1:2D-1}} \sum_{n=1}^{NP} \max_{s_n\in \mathsf{A}_M} \Re\{\dot{s}_n^* \mathbf{V}_{n,1:D}\mathbf{c}(\boldsymbol{\phi}) + \dot{s}_{n+NP}^* \mathbf{V}_{n+NP,1:D}\mathbf{c}(\boldsymbol{\phi})\}. \end{aligned}$$

By the definition given for $\dot{\mathbf{s}}$ we know that , $\forall i : 1 \le i \le NP$, $\dot{s}_i = s_i^*$ and $\dot{s}_{i+NP} = s_i$.
Thus, we get

$$\begin{aligned}\hat{\mathbf{s}}_{opt} &= \arg\max_{\phi_{1:2D-1}} \sum_{n=1}^{NP} \max_{s_n\in \mathsf{A}_M} \Re\{s_n \mathbf{V}_{n,1:D}\mathbf{c}(\boldsymbol{\phi}) + s_n^* \mathbf{V}_{n+NP,1:D}\mathbf{c}(\boldsymbol{\phi})\} \\ &= \arg\max_{\phi_{1:2D-1}} \sum_{n=1}^{NP} \max_{s_n\in \mathsf{A}_M} \{\Re\{s_n \mathbf{V}_{n,1:D}\mathbf{c}(\boldsymbol{\phi})\} + \Re\{s_n^* \mathbf{V}_{n+NP,1:D}\mathbf{c}(\boldsymbol{\phi})\}\} \\ &= \arg\max_{\phi_{1:2D-1}} \sum_{n=1}^{NP} \max_{s_n\in \mathsf{A}_M} \{\Re\{s_n \mathbf{V}_{n,1:D}\mathbf{c}(\boldsymbol{\phi})\} + \Re\{s_n \mathbf{V}_{n+NP,1:D}^* \mathbf{c}^*(\boldsymbol{\phi})\}\}. \end{aligned}$$

Hence,

$$\begin{aligned}\hat{\mathbf{s}}_{opt} &= \arg\max_{\phi_{1:2D-1}} \sum_{n=1}^{NP} \max_{s_n\in \mathsf{A}_M} \Re\{s_n \underbrace{\left[\begin{array}{cc} \mathbf{V}_{n,1:D} & \mathbf{V}_{n+NP,1:D}^* \end{array}\right]}_{\mathbf{W}_{n,1:2D}} \underbrace{\left[\begin{array}{c} \mathbf{c}(\boldsymbol{\phi}) \\ \mathbf{c}^*(\boldsymbol{\phi}) \end{array}\right]}_{\mathbf{d}(\boldsymbol{\phi})}\} \\ &= \arg\max_{\phi_{1:2D-1}} \sum_{n=1}^{NP} \max_{s_n\in \mathsf{A}_M} \Re\{s_n \mathbf{W}_{n,1:2D}\mathbf{d}(\boldsymbol{\phi})\} \\ &= \arg \max_{\mathbf{s}\in \mathsf{A}_M{}^{NP}} \max_{\phi_{1:2D-1}} \Re\{\mathbf{s}^T \mathbf{W}\mathbf{d}(\boldsymbol{\phi})\}. \end{aligned} \tag{32}$$

where $\mathbf{W} \stackrel{\triangle}{=} \left[\begin{array}{cc} \mathbf{V}_{1:NP,:} & \mathbf{V}_{1+NP:2NP,:}^* \end{array}\right] \in \mathbb{C}^{NP\times 2D}$.
As already stated, $\mathbf{V} = \check{\mathbf{Z}}(\mathbf{I}_{M_t} \otimes \mathbf{Y}^H)$ and $\check{\mathbf{Z}} = \left[\begin{array}{c} \mathbf{Z}_A \\ \mathbf{Z}_B \end{array}\right]$. Let us now divide $\mathbf{V}$ into two submatrices $\mathbf{V}_1$ and $\mathbf{V}_2$ as

$$\begin{aligned}\mathbf{V}_1 &\stackrel{\triangle}{=} \mathbf{V}_{1:NP,:} = \mathbf{Z}_A(\mathbf{I}_{M_t} \otimes \mathbf{Y}^H) \\ \mathbf{V}_2 &\stackrel{\triangle}{=} \mathbf{V}_{NP+1:2NP,:} = \mathbf{Z}_B(\mathbf{I}_{M_t} \otimes \mathbf{Y}^H),\end{aligned}$$

and separate the matrices $\mathbf{Z}_A$ and $\mathbf{Z}_B$ into M_t submatrices each. Every submatrix will have NP rows and PT columns.

$$\mathbf{Z}_A = \left[\begin{array}{ccc} \mathbf{Z}_A^{(1)} & \dots & \mathbf{Z}_A^{(Mt)} \end{array} \right]$$

and

$$\mathbf{Z}_B = \left[\begin{array}{ccc} \mathbf{Z}_B^{(1)} & \dots & \mathbf{Z}_B^{(Mt)} \end{array} \right]$$

Lemma 2: Having fixed $x \in \{1,\dots,M_t\}$, either $\mathbf{Z}_A^{(x)}$ or $\mathbf{Z}_B^{(x)}$ is non-zero.

- if $x \in \mathbb{X}$, then $\mathbf{Z}_A^{(x)}$ is non-zero.
- if $x \in \mathbb{Y}$, then $\mathbf{Z}_B^{(x)}$ is non-zero.

□

As defined in (32), $\mathbf{W}$ gets constructed from $\mathbf{V}$ as

$$\begin{aligned} \mathbf{W} &= [\mathbf{V}_1 \;\; \mathbf{V}_2^*] = [\mathbf{Z}_A(\mathbf{I}_{M_t} \otimes \mathbf{Y}^H) \;\; \mathbf{Z}_B^*(\mathbf{I}_{M_t} \otimes \mathbf{Y}^T)] \\ &= [\mathbf{Z}_A^{(1)}\mathbf{Y}^H \dots \mathbf{Z}_A^{(M_t)}\mathbf{Y}^H \;\; \mathbf{Z}_B^{(1)*}\mathbf{Y}^T \dots \mathbf{Z}_B^{(M_t)*}\mathbf{Y}^T]. \end{aligned}$$

In this point, we separate $\mathbf{W}$ in $2M_t$ submatrices as

$$\mathbf{W} = [\mathbf{W}^{(1)} \dots \mathbf{W}^{(2M_t)}]$$

where $\forall x \in \{1,\dots,M_t\}$, $\mathbf{W}^{(x)} = \mathbf{Z}_A^{(x)}\mathbf{Y}^H$ and $\mathbf{W}^{(x+M_t)} = \mathbf{Z}_B^{(x)*}\mathbf{Y}^T$.

Lemma 3: Having fixed $x \in \{1,\dots,M_t\}$, either $\mathbf{W}^{(x)}$ or $\mathbf{W}^{(x+M_t)}$ is non-zero.

- if $x \in \mathbb{X}$, then $\mathbf{W}^{(x)}$ is non-zero.
- if $x \in \mathbb{Y}$, then $\mathbf{W}^{(x+M_t)}$ is non-zero.

□

We also separate $\mathbf{d}(\boldsymbol{\phi})$ in $2M_t$ subvectors as

$$\mathbf{d}(\boldsymbol{\phi}) = \left[\begin{array}{c} \mathbf{d}^{(1)}(\boldsymbol{\phi}) \\ \vdots \\ \mathbf{d}^{(2M_t)}(\boldsymbol{\phi}) \end{array} \right].$$

Note that $\forall x \in \{1,\dots,M_t\}$, $\mathbf{d}^{(x)}(\boldsymbol{\phi}) = \mathbf{c}^{(x)}(\boldsymbol{\phi})$ and $\mathbf{d}^{(x+M_t)}(\boldsymbol{\phi}) = \mathbf{c}^{(x)^*}(\boldsymbol{\phi})$.
Hence, the maximization in (32) can be rewritten as

$$\begin{aligned} \hat{\mathbf{s}}_{opt} &= \arg \max_{\mathbf{s} \in \mathsf{A}_M{}^{NP}} \max_{\phi_{1:2D-1}} \Re\{\mathbf{s}^T [\mathbf{W}^{(1)} \dots \mathbf{W}^{(2M_t)}] \left[\begin{array}{c} \mathbf{d}^{(1)}(\boldsymbol{\phi}) \\ \vdots \\ \mathbf{d}^{(2M_t)}(\boldsymbol{\phi}) \end{array} \right] \} \\ &= \arg \max_{\mathbf{s} \in \mathsf{A}_M{}^{NP}} \max_{\phi_{1:2D-1}} \Re\{\mathbf{s}^T (\sum_{x=1}^{M_t} \mathbf{W}^{(x)}\mathbf{c}^{(x)}(\boldsymbol{\phi}) + \sum_{x=1}^{M_t} \mathbf{W}^{(x+M_t)}\mathbf{c}^{(x)^*}(\boldsymbol{\phi}))\}. \end{aligned} \quad (33)$$

According to *lemma 3* and the definition of vector $\mathbf{c}(\boldsymbol{\phi})$, for a specific x, we get two disjoint cases

- if $x \in \mathbb{X}$, then $\mathbf{W}^{(x)}$ will be non-zero and $\mathbf{c}^{(x)} = \mathbf{c}'^{(x)}$.
- if $x \in \mathbb{Y}$, then $\mathbf{W}^{(x+M_t)}$ will be non-zero and $\mathbf{c}^{(x)^*} = \mathbf{c}'^{(x)}$.

Thus, the maximization above is equivalent to

$$\begin{aligned}
\hat{\mathbf{s}}_{opt} &= \arg \max_{\mathbf{s} \in \mathsf{A}_M{}^{NP}} \max_{\phi_{1:2D-1}} \Re\{\mathbf{s}^T (\sum_{x=1}^{M_t} \underbrace{\left(\mathbf{W}^{(x)} + \mathbf{W}^{(x+M_t)}\right)}_{\check{\mathbf{W}}^{(x)}} \mathbf{c}'^{(x)}(\boldsymbol{\phi}))\} \\
&= \arg \max_{\mathbf{s} \in \mathsf{A}_M{}^{NP}} \max_{\phi_{1:2D-1}} \Re\{\mathbf{s}^T (\sum_{x=1}^{M_t} \check{\mathbf{W}}^{(x)} \mathbf{c}'^{(x)}(\boldsymbol{\phi}))\} \\
&= \arg \max_{\mathbf{s} \in \mathsf{A}_M{}^{NP}} \max_{\phi_{1:2D-1}} \Re\{\mathbf{s}^T \check{\mathbf{W}} \mathbf{c}'(\boldsymbol{\phi})\}. \qquad (34)
\end{aligned}$$

Due the Cauchy-Schwartz Inequality [26]

$$\Re\{\mathbf{s}^T \check{\mathbf{W}} \mathbf{c}'(\boldsymbol{\phi})\} \leq |\mathbf{s}^T \check{\mathbf{W}} \mathbf{c}'(\boldsymbol{\phi})| \leq \|\check{\mathbf{W}}^H \mathbf{s}^*\| \|\mathbf{c}'(\boldsymbol{\phi})\| = \|\check{\mathbf{W}}^H \mathbf{s}^*\| = \|\check{\mathbf{W}}^T \mathbf{s}\|,$$

the maximization of (26) can be rewritten as

$$\hat{\mathbf{s}}_{opt} = \arg \max_{\mathbf{s} \in \mathsf{A}_M{}^{NP}} \|\check{\mathbf{W}}^H \mathbf{s}^*\| = \arg \max_{\mathbf{s} \in \mathsf{A}_M{}^{NP}} \|\check{\mathbf{W}}^T \mathbf{s}\| = \arg \max_{\mathbf{s} \in \mathsf{A}_M{}^{NP}} \|\boldsymbol{\Gamma}^H \mathbf{s}\| = \arg \max_{\mathbf{s} \in \mathsf{A}_M{}^{NP}} \mathbf{s}^H \boldsymbol{\Gamma} \boldsymbol{\Gamma}^H \mathbf{s} \qquad (35)$$

where $\boldsymbol{\Gamma} \overset{\triangle}{=} \check{\mathbf{W}}^* \in \mathbb{C}^{D \times NP}$.

The computation of $\hat{\mathbf{s}}_{opt}$ in (35) can be implemented with complexity $\mathsf{O}((\frac{MNP}{2})^{2D})$ if we follow the multiple-auxiliary-angle methodology that has been introduced in [32] for the problem of rank-deficient quadratic form maximization.

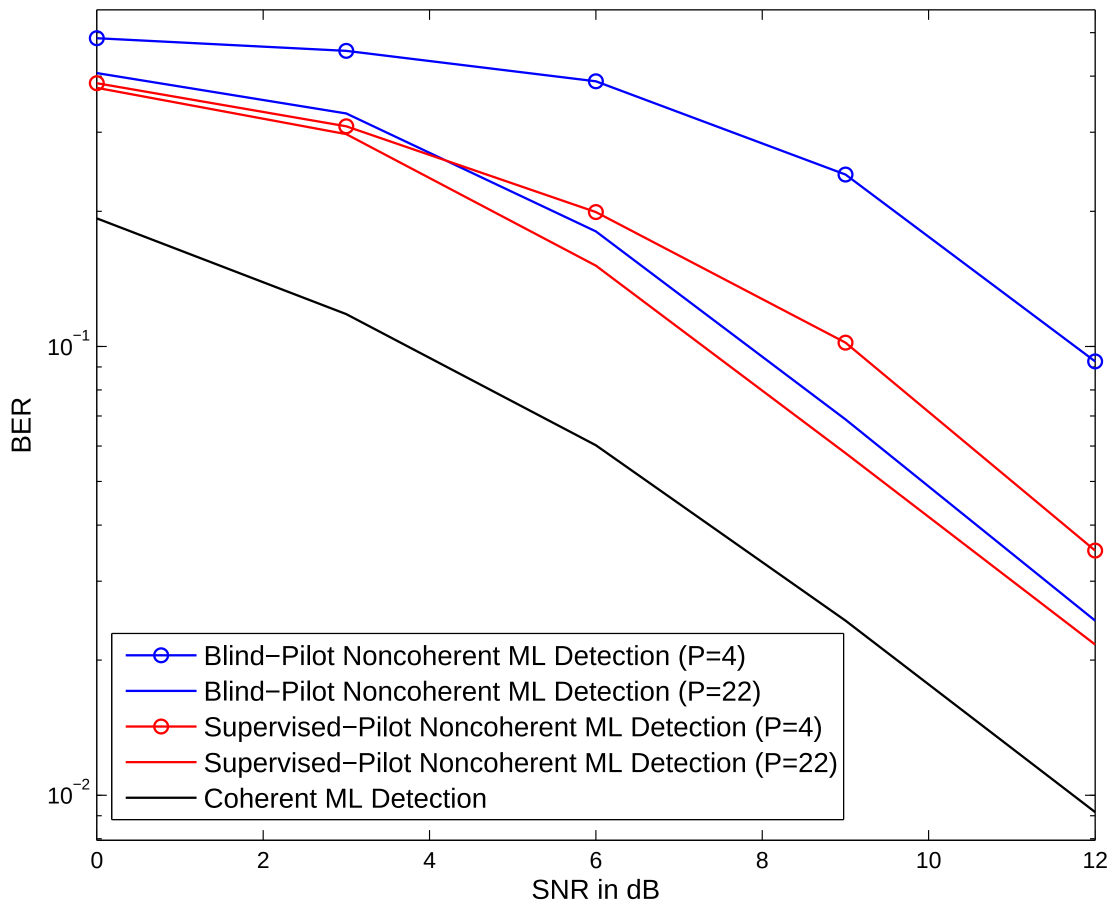

Figure 1: BER versus SNR for ML coherent OSTBC receivers and proposed pilot-assisted ("blind" and "supervised") ML noncoherent OSTBC receivers with sequence length $P = 4$ and $P = 22$ upon static Rayleigh channel.

IV. Simulation Studies

We consider a 2×1 MISO system that employs Alamouti space-time coding (with rate $R = \frac{N}{T} = 1$ since $N = T = 2$) to transmit binary data in an unknown Rayleigh fading channel environment. Space-time ambiguity induced by the rotatability of the Alamouti code is resolved by means of pilot-assisted transmission. In Figure 1 we study the Rayleigh fading channel (i.i.d. channel coefficients) case and present the bit error rate (BER) of the one-shot coherent ML receiver and the pilot-assited ML noncoherent receiver implemented by the proposed algorithm as a function of the information SNR for sequence lengths $P = 4$ and $P = 22$. The rank of the channel covariance matrix is 2 since it is a scaled version of $\mathbf{I}_2$. Therefore, the overall complexity of the proposed pilot-assisted ML receiver becomes $\mathrm{O}((\frac{MNP}{2})^{2D})$. We present results averaged over 300 channel realizations.

Note that, for comparison purposes, we apply both "blind" and "supervised" pilot-assisted transmission. In the first one ("blind" pilot-assisted transmission), we regard the transmitter to send N

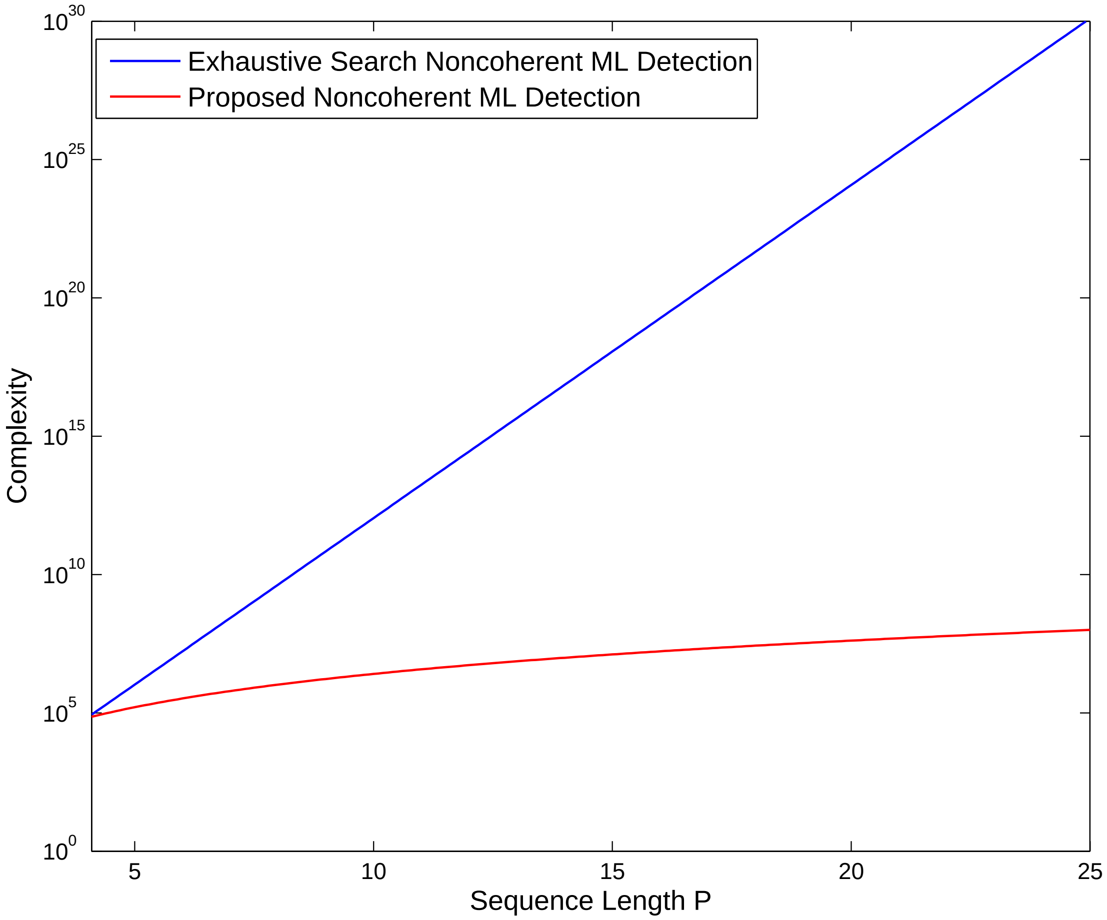

Figure 2: Complexity versus sequence length P.

pilot symbols only in the start of the symbol sequence. The receiver uses this knowledge to make a correct decision for the following NP symbols; in the following received symbol sequences it is the last N detected symbols from the previous one which are considered as pilot symbols. In the second one ("supervised" pilot-assisted transmission), the receiver has knowledge over the first N symbols of every transmission. This means that P pilot symbol sequences are being used and the data rate decreases considerably, while performance improves.

However, a quite interesting observation in Figure 1 is that the longer the symbol sequence is, the smaller the loss that the "blind" pilot-assisted ML noncoherent receiver exhibits in comparison with the "supervised" one. For example, for $P = 22$ the SNR loss of the "blind" pilot-assisted receiver is approximately 0.5 dB in comparison to the "supervised". Moreover, by increasing the length of the symbol sequence, the SNR loss of the "blind" pilot-assisted receiver in contrary to the ML coherent one decreases as well.

Figure 2 demonstrates the significant complexity gain offered by the proposed algorithm. For example, for $P = 22$ conventional ML noncoherent detection would demand an exhaustive search among $M^{NP} = 4^{44} = 2^{88}$ (for QPSK Alamouti 2×1) binary vectors, while a receiver which applies the pro-

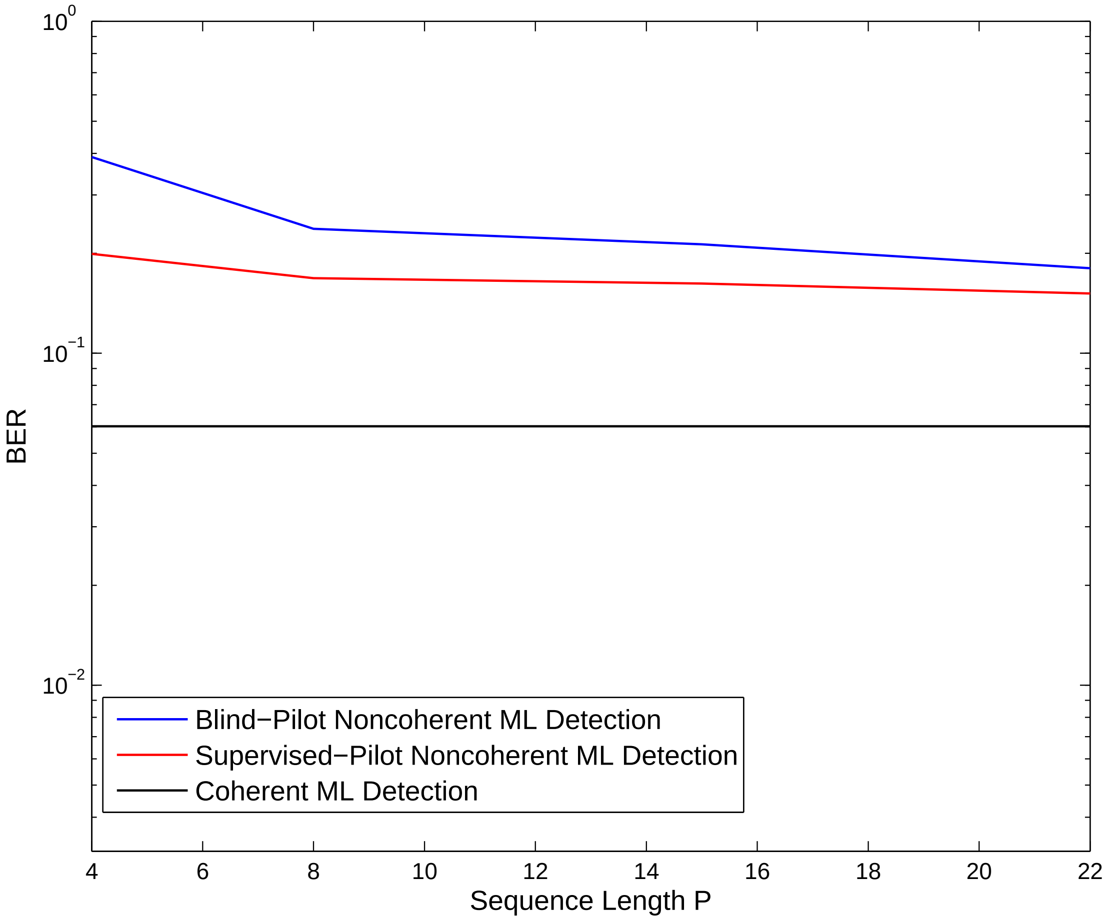

Figure 3: BER versus sequence length P.

posed algorithm would search among $(\frac{MNP}{2})^{2D} = 88^4 \approx \lceil 2^{25.8} \rceil$ binary vectors.
In Fig. 3, we set the SNR to 6 dB and present BER curves of the proposed ML noncoherent receiver versus sequence length P. We observe that, if the "blind" pilot-assisted receiver operates with the same complexity with the "supervised" pilot-assisted receiver, then its performance becomes better as the sequence length P increases. The BER of the coherent ML receiver is also presented as a performance lower bound.

V. Conclusion

We proved that ML noncoherent sequence detection is polynomially solvable with respect to the sequence length for OSTBC and static Rayleigh channels (i.i.d) and presented the constraints under which this is achievable. Complexity exponent is only a function of the rank of the product of the number of antennas at the transmitter and the receiver. This work comes as an expansion of [31] of D. S. Papailiopoulos and G. N. Karystinos and generalizes the idea of building a ML OSTBC polynomial-complexity detector when an MPSK instead of a BPSK constellation is being utilized.

Appendix

Proof of Lemma 2

Let us fix $x \in \{1,\ldots,M_t\}$. Then $\forall i \in \{1,\ldots,NP\}$ and $\forall k \in \{1,\ldots,N\}$, either

$$\begin{aligned}
x \in \mathbb{X} \quad &\text{or} \quad x \in \mathbb{Y} \\
\Leftrightarrow [\check{\mathbf{A}}_k]_{x,:} \neq \mathbf{0}_T^T \quad &\text{or} \quad [\check{\mathbf{B}}_k]_{x,:} \neq \mathbf{0}_T^T \\
\Leftrightarrow [\mathbf{X}_k]_{x,1:T} \neq \mathbf{0}_T^T \quad &\text{or} \quad [\mathbf{X}_k]_{x,T+1:2T} \neq \mathbf{0}_T^T \\
\Leftrightarrow [\tilde{\mathbf{X}}_x]_{k,1:T} \neq \mathbf{0}_T^T \quad &\text{or} \quad [\tilde{\mathbf{X}}_x]_{k,T+1:2T} \neq \mathbf{0}_T^T \\
\Leftrightarrow \left[\ [\tilde{\mathbf{X}}_x]_{:,1:T}\ \right]_{k,:} \neq \mathbf{0}_T^T \quad &\text{or} \quad \left[\ [\tilde{\mathbf{X}}_x]_{:,T+1:2T}\ \right]_{k,:} \neq \mathbf{0}_T^T \\
\Leftrightarrow [\tilde{\mathbf{X}}_x(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \mathbf{I}_T)]_{k,:} \neq \mathbf{0}_T^T \quad &\text{or} \quad [\tilde{\mathbf{X}}_x(\begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \mathbf{I}_T)]_{k,:} \neq \mathbf{0}_T^T \\
\Leftrightarrow \mathbf{I}_P \otimes [\tilde{\mathbf{X}}_x(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \mathbf{I}_T)]_{k,:} \neq \mathbf{0}_{PT}^T \quad &\text{or} \quad \mathbf{I}_P \otimes [\tilde{\mathbf{X}}_x(\begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \mathbf{I}_T)]_{k,:} \neq \mathbf{0}_{PT}^T \\
\Leftrightarrow [\mathbf{Z}_A^{(x)}]_{i,:} \neq \mathbf{0}_{PT}^T \quad &\text{or} \quad [\mathbf{Z}_B^{(x)}]_{i,:} \neq \mathbf{0}_{PT}^T \\
\Leftrightarrow \mathbf{Z}_A^{(x)} \neq \mathbf{0}_{NP\times PT} \quad &\text{or} \quad \mathbf{Z}_B^{(x)} \neq \mathbf{0}_{NP\times PT}.
\end{aligned}$$

Proof of Lemma 3

Consider a fixed $x \in \{1,\ldots,M_t\}$. Then, $\mathbf{W}^{(x)} = \mathbf{Z}_A^{(x)}\mathbf{Y}^H$ and $\mathbf{W}^{(x+M_t)} = \mathbf{Z}_B^{(x)}\mathbf{Y}^T$.
As stated in *lemma 2*, either $\mathbf{Z}_A^{(x)}$ or $\mathbf{Z}_B^{(x)}$ is non-zero.
Thus, either $\mathbf{W}^{(x)}$ or $\mathbf{W}^{(x+M_t)}$ are non-zero.

Proof of (23)

$$
\begin{aligned}
\mathrm{vec}((\tilde{\mathbf{S}}^H \otimes \mathbf{I}_T)\mathbf{X}^H) &= \mathrm{vec}(\begin{bmatrix} (\tilde{\mathbf{S}}^H \otimes \mathbf{I}_T)\mathrm{vec}(\tilde{\mathbf{X}}_1^H) & \cdots & (\tilde{\mathbf{S}}^H \otimes \mathbf{I}_T)\mathrm{vec}(\tilde{\mathbf{X}}_{M_t}^H) \end{bmatrix}) \\
&= \mathrm{vec}(\begin{bmatrix} ((\mathbf{S}^H \otimes [1\ 0] + \mathbf{S}^T \otimes [0\ 1]) \otimes \mathbf{I}_T)\mathrm{vec}(\tilde{\mathbf{X}}_1^H) \\ \vdots \\ ((\mathbf{S}^H \otimes [1\ 0] + \mathbf{S}^T \otimes [0\ 1]) \otimes \mathbf{I}_T)\mathrm{vec}(\tilde{\mathbf{X}}_{M_t}^H) \end{bmatrix}) \\
&= \mathrm{vec}(\begin{bmatrix} (\mathbf{S}^H \otimes [1\ 0] \otimes \mathbf{I}_T)\mathrm{vec}(\tilde{\mathbf{X}}_1^H) + (\mathbf{S}^T \otimes [0\ 1] \otimes \mathbf{I}_T)\mathrm{vec}(\tilde{\mathbf{X}}_1^H) \\ \vdots \\ (\mathbf{S}^H \otimes [1\ 0] \otimes \mathbf{I}_T)\mathrm{vec}(\tilde{\mathbf{X}}_{M_t}^H) + (\mathbf{S}^T \otimes [0\ 1] \otimes \mathbf{I}_T)\mathrm{vec}(\tilde{\mathbf{X}}_{M_t}^H) \end{bmatrix}) \\
&= \mathrm{vec}(\begin{bmatrix} (\mathbf{S}^H \otimes [1\ 0] \otimes \mathbf{I}_T)\mathrm{vec}(\tilde{\mathbf{X}}_1^H) \\ \vdots \\ (\mathbf{S}^H \otimes [1\ 0] \otimes \mathbf{I}_T)\mathrm{vec}(\tilde{\mathbf{X}}_{M_t}^H) \end{bmatrix}) + \mathrm{vec}(\begin{bmatrix} (\mathbf{S}^T \otimes [0\ 1] \otimes \mathbf{I}_T)\mathrm{vec}(\tilde{\mathbf{X}}_1^H) \\ \vdots \\ (\mathbf{S}^T \otimes [0\ 1] \otimes \mathbf{I}_T)\mathrm{vec}(\tilde{\mathbf{X}}_{M_t}^H) \end{bmatrix}) \\
&= \begin{bmatrix} (\mathbf{S}^H \otimes [1\ 0] \otimes \mathbf{I}_T)\mathrm{vec}(\tilde{\mathbf{X}}_1^H) \\ \vdots \\ (\mathbf{S}^H \otimes [1\ 0] \otimes \mathbf{I}_T)\mathrm{vec}(\tilde{\mathbf{X}}_{M_t}^H) \end{bmatrix} + \begin{bmatrix} (\mathbf{S}^T \otimes [0\ 1] \otimes \mathbf{I}_T)\mathrm{vec}(\tilde{\mathbf{X}}_1^H) \\ \vdots \\ (\mathbf{S}^T \otimes [0\ 1] \otimes \mathbf{I}_T)\mathrm{vec}(\tilde{\mathbf{X}}_{M_t}^H) \end{bmatrix} \\
&= \begin{bmatrix} \mathrm{vec}(([1\ 0] \otimes \mathbf{I}_T)\tilde{\mathbf{X}}_1^H \mathbf{S}^*) \\ \vdots \\ \mathrm{vec}(([1\ 0] \otimes \mathbf{I}_T)\tilde{\mathbf{X}}_{M_t}^H \mathbf{S}^*) \end{bmatrix} + \begin{bmatrix} \mathrm{vec}(([0\ 1] \otimes \mathbf{I}_T)\tilde{\mathbf{X}}_1^H \mathbf{S}) \\ \vdots \\ \mathrm{vec}(([0\ 1] \otimes \mathbf{I}_T)\tilde{\mathbf{X}}_{M_t}^H \mathbf{S}) \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{I}_P \otimes (([1\ 0] \otimes \mathbf{I}_T)\tilde{\mathbf{X}}_1^H) \\ \vdots \\ \mathbf{I}_P \otimes (([1\ 0] \otimes \mathbf{I}_T)\tilde{\mathbf{X}}_{M_t}^H) \end{bmatrix} \mathbf{s}^* + \begin{bmatrix} \mathbf{I}_P \otimes (([0\ 1] \otimes \mathbf{I}_T)\tilde{\mathbf{X}}_1^H) \\ \vdots \\ \mathbf{I}_P \otimes (([0\ 1] \otimes \mathbf{I}_T)\tilde{\mathbf{X}}_{M_t}^H) \end{bmatrix} \mathbf{s}.
\end{aligned}
$$

Sylvester's Determinant Theorem

Consider a $M \times N$ matrix $\mathbf{A}$ and a $N \times M$ matrix $\mathbf{B}$. Then,

$$|\mathbf{I}_M + \mathbf{AB}| = |\mathbf{I}_N + \mathbf{BA}|.$$

Sherman-Morrison-Woodbury Formula

Consider a $N \times N$ matrix $\mathbf{A}$, a $N \times K$ matrix $\mathbf{U}$, a $K \times K$ matrix $\mathbf{C}$ and a $K \times N$ matrix $\mathbf{V}$. Then,

$$(\mathbf{A} + \mathbf{UCV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}\left(\mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U}\right)^{-1}\mathbf{VA}^{-1}.$$

This formula can be derived using blockwise matrix inversion.

References

[1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585-595, Nov./Dec. 1999.

[2] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Y. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, pp. 527-537, Apr. 1999.

[3] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, pp. 744-765, Mar. 1998.

[4] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 311-335, Mar. 1998.

[5] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 45, pp. 139-157, Jan. 1999.

[6] D. Gesbert, M. Sha, D.-S. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 281-302, Apr. 2003.

[7] S. N. Diggavi, N. Al-Dhahir, A. Stamoulis, and A. R. Calderbank, "Great expectations: The value of spatial diversity in wireless networks," in *Proc. IEEE*, vol. 92, pp. 219-270, Feb. 2004.

[8] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451-1458, Oct. 1998.

[9] 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.

[10] E. G. Larsson, P. Stoica, and J. Li, "On maximum-likelihood detection and decoding for space-time coding systems," *IEEE Trans. Signal Process.*, vol. 50, pp. 937-944, Apr. 2002.

[11] A. Sezgin, E. A. Jorswieck, O. Henkel, S. Pereira, and A. Paulraj, "On the relation of OSTBC and code rate one QSTBC: Average rate, BER, and coding gain," *IEEE Trans. Signal Process.*, vol. 56, pp. 4879-4891, Oct. 2008.

[12] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 48, pp. 359-383, Feb. 2002.

[13] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Inf. Theory*, vol. 46, pp. 1962-1973, Sept. 2000.

[14] 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, pp. 362-372, Feb. 2003.

[15] W.-K. Ma, B.-N. Vo, T. N. Davidson, and P.-C. Ching, "Blind ML detection of orthogonal space-time block codes: Efficient high-performance implementations," *IEEE Trans. Signal Process.*, vol. 54, pp. 738-751, Feb. 2006.

[16] T. Cui and C. Tellambura, "Efficient blind receiver design for orthogonal space-time block codes," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1890-1899, May 2007.

[17] M.-T. Le, V.-S. Pham, L. Mai, and G. Yoon, "Efficient algorithm for blind detection of orthogonal space-time block codes," *IEEE Signal Process. Lett.*, vol. 14, pp. 301-304, May 2007.

[18] M. Grötschel, L. Lovasz, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, 2nd ed. New York: Springer-Verlag, 1993.

[19] M. Ajtai, "The shortest vector problem in L_2 is NP-hard for randomized reductions," in *Proc. 30th Annu. ACM Symp. Theory of Computing*, Dallas, TX, May 1998, pp. 10-19.

[20] J. Jaldén and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Trans. Signal Process.*, vol. 53, pp. 1474-1484, Apr. 2005.

[21] A. Krishnamoorthy and A. Anastasopoulos, "Code and receiver design for the noncoherent fast-fading channel," *IEEE Trans. Commun.*, vol. 23, pp. 1769-1778, Sept. 2005.

[22] P. Dayal, M. Brehler, and M. K. Varanasi, "Leveraging coherent space-time codes for noncoherent communication via training," *IEEE Trans. Inf. Theory*, vol. 50, pp. 2058-2080, Sept. 2004.

[23] I. Motedayen-Aval, A. Krishnamoorthy, and A. Anastasopoulos, "Optimal joint detection/estimation in fading channels with polynomial complexity," *IEEE Trans. Inf. Theory*, vol. 53, pp. 209-223, Jan. 2007.

[24] V. Pauli, L. Lampe, R. Schober, and K. Fukuda, "Multiple-symbol differential detection based on combinatorial geometry," *IEEE Trans. Commun.*, vol. 56, pp. 1596-1600, Oct. 2008.

[25] A. Graham, *Kronecker Products and Matrix Calculus with Applications*. Chichester, U.K.: Ellis Horwood, 1981.

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

[27] H. El Gamal, D. Aktas, and M. O. Damen, "Noncoherent space-time coding: An algebraic perspective," *IEEE Trans. Inf. Theory*, vol. 51, pp. 2380-2390, July 2005.

[28] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, pp. 2740-2746, June 2007.

[29] W.-K. Ma, "Blind ML detection of orthogonal space-time block codes: Identifiability and code construction," *IEEE Trans. Signal Process.*, vol. 55, pp. 3312-3324, July 2007.

[30] D. S. Papailiopoulos and G. N. Karystinos, "Efficient computation of the M-phase vector that maximizes a rank-deficient quadratic form," *42nd Annual Conference on Information Sciences and Systems*, 2008.

[31] D. S. Papailiopoulos and G. N. Karystinos, "Maximum-likelihood noncoherent OSTBC detection with polynomial complexity," *IEEE Trans. Wireless Commun.*, vol. 9, June 2010.

[32] A. Kyrillidis, "Polynomial-complexity computation of the M-phase vector that maximizes a rank-defficient quadratic form," M.Sc. Thesis, Electronic and Computer Engineering Dept., Technical University of Crete, Sept. 2010.

[33] W. Su, S. N. Batalama and D. A. Pados, "On orthogonal space-time block codes and transceiver signal linearization," *IEEE Commun. Letters*, vol. 10, Feb. 2006.

[34] W. Su, X.-G. Xia and K. J. R. Liu, "A systematic design of high-rate complex orthogonal space-time block codes," *IEEE Commun. Letters*, vol. 8, June 2004.

[35] E. G. Larsson and P. Stoica, *Space-time block coding for wireless communications*. Cambridge, U.K.: Cambridge University Press, 2003.