Maximum-Likelihood Noncoherent OSTBC Detection with Polynomial Complexity

Dimitris S. Papailiopoulos, *Student Member, IEEE*, and George N. Karystinos, *Member, IEEE*

***Abstract*—We prove that maximum-likelihood (ML) noncoherent sequence detection of orthogonal space-time block coded signals can always be performed in polynomial time with respect to the sequence length for static Rayleigh, correlated (in general) channels. Moreover, using recent results on efficient maximization of rank-deficient quadratic forms over finite alphabets, we develop an algorithm that performs ML noncoherent sequence detection with polynomial complexity. The order of the polynomial complexity of the proposed receiver equals two times the rank of the covariance matrix of the vectorized channel matrix. Therefore, the lower the Rayleigh channel covariance rank the lower the receiver complexity. Instead, for Ricean channel distribution, we prove that polynomial complexity is attained through the proposed receiver as long as the mean channel vector is in the range of the covariance matrix of the vectorized channel matrix. Hence, full-rank channel correlation is desired to guarantee polynomial ML noncoherent detection complexity for the case of static Ricean fading. Our results are presented for the general case of block-fading Rayleigh or Ricean channels where we provide conditions under which ML noncoherent sequence detection can be performed in polynomial time through our algorithm.**

***Index Terms*—Blind detection, maximum-likelihood (ML), multiple-input multiple-output (MIMO), orthogonal space-time block code (OSTBC), quadratic form maximization.**

I. Introduction

PRESENT and next generation wireless standards aim at high data rates and reliable communications, features afforded by multiple antenna systems that are proven to attain higher channel capacity than single antenna setups while lowering the error probability [1]-[7]. Elaborate information-theoretic results tailored to Rayleigh fading [4] prove that channel capacity actually grows linearly when the number of receive and transmit antennas (simultaneously) increases.

It is, however, natural that antenna arrays are costly and space demanding, thus being a more plausible setup at base stations rather than remote terminals. As a result, transmit diversity techniques have enjoyed primary focus, with the first pioneering work coming from Alamouti [8] that delivered the first full-diversity, full-rate space-time block code (STBC) for two transmit antennas. Tarokh, Jafarkhani, and Calderbank generalized the design of the work in [8] to more than two transmit antennas introducing a paradigm for the construction of space-time block codes based on orthogonal designs [9]. The so-called orthogonal STBCs (OSTBCs) are proven to achieve full antenna diversity gain with linear-complexity single-symbol maximum-likelihood (ML) coherent detection [9], [10]. OSTBCs outperform nonorthogonal 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 [11].

Such an error rate is attainable with linear complexity, if the channel state information (CSI) is available at the receiver. However, the very 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 [3], with the direct implication of reduced effective 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 [12] shows that when CSI is not available the capacity of multi-antenna 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 [5], [10], [13]-[15]. However, if sequence detection is performed through exhaustive search among all possible data sequences [5], [10], [13], [15], then exponential computational complexity is required. 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 [5] and later proven to also take the form of a binary quadratic form maximization problem [16]-[19] that in the general case is NP-hard [20], [21]. In [17], [19] 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 [22]. To avoid the exponential complexity of the optimal

Manuscript received December 29, 2008; revised June 22, 2009 and February 7, 2010; accepted February 24, 2010. The associate editor coordinating the review of this paper and approving it for publication was L. Lampe.

This work was supported in part by the Sixth Framework Program of the European Commission under Grant MIRG-CT-2007-46563. The material in this paper was presented in part at the 46th Annual Allerton Conference on Communication, Control, and Computing, Allerton House, Monticello, IL, September 2008, and in part at the 43rd Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, November 2009.

D. S. Papailiopoulos was with the Department of Electronic and Computer Engineering, Technical University of Crete, Chania, 73100, Greece. He is now with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089 USA (e-mail: papailio@usc.edu).

G. N. Karystinos is with the Department of Electronic and Computer Engineering, Technical University of Crete, Chania, 73100, Greece (e-mail: karystinos@telecom.tuc.gr).

Digital Object Identifier 10.1109/TWC.2010.06.081715

receiver many suboptimal schemes have been proposed in the literature such as differential detection schemes [23]-[28], the cyclic ML receiver proposed in [15] based on alternating optimization, and semidefine relaxation approaches [16]-[19]. To combat the exponential cost of the optimal noncoherent receiver the aid of pilot symbols was considered in [29]-[30] at the expense of information rate.

In this work, we consider the case of block-fading Ricean channels and prove that ML noncoherent OSTBC detection can be performed in polynomial time when the rank of the covariance matrix of the vectorized channel matrix is not a function of the sequence length and the mean channel vector is in the range of the channel covariance matrix; the polynomial complexity order is completely determined by the rank of the channel covariance matrix, namely, it equals two times the rank of the covariance matrix. Furthermore, motivated by the works in [31]-[36] which treat the problem of rank-deficient quadratic form maximization, we provide an algorithm that solves the ML noncoherent OSTBC detection problem in polynomial time. We tailor to our detection problem the algorithm originally introduced in [33]-[36] and observe that the polynomial in time solution lies in the utilization of multiple auxiliary spherical variables. The optimal data sequence is proven to belong to a polynomial in size set of binary vectors that is built in polynomial time, altogether resulting in an efficient, fixed-complexity algorithm.

Especially for the static Rayleigh fading channel, the channel mean is zero and the channel covariance matrix rank is always less than or equal to the product of the numbers of transmit and receive antennas, hence polynomial-complexity detection is always guaranteed. In contrast to Rayleigh fading, full-rank channel correlation is desired to guarantee polynomial-complexity ML noncoherent detection upon static Ricean fading "channel processing." For illustration purposes, we operate the proposed receiver for sequence lengths up to more than 100 bits in the context of plain 2×2 Alamouti transmissions in unknown Rayleigh or Ricean fading channel environments. Even when the bit-sequence length exceeds 100, the polynomial-complexity feature of our algorithm allows ML noncoherent detection without a prohibitive computational cost.

II. System Model and Problem Statement

We consider a multiple-input multiple-output (MIMO) system with M_{t} transmit and M_{r} receive antennas that employs orthogonal space-time coded transmission of size $M_{\mathrm{t}} \times T$ and rate $R = \frac{N}{T}$, $N \leq T$. We assume transmission of binary data that are split into vectors of N bits. Each bit vector is used to generate a corresponding space-time block (matrix) of size $M_{\mathrm{t}} \times T$. The $M_{\mathrm{t}} \times T$ space-time block $\mathbf{C}(\mathbf{s}) \in \mathbb{C}^{M_{\mathrm{t}} \times T}$ that corresponds to the $N \times 1$ data vector $\mathbf{s} \in \{\pm 1\}^N$ is given by

$$\mathbf{C}(\mathbf{s}) = \sum_{n=1}^{N} \mathbf{X}_n s_n \tag{1}$$

where $s_n = \pm 1$ denotes the nth element (bit) of $\mathbf{s}$, $n = 1, 2, \ldots, N$, and $\mathbf{X}_n \in \mathbb{C}^{M_{\mathrm{t}} \times T}$, $n = 1, 2, \ldots, N$, are orthogonal space-time encoding matrices that enforce the property

$$\mathbf{C}(\mathbf{s})\mathbf{C}^H(\mathbf{s}) = \|\mathbf{s}\|^2 \mathbf{I}_{M_{\mathrm{t}}} = N\mathbf{I}_{M_{\mathrm{t}}}, \tag{2}$$

for any $\mathbf{s} \in \{\pm 1\}^N$. Eq. (2) denotes orthogonality and leads to maximum spatial diversity gain [9].

Let $\mathbf{s}^{(p)} = \left[s_1^{(p)}\ s_2^{(p)}\ \ldots\ s_N^{(p)}\right]^T$ denote the data vector contained in the pth transmitted code block, $p = 1, 2, 3, \ldots$. The downconverted and pulse-matched equivalent pth received block of size $M_{\mathrm{r}} \times T$ is

$$\mathbf{Y}^{(p)} = \mathbf{H}^{(p)}\mathbf{C}\left(\mathbf{s}^{(p)}\right) + \mathbf{W}^{(p)}. \tag{3}$$

In (3), $\mathbf{H}^{(p)} \in \mathbb{C}^{M_{\mathrm{r}} \times M_{\mathrm{t}}}$ refers to the pth transmission and represents the channel matrix between the M_{t} transmit and M_{r} receive antennas. In general, $\mathbf{H}^{(p)}$ consists of correlated coefficients that are modeled as circular complex Gaussian random variables and account for flat fading. We assume that all collected energy is absorbed by the channel matrix $\mathbf{H}^{(p)}$. In addition, $\mathbf{W}^{(p)} \in \mathbb{C}^{M_{\mathrm{r}} \times T}$ denotes zero-mean additive spatially and temporally white circular complex Gaussian noise with variance σ_v^2. The channel and noise matrices $\mathbf{H}^{(p)}$ and $\mathbf{W}^{(p)}$, respectively, $p = 1, 2, 3, \ldots$, 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

$$\hat{\mathbf{s}}^{(p)} = \underset{\mathbf{s}^{(p)} \in \{\pm 1\}^N}{\arg\min} \left\|\mathbf{Y}^{(p)} - \mathbf{H}^{(p)}\mathbf{C}\left(\mathbf{s}^{(p)}\right)\right\|_{\mathrm{F}}^2,\ p = 1, 2, 3, \ldots. \tag{4}$$

Since orthogonal space-time codes are utilized, exhaustive search among the 2^N possible bit vectors $\mathbf{s}^{(p)} \in \{\pm 1\}^N$ need not be performed because the detector in (4) is equivalent to linear-complexity single-bit decisions of the form

$$\hat{s}_n^{(p)} = \mathrm{sign}\left(\Re\left\{\mathrm{tr}\left\{\mathbf{Y}^{(p)}\mathbf{X}_n^H\left(\mathbf{H}^{(p)}\right)^H\right\}\right\}\right), \tag{5}$$

$n = 1, 2, \ldots, N$, $p = 1, 2, 3, \ldots$. In this work, we assume that the channel matrices $\mathbf{H}^{(p)}$, $p = 1, 2, 3, \ldots$, are not available to the receiver. Hence, coherent detection in (5) 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 $\mathbf{Y}^{(1)}, \ldots, \mathbf{Y}^{(P)}$, and form the $M_{\mathrm{r}} \times TP$ observation matrix

$$\begin{aligned} \mathbf{Y} &\triangleq \left[\mathbf{Y}^{(1)} \ldots \mathbf{Y}^{(P)}\right] \\ &= \left[\mathbf{H}^{(1)}\mathbf{C}\left(\mathbf{s}^{(1)}\right) \ldots \mathbf{H}^{(P)}\mathbf{C}\left(\mathbf{s}^{(P)}\right)\right] + \left[\mathbf{W}^{(1)} \ldots \mathbf{W}^{(P)}\right]. \end{aligned} \tag{6}$$

In the sequel, based on the observation of P blocks at the receiver we present ML noncoherent detection developments.

III. Maximum-likelihood Noncoherent Detection

We consider a block-fading Ricean fading MIMO channel, derive an efficient algorithm for the implementation of the ML noncoherent receiver, and prove that the complexity of the proposed ML receiver implementation is polynomial in the sequence length P if the mean channel vector is in the range of the channel covariance matrix and the rank of the latter 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. As a consequence, for the special case of static Ricean fading, full-rank channel

$$\hat{\tilde{\mathbf{s}}}_{\text{opt}} = \underset{\substack{\tilde{\mathbf{s}}\in\{\pm1\}^{NP+1}\\ \tilde{s}_{NP+1}=1}}{\arg\max}\ \tilde{\mathbf{s}}^T \begin{bmatrix} \frac{1}{\sigma_v^2}\mathbf{Z}\mathbf{E}^T\left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}^H\right)\mathbf{U}\mathbf{U}^H\left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}\right)\mathbf{E}\mathbf{Z}^H & \mathbf{Z}\mathbf{E}^T\left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}^H\right)\left(\mathbf{I}_{M_tM_rP}-\frac{N}{\sigma_v^2}\mathbf{U}\mathbf{U}^H\right)\boldsymbol{\mu} \\ \boldsymbol{\mu}^H\left(\mathbf{I}_{M_tM_rP}-\frac{N}{\sigma_v^2}\mathbf{U}\mathbf{U}^H\right)\left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}\right)\mathbf{E}\mathbf{Z}^H & 0 \end{bmatrix}\tilde{\mathbf{s}}. \quad (17)$$

correlation guarantees polynomial ML noncoherent detection complexity independently of the mean channel vector.

We assume that the channel matrix $\mathbf{H}^{(p)}$ varies during different transmissions and define the concatenated channel matrix $\mathbf{H} \triangleq \left[\mathbf{H}^{(1)}\ldots\mathbf{H}^{(P)}\right] \in \mathbb{C}^{M_r\times PM_t}$. Due to Ricean fading, the vectorized[1] channel matrix $\mathbf{h} \triangleq \text{vec}(\mathbf{H})$ is a circular complex Gaussian vector of length M_tM_rP with mean vector $\boldsymbol{\mu} \in \mathbb{C}^{M_tM_rP}$ and covariance matrix $\mathbf{C}_h = \text{E}\left\{(\mathbf{h}-\boldsymbol{\mu})(\mathbf{h}-\boldsymbol{\mu})^H\right\} = \mathbf{Q}\mathbf{Q}^H \in \mathbb{C}^{M_tM_rP\times M_tM_rP}$ where $\mathbf{Q} \in \mathbb{C}^{M_tM_rP\times D}$ consists of orthogonal columns and $D \leq M_tM_rP$. Given the $M_r \times TP$ observation matrix $\mathbf{Y}$, the ML detector for the bit sequence $\mathbf{s} = \left[\left(\mathbf{s}^{(1)}\right)^T\ldots\left(\mathbf{s}^{(P)}\right)^T\right]^T \in \{\pm1\}^{NP}$ maximizes the conditional probability density function (pdf) of $\mathbf{Y}$ given $\mathbf{s}$. Thus, the optimal decision is given by

$$\hat{\mathbf{s}}_{\text{opt}} = \underset{\mathbf{s}\in\{\pm1\}^{NP}}{\arg\max} f(\mathbf{Y}|\mathbf{s}) = \underset{\mathbf{s}\in\{\pm1\}^{NP}}{\arg\max} f(\mathbf{y}|\mathbf{s}) \quad (7)$$

where $\mathbf{y} \triangleq \text{vec}(\mathbf{Y}) \in \mathbb{C}^{M_rTP}$ and $f(\cdot|\cdot)$ represents the pertinent matrix/vector probability density function of the channel output conditioned on a bit sequence. If we define the block-diagonal matrix $\mathbf{D}(\mathbf{s}) \triangleq \text{diag}\left(\left[\mathbf{C}\left(\mathbf{s}^{(1)}\right),\ldots,\mathbf{C}\left(\mathbf{s}^{(P)}\right)\right]\right) \in \mathbb{C}^{M_tP\times TP}$, then the optimization problem in (7) is rewritten as

$$\hat{\mathbf{s}}_{\text{opt}} = \underset{\mathbf{s}\in\{\pm1\}^{NP}}{\arg\max} \frac{1}{\pi^{M_rTP}\left|\mathbf{C}_y(\mathbf{s})\right|} \times e^{-\left(\mathbf{y}-\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_r}\right)\boldsymbol{\mu}\right)^H\mathbf{C}_y^{-1}(\mathbf{s})\left(\mathbf{y}-\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_r}\right)\boldsymbol{\mu}\right)} \quad (8)$$

where $\mathbf{C}_y(\mathbf{s})$ is the covariance matrix of $\mathbf{y}$ given $\mathbf{s}$ and operator $\otimes$ denotes Kronecker tensor product. For a proof of (8), see the Appendix. A natural approach to (8) would be an exhaustive search among all 2^{NP} data sequences $\mathbf{s} \in \{\pm1\}^{NP}$, but such a receiver is impractical even for moderate values of P, since its complexity grows exponentially with P. In the sequel, we present an efficient algorithm that performs the maximization in (8) with $\mathcal{O}(P^{2D})$ calculations if $\boldsymbol{\mu}$ is in the span of $\mathbf{C}_h$.

For notation simplicity, we define the $D\times D$ diagonal matrix $\boldsymbol{\Sigma} \triangleq \mathbf{Q}^H\mathbf{Q}$ which contains the D positive eigenvalues of $\mathbf{C}_h$ on its diagonal and the $M_tM_rP \times D$ matrix

$$\mathbf{U} \triangleq \mathbf{Q}\left(\mathbf{I}_D + \frac{N}{\sigma_v^2}\boldsymbol{\Sigma}\right)^{-\frac{1}{2}}. \quad (9)$$

Then, as shown in the Appendix, the optimization problem in (8) becomes

$$\hat{\mathbf{s}}_{\text{opt}} = \underset{\mathbf{s}\in\{\pm1\}^{NP}}{\arg\max}\left\{\frac{1}{\sigma_v^2}\mathbf{y}^H\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_r}\right)\mathbf{U}\mathbf{U}^H\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_r}\right)\mathbf{y} + \mathbf{y}^H\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_r}\right)\left(\mathbf{I}_{M_tM_rP}-\frac{N}{\sigma_v^2}\mathbf{U}\mathbf{U}^H\right)\boldsymbol{\mu} + \boldsymbol{\mu}^H\left(\mathbf{I}_{M_tM_rP}-\frac{N}{\sigma_v^2}\mathbf{U}\mathbf{U}^H\right)\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_r}\right)\mathbf{y}\right\}. \quad (10)$$

We continue our algorithmic developments by defining the matrix $\mathbf{G}(\mathbf{s}) \triangleq \left[\mathbf{C}\left(\mathbf{s}^{(1)}\right)\ \ldots\ \mathbf{C}\left(\mathbf{s}^{(P)}\right)\right] \in \mathbb{C}^{M_t\times TP}$ and observing that $\mathbf{G}(\mathbf{s})\mathbf{G}^H(\mathbf{s}) = NP\mathbf{I}_{M_t}$ due to (2). Moreover, we denote by $\mathbf{e}_p$ the pth column of $\mathbf{I}_P$, $p = 1, 2, \ldots, P$, and define the matrix $\mathbf{E} \triangleq \begin{bmatrix}\mathbf{I}_{M_t}\otimes\mathbf{e}_1\mathbf{e}_1^T\\ \vdots \\ \mathbf{I}_{M_t}\otimes\mathbf{e}_P\mathbf{e}_P^T\end{bmatrix}\otimes\mathbf{I}_T \in \{0,1\}^{M_tP^2T\times M_tPT}$. We also denote by $\tilde{\mathbf{X}}_m$ the matrix that contains the mth rows of all N space-time matrices, that is

$$\tilde{\mathbf{X}}_m \triangleq \begin{bmatrix}[\mathbf{X}_1]_{m,:}\\ \vdots \\ [\mathbf{X}_N]_{m,:}\end{bmatrix} \in \mathbb{C}^{N\times T},\ m = 1,\ldots,M_t, \quad (11)$$

and define the matrices $\mathbf{Z}_m \triangleq \mathbf{I}_P\otimes\tilde{\mathbf{X}}_m \in \mathbb{C}^{NP\times TP}$, $m = 1,\ldots,M_t$, and $\mathbf{Z} \triangleq [\mathbf{Z}_1\ \ldots\ \mathbf{Z}_{M_t}] \in \mathbb{C}^{NP\times M_tTP}$. Then, the vector $\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_r}\right)\mathbf{y}$ that appears in the maximization problem in (10) is reexpressed as

$$\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_r}\right)\mathbf{y} = \left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}\right)\mathbf{E}\mathbf{Z}^H\mathbf{s}. \quad (12)$$

A proof of (12) is provided in the Appendix.

Due to (12), the first, second, and third part of the maximization argument in (10) become

$$\frac{1}{\sigma_v^2}\mathbf{s}^T\mathbf{Z}\mathbf{E}^T\left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}^H\right)\mathbf{U}\mathbf{U}^H\left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}\right)\mathbf{E}\mathbf{Z}^H\mathbf{s}, \quad (13)$$

$$\mathbf{s}^T\mathbf{Z}\mathbf{E}^T\left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}^H\right)\left(\mathbf{I}_{M_tM_rP}-\frac{N}{\sigma_v^2}\mathbf{U}\mathbf{U}^H\right)\boldsymbol{\mu}, \quad (14)$$

$$\text{and}\quad \boldsymbol{\mu}^H\left(\mathbf{I}_{M_tM_rP}-\frac{N}{\sigma_v^2}\mathbf{U}\mathbf{U}^H\right)\left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}\right)\mathbf{E}\mathbf{Z}^H\mathbf{s}, \quad (15)$$

respectively. We append a 1 to the end of the data vector $\mathbf{s}$, define $\tilde{\mathbf{s}} \triangleq \left[\mathbf{s}^T\ 1\right]^T$, and obtain

$$\hat{\mathbf{s}}_{\text{opt}} = \left[\hat{\tilde{\mathbf{s}}}_{\text{opt}}\right]_{1:NP,1} \quad (16)$$

where, using (13), (14), and (15) in (10), $\hat{\tilde{\mathbf{s}}}_{\text{opt}}$ is given by (17). In the sequel, we show that (17) is order-$2D$ polynomially solvable when the matrix inside the quadratic form of (17) -up to diagonal manipulations- has at most D nonzero eigenvalues that are also positive. For this purpose, we present the following two lemmas.

Lemma 1: Every matrix $\mathbf{A} \in \mathbb{C}^{N\times N}$ is binary-quadratic-form-optimization-equivalent (BQFO-equivalent) to $\dot{\mathbf{A}} = \mathbf{A} +$

[1]Operator $\text{vec}(\cdot)$ accounts for column-by-column vectorization of a matrix.

$\mathrm{diag}(\mathbf{x})$, for any $\mathbf{x} \in \mathbb{R}^{N\times 1}$, i.e. $\mathbf{s}^T\mathbf{A}\mathbf{s}$ and $\mathbf{s}^T\dot{\mathbf{A}}\mathbf{s}$ are maximized (minimized) by the same binary vector $\mathbf{s} \in \{\pm 1\}^N$.

Proof: For any $\mathbf{s} \in \{\pm 1\}^N$, $\mathbf{s}^T\dot{\mathbf{A}}\mathbf{s} = \mathbf{s}^T(\mathbf{A} + \mathrm{diag}(\mathbf{x}))\mathbf{s} = \mathbf{s}^T\mathbf{A}\mathbf{s} + \sum_{n=1}^{N} x_n s_n^2 = \mathbf{s}^T\mathbf{A}\mathbf{s} + \sum_{n=1}^{N} x_n$ where $\sum_{n=1}^{N} x_n$ is a real constant that does not affect the maximization (minimization) of the quadratic form. □

Lemma 2: Let $\mathbf{B} \in \mathbb{C}^{(N-1)\times M}$, $\mathbf{C} \in \mathbb{C}^{M\times D}$, $\mathbf{x} \in \mathbb{C}^{M\times 1}$, and

$$\mathbf{A} \triangleq \begin{bmatrix} \mathbf{B}\mathbf{C}\mathbf{C}^H\mathbf{B}^H & \mathbf{B}\left(\mathbf{I}_M - \mathbf{C}\mathbf{C}^H\right)\mathbf{x} \\ \mathbf{x}^H\left(\mathbf{I}_M - \mathbf{C}\mathbf{C}^H\right)\mathbf{B}^H & 0 \end{bmatrix} \in \mathbb{C}^{N\times N}. \tag{18}$$

If $\mathbf{x}$ is in the range of $\mathbf{C}$, i.e. $\mathbf{x} = \mathbf{C}\mathbf{a}$, $\mathbf{a} \in \mathbb{C}^{D\times 1}$, then $\mathbf{A}$ is BQFO-equivalent to the positive (semi)definite matrix

$$\dot{\mathbf{A}} \triangleq \begin{bmatrix} \mathbf{B}\mathbf{C} \\ \mathbf{a}^H\left(\mathbf{I}_D - \mathbf{C}^H\mathbf{C}\right) \end{bmatrix} \begin{bmatrix} \mathbf{C}^H\mathbf{B}^H & \left(\mathbf{I}_D - \mathbf{C}^H\mathbf{C}\right)\mathbf{a} \end{bmatrix}. \tag{19}$$

Proof: If $\mathbf{x} = \mathbf{C}\mathbf{a}$, $\mathbf{a} \in \mathbb{C}^{D\times 1}$, then

$$\begin{aligned}
\mathbf{A} &= \begin{bmatrix} \mathbf{B}\mathbf{C}\mathbf{C}^H\mathbf{B}^H & \mathbf{B}\left(\mathbf{I}_M - \mathbf{C}\mathbf{C}^H\right)\mathbf{C}\mathbf{a} \\ \mathbf{a}^H\mathbf{C}^H\left(\mathbf{I}_M - \mathbf{C}\mathbf{C}^H\right)\mathbf{B}^H & 0 \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{B}\mathbf{C}\mathbf{C}^H\mathbf{B}^H & \mathbf{B}\left(\mathbf{C} - \mathbf{C}\mathbf{C}^H\mathbf{C}\right)\mathbf{a} \\ \mathbf{a}^H\left(\mathbf{C}^H - \mathbf{C}^H\mathbf{C}\mathbf{C}^H\right)\mathbf{B}^H & 0 \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{B}\mathbf{C}\mathbf{C}^H\mathbf{B}^H & \mathbf{B}\mathbf{C}\left(\mathbf{I}_D - \mathbf{C}^H\mathbf{C}\right)\mathbf{a} \\ \mathbf{a}^H\left(\mathbf{I}_D - \mathbf{C}^H\mathbf{C}\right)\mathbf{C}^H\mathbf{B}^H & 0 \end{bmatrix} \\
&= \dot{\mathbf{A}} + \mathrm{diag}\left(\begin{bmatrix} \mathbf{0}_{(N-1)\times 1} \\ -\left\|\left(\mathbf{I}_D - \mathbf{C}^H\mathbf{C}\right)\mathbf{a}\right\|^2 \end{bmatrix}\right).
\end{aligned} \tag{20}$$

Due to Lemma 1, $\mathbf{A}$ is BQFO-equivalent to $\dot{\mathbf{A}}$. □

If $\boldsymbol{\mu}$ is in the range of $\mathbf{C}_h$, then, due to Lemma 2, we can rewrite (17) as

$$\begin{aligned}
\hat{\tilde{\mathbf{s}}}_{\text{opt}} &= \underset{\substack{\tilde{\mathbf{s}}\in\{\pm 1\}^{NP+1} \\ \tilde{s}_{NP+1}=1}}{\arg\max}\ \tilde{\mathbf{s}}^T\left[\Re\{\mathbf{A}\}\ \Im\{\mathbf{A}\}\right]\left[\Re\{\mathbf{A}\}\ \Im\{\mathbf{A}\}\right]^T\tilde{\mathbf{s}} \\
&= \underset{\substack{\tilde{\mathbf{s}}\in\{\pm 1\}^{NP+1} \\ \tilde{s}_{NP+1}=1}}{\arg\max}\ \left\|\mathbf{V}^T\tilde{\mathbf{s}}\right\|
\end{aligned} \tag{21}$$

where

$$\mathbf{A} \triangleq \begin{bmatrix} \frac{1}{\sigma_v}\mathbf{Z}\mathbf{E}^T\left(\mathbf{I}_{M_tP} \otimes \mathbf{Y}^H\right)\mathbf{Q} \\ \sigma_v\boldsymbol{\mu}^H\mathbf{Q}\boldsymbol{\Sigma}^{-1} \end{bmatrix}\left(\mathbf{I}_D + \frac{N}{\sigma_v^2}\boldsymbol{\Sigma}\right)^{-\frac{1}{2}} \in \mathbb{C}^{(NP+1)\times D} \tag{22}$$

and $\mathbf{V} \triangleq \left[\Re\{\mathbf{A}\}\ \ \Im\{\mathbf{A}\}\right] \in \mathbb{R}^{(NP+1)\times 2D}$. A proof of (21) is provided in the Appendix. The computation of $\hat{\tilde{\mathbf{s}}}_{\text{opt}}$ in (21) (hence, $\hat{\mathbf{s}}_{\text{opt}}$ in (16)) can be implemented with complexity $\mathcal{O}\left(P^{2D}\right)$ if we follow the multiple-auxiliary-angle methodology that has been introduced in [33]-[36] for the problem of rank-deficient quadratic form maximization and is presented below.

We introduce the spherical variables $\phi_1 \in (-\pi, \pi]$, $\phi_2, \ldots, \phi_{2D-1} \in (-\frac{\pi}{2}, \frac{\pi}{2}]$ and define the spherical variable vector $\boldsymbol{\phi} \triangleq [\phi_1, \phi_2, \ldots, \phi_{2D-1}]^T$ and the $2D \times 1$ hyperspherical vector

$$\mathbf{c}(\boldsymbol{\phi}) \triangleq \begin{bmatrix} \sin\phi_1 \\ \cos\phi_1\sin\phi_2 \\ \vdots \\ \cos\phi_1\ldots\cos\phi_{2D-2}\sin\phi_{2D-1} \\ \cos\phi_1\ldots\cos\phi_{2D-2}\cos\phi_{2D-1} \end{bmatrix}. \tag{23}$$

Due to Cauchy-Schwartz Inequality, for the optimization problem in (21) we obtain

$$\begin{aligned}
&\max_{\tilde{\mathbf{s}}\in\{\pm 1\}^{NP+1}} \left\|\mathbf{V}^T\tilde{\mathbf{s}}\right\| \\
&= \max_{\tilde{\mathbf{s}}\in\{\pm 1\}^{NP+1}}\ \max_{\phi_1\in[-\pi,\pi]}\ \max_{\phi_2,\ldots,\phi_{2D-1}\in(-\frac{\pi}{2},\frac{\pi}{2}]} \left\{\tilde{\mathbf{s}}^T\mathbf{V}\mathbf{c}(\boldsymbol{\phi})\right\} \\
&= \max_{\phi_1\in(-\pi,\pi]}\ \max_{\phi_2,\ldots,\phi_{2D-1}\in(-\frac{\pi}{2},\frac{\pi}{2}]}\ \sum_{n=1}^{NP+1}\ \max_{\tilde{s}_n=\pm 1}\left\{\tilde{s}_n\mathbf{V}_{n,:}\mathbf{c}(\boldsymbol{\phi})\right\}
\end{aligned} \tag{24}$$

by interchanging the maximizations in (24). Eq. (24) shows that $\hat{\tilde{\mathbf{s}}}_{\text{opt}}$ can be obtained by determining the optimal vector $\hat{\tilde{\mathbf{s}}}(\boldsymbol{\phi})$ for every point $\boldsymbol{\phi} \in (-\pi, \pi] \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]^{2D-2}$. We note that, for any point $\boldsymbol{\phi}$, the maximizing argument of each term of the sum in (24) depends only on the corresponding row of $\mathbf{V}$ and is determined by $\tilde{s}_n = \mathrm{sgn}\left(\mathbf{V}_{n,:}\,\mathbf{c}(\boldsymbol{\phi})\right), n = 1, \ldots, NP+1$. Then, for the given $(NP+1)\times 2D$ matrix $\mathbf{V}$, each point $\boldsymbol{\phi} \in (-\pi, \pi] \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]^{2D-2}$ is associated with a candidate binary vector $\tilde{\mathbf{s}}(\boldsymbol{\phi}) = \mathrm{sgn}(\mathbf{V}\mathbf{c}(\boldsymbol{\phi}))$ and the optimal vector $\hat{\tilde{\mathbf{s}}}_{\text{opt}}$ in (21) is met if we scan the entire set $(-\pi, \pi] \times (-\frac{\pi}{2}, \frac{\pi}{2}]^{2D-2}$. Moreover, opposite binary vectors $\tilde{\mathbf{s}}$ and $-\tilde{\mathbf{s}}$ result in the same metric in (21), thus, we can ignore the values of ϕ_1 in $\left(-\pi, -\frac{\pi}{2}\right] \cup \left(\frac{\pi}{2}, \pi\right]$ and rewrite the optimization problem in (24) as $\max_{\boldsymbol{\phi}\in\Phi}\sum_{n=1}^{NP+1}\max_{\tilde{s}_n=\pm 1}\left\{\tilde{s}_n\mathbf{V}_{n,:}\mathbf{c}(\boldsymbol{\phi})\right\}$ where $\Phi \triangleq (-\frac{\pi}{2}, \frac{\pi}{2}]^{2D-1}$. Finally, we collect all candidate binary vectors into set $\mathcal{S} \triangleq \bigcup_{\boldsymbol{\phi}\in\Phi}\left\{\tilde{\mathbf{s}}(\boldsymbol{\phi})\right\} \subseteq \{\pm 1\}^{NP+1}$ and observe that $\underset{\tilde{\mathbf{s}}\in\{\pm 1\}^{NP+1}}{\arg\max}\left\|\mathbf{V}^T\tilde{\mathbf{s}}\right\| \in \mathcal{S}$, hence $\hat{\tilde{\mathbf{s}}}_{\text{opt}} = \hat{\tilde{\mathbf{s}}}'_{\text{opt}} \cdot \left[\hat{\tilde{\mathbf{s}}}'_{\text{opt}}\right]_{NP+1,1}$ where $\hat{\tilde{\mathbf{s}}}'_{\text{opt}} = \underset{\tilde{\mathbf{s}}\in\mathcal{S}}{\arg\max}\left\|\mathbf{V}^T\tilde{\mathbf{s}}\right\|$.

In [35], it was shown that the decision $\tilde{s}_n = \mathrm{sgn}\left(\mathbf{V}_{n,:}\,\mathbf{c}(\boldsymbol{\phi})\right)$ is equivalent to

$$\tilde{s}_n = \begin{cases} -\mathrm{sgn}(V_{n,1}), & \phi_1 \in \left(-\frac{\pi}{2}, \tan^{-1}\!\left(-\frac{\mathbf{V}_{n,2:2D}\mathbf{c}(\boldsymbol{\phi}_{2:2D-1})}{V_{n,1}}\right)\right] \\ \mathrm{sgn}(V_{n,1}), & \phi_1 \in \left(\tan^{-1}\!\left(-\frac{\mathbf{V}_{n,2:2D}\mathbf{c}(\boldsymbol{\phi}_{2:2D-1})}{V_{n,1}}\right), \frac{\pi}{2}\right] \end{cases}. \tag{25}$$

The function $\phi_1 = \tan^{-1}\left(-\frac{\mathbf{V}_{n,2:2D}\mathbf{c}(\boldsymbol{\phi}_{2:2D-1})}{V_{n,1}}\right)$ determines a hypersurface which partitions Φ into two regions. One region corresponds to $\tilde{s}_n = -1$ and the other one corresponds to $\tilde{s}_n = +1$. Therefore, the $NP+1$ rows of $\mathbf{V}$ are associated with $NP+1$ corresponding hypersurfaces that partition the hypercube Φ into K cells $C_1, C_2, \ldots, C_K$ such that $\bigcup_{k=1}^{K} C_k = \Phi$, $C_k \cap C_j = \emptyset\ \forall\ k \neq j$, and each cell C_k corresponds to a distinct vector $\tilde{\mathbf{s}}_k \in \{\pm 1\}^{NP+1}$. See, for example, Fig. 1 where we present the cells and associated candidate vectors that are formed by an eight-row three-column matrix $\mathbf{V}$ and a spherical vector $\boldsymbol{\phi} = [\phi_1\ \phi_2]^T$.[2] In [35], it was shown that $K = \sum_{d=0}^{2D-1}\binom{NP}{d}$, therefore all candidate vectors form set $\mathcal{S}$ with cardinality $|\mathcal{S}| = \sum_{d=0}^{2D-1}\binom{NP}{d} = \mathcal{O}\left(P^{2D-1}\right)$. The construction of $\mathcal{S}$ is of special interest since it determines the overall performance of the proposed method. An algorithm for the construction of $\mathcal{S}$ was developed in [35] and is available at http://www.telecom.tuc.gr/∼karystinos. Interestingly, the algorithm's complexity for the construction of $\mathcal{S}$ is $\mathcal{O}(P^{2D})$ for any given matrix $\mathbf{V}$.

[2]We selected an odd number of columns for matrix $\mathbf{V}$ and even number of elements for $\boldsymbol{\phi}$ just for illustration purposes. It should be mentioned that $\mathbf{V}$ and $\boldsymbol{\phi}$ always consist of an even number of columns and odd number of elements, respectively.

In our developments in this section, $\mathbf{V}$ (which is a function of $\mathbf{A}$) has to be computed by the receiver and subsequently fed to the algorithm of [35]. In (22), it is seen that $\mathbf{A}$ is a function of the received data matrix $\mathbf{Y}$, matrices $\mathbf{Z}$, $\mathbf{E}$, $\mathbf{Q}$, and $\mathbf{\Sigma}$, vector $\boldsymbol{\mu}$, and scalar σ_v. We note that

$$\begin{aligned}
&\mathbf{Z}\mathbf{E}^T\left(\mathbf{I}_{M_\mathrm{t}P}\otimes\mathbf{Y}^H\right)\mathbf{Q}\\
&=\mathbf{Z}\left[\mathbf{I}_{M_\mathrm{t}}\otimes\mathbf{e}_1\mathbf{e}_1^T\otimes\mathbf{I}_T\ \ldots\ \mathbf{I}_{M_\mathrm{t}}\otimes\mathbf{e}_P\mathbf{e}_P^T\otimes\mathbf{I}_T\right]\left(\mathbf{I}_P\otimes\mathbf{I}_{M_\mathrm{t}}\otimes\mathbf{Y}^H\right)\mathbf{Q}\\
&=\mathbf{Z}\left[\left(\mathbf{I}_{M_\mathrm{t}}\otimes\mathbf{e}_1\mathbf{e}_1^T\otimes\mathbf{I}_T\right)\left(\mathbf{I}_{M_\mathrm{t}}\otimes\mathbf{Y}^H\right)\right.\\
&\qquad\left.\ldots\left(\mathbf{I}_{M_\mathrm{t}}\otimes\mathbf{e}_P\mathbf{e}_P^T\otimes\mathbf{I}_T\right)\left(\mathbf{I}_{M_\mathrm{t}}\otimes\mathbf{Y}^H\right)\right]\mathbf{Q}\\
&=\mathbf{Z}\left[\mathbf{I}_{M_\mathrm{t}}\otimes\left(\mathbf{e}_1\mathbf{e}_1^T\otimes\mathbf{I}_T\right)\mathbf{Y}^H\ \ldots\ \mathbf{I}_{M_\mathrm{t}}\otimes\left(\mathbf{e}_P\mathbf{e}_P^T\otimes\mathbf{I}_T\right)\mathbf{Y}^H\right]\\
&\qquad\times\begin{bmatrix}[\mathbf{Q}]_{1:M_\mathrm{t}M_\mathrm{r},:}\\ \vdots\\ [\mathbf{Q}]_{(P-1)M_\mathrm{t}M_\mathrm{r}+1:PM_\mathrm{t}M_\mathrm{r},:}\end{bmatrix}\\
&=\sum_{p=1}^{P}\left[\mathbf{Z}_1\ \ldots\ \mathbf{Z}_{M_\mathrm{t}}\right]\left(\mathbf{I}_{M_\mathrm{t}}\otimes\left(\mathbf{e}_p\mathbf{e}_p^T\otimes\mathbf{I}_T\right)\mathbf{Y}^H\right)\mathbf{Q}_p \qquad (26)
\end{aligned}$$

where $\mathbf{Q}_p \triangleq [\mathbf{Q}]_{(p-1)M_\mathrm{t}M_\mathrm{r}+1:pM_\mathrm{t}M_\mathrm{r},:}$. We observe that, for any $p=1,2,\ldots,P$,

$$\begin{aligned}
&\left[\mathbf{Z}_1\ \ldots\ \mathbf{Z}_{M_\mathrm{t}}\right]\left(\mathbf{I}_{M_\mathrm{t}}\otimes\left(\mathbf{e}_p\mathbf{e}_p^T\otimes\mathbf{I}_T\right)\mathbf{Y}^H\right)\mathbf{Q}_p\\
&=\left[\mathbf{I}_P\otimes\tilde{\mathbf{X}}_1\ \ldots\ \mathbf{I}_P\otimes\tilde{\mathbf{X}}_{M_\mathrm{t}}\right]\left(\mathbf{I}_{M_\mathrm{t}}\otimes\left(\mathbf{e}_p\mathbf{e}_p^T\otimes\mathbf{I}_T\right)\mathbf{Y}^H\right)\mathbf{Q}_p\\
&=\left[\left(\mathbf{I}_P\otimes\tilde{\mathbf{X}}_1\right)\left(\mathbf{e}_p\mathbf{e}_p^T\otimes\mathbf{I}_T\right)\mathbf{Y}^H\ldots\left(\mathbf{I}_P\otimes\tilde{\mathbf{X}}_{M_\mathrm{t}}\right)\left(\mathbf{e}_p\mathbf{e}_p^T\otimes\mathbf{I}_T\right)\mathbf{Y}^H\right]\mathbf{Q}_p\\
&=\left[\left(\mathbf{e}_p\mathbf{e}_p^T\otimes\tilde{\mathbf{X}}_1\right)\mathbf{Y}^H\ \ldots\ \left(\mathbf{e}_p\mathbf{e}_p^T\otimes\tilde{\mathbf{X}}_{M_\mathrm{t}}\right)\mathbf{Y}^H\right]\mathbf{Q}_p\\
&=\sum_{m=1}^{M_\mathrm{t}}\left(\mathbf{e}_p\mathbf{e}_p^T\otimes\tilde{\mathbf{X}}_m\right)\mathbf{Y}^H\left[\mathbf{Q}_p\right]_{(m-1)M_\mathrm{r}+1:mM_\mathrm{r},:}\\
&=\sum_{m=1}^{M_\mathrm{t}}\begin{bmatrix}\mathbf{0}_{(p-1)N\times M_\mathrm{r}}\\ \tilde{\mathbf{X}}_m\left[\mathbf{Y}^H\right]_{(p-1)T+1:pT,:}\\ \mathbf{0}_{(P-p)N\times M_\mathrm{r}}\end{bmatrix}\left[\mathbf{Q}_p\right]_{(m-1)M_\mathrm{r}+1:mM_\mathrm{r},:}\\
&=\begin{bmatrix}\mathbf{0}_{(p-1)N\times D}\\ \sum_{m=1}^{M_\mathrm{t}}\tilde{\mathbf{X}}_m\left[\mathbf{Y}^H\right]_{(p-1)T+1:pT,:}\left[\mathbf{Q}_p\right]_{(m-1)M_\mathrm{r}+1:mM_\mathrm{r},:}\\ \mathbf{0}_{(P-p)N\times D}\end{bmatrix}. \qquad (27)
\end{aligned}$$

Computation of $\tilde{\mathbf{X}}_m\left[\mathbf{Y}^H\right]_{(p-1)T+1:pT,:}\left[\mathbf{Q}_p\right]_{(m-1)M_\mathrm{r}+1:mM_\mathrm{r},:}$, $m=1,\ldots,M_\mathrm{t}$, requires $\mathcal{O}(NTM_\mathrm{r}+NM_\mathrm{r}D)$ calculations and the sum in (27) consists of M_t such products resulting in a total of $\mathcal{O}(M_\mathrm{t}(NTM_\mathrm{r}+NM_\mathrm{r}D))$ calculations. In addition, (26) contains P such sums, hence the computational complexity of $\mathbf{Z}\mathbf{E}^T\left(\mathbf{I}_{M_\mathrm{t}P}\otimes\mathbf{Y}^H\right)\mathbf{Q}$ is $\mathcal{O}(PM_\mathrm{t}(NTM_\mathrm{r}+NM_\mathrm{r}D))$. Computation of the row vector $\sigma_v\boldsymbol{\mu}^H\mathbf{Q}\mathbf{\Sigma}^{-1}$ that appears in the bottom row of $\mathbf{A}$ requires $\mathcal{O}(M_\mathrm{t}M_\mathrm{r}PD+D)$ calculations. Finally, the multiplication of the leftmost matrix in (22) with $\left(\mathbf{I}_D+\frac{N}{\sigma_v^2}\mathbf{\Sigma}\right)^{-\frac{1}{2}}$ costs $\mathcal{O}((NP+1)D)$. Therefore, the overall complexity overhead for the computation of $\mathbf{V}$ becomes $\mathcal{O}(PM_\mathrm{t}(NTM_\mathrm{r}+NM_\mathrm{r}D)+M_\mathrm{t}M_\mathrm{r}PD+D+(NP+1)D)$ which is linear in the sequence length P for constant values of M_t, M_r, N, T, and D (that is, fixed number of antennas, space-time coding rate, and channel covariance rank). We conclude that the overall complexity of the proposed receiver is $\mathcal{O}\left(P^{2D}\right)$.

As a brief summary, the sequence of calculations of the proposed ML noncoherent receiver is as follows. The whole data record of TP received vectors is utilized to form the received matrix $\mathbf{Y}$. Then, matrix $\mathbf{V}$ is computed as a function of $\mathbf{A}$ in (22) with complexity $\mathcal{O}(P)$. Finally, the quadratic-form-maximization algorithm of [35] is operated on $\mathbf{V}$ to produce the optimal data bit sequence $\hat{\tilde{\mathbf{s}}}_{\mathrm{opt}}$ with complexity $\mathcal{O}\left(P^{2D}\right)$.

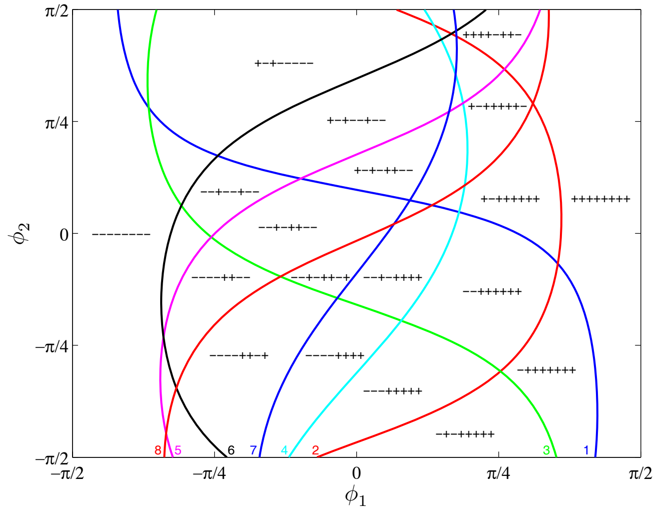

Fig. 1. Partition of Φ into cells.

IV. Special Channel Model Cases

In the previous section we considered a block-fading Ricean MIMO channel model and showed that ML noncoherent OSTBC detection is attained with polynomial complexity if the mean channel vector is in the range of the channel covariance matrix and the rank of the covariance matrix is not a function of the sequence length; provided the latter condition, the polynomial complexity of our proposed optimal receiver depends strictly on the rank of the channel covariance matrix. Since the static Ricean, block-fading Rayleigh, and static Rayleigh channel models are special cases of the general model that we considered, we immediately conclude that polynomial ML noncoherent detection complexity is also attained provided that the same conditions with the block-fading Ricean channel model case hold. Especially for (block-fading or static) Rayleigh fading, the channel mean is zero which always is in the range of the channel covariance matrix, hence polynomial-complexity ML noncoherent detection is always attained provided that the channel covariance rank is not a function of the sequence length. In this section, we examine separately the three special cases of the general model of the previous section, identify conditions for polynomial solvability of the ML noncoherent detection problem, and report exact complexity requirements of the proposed polynomial-complexity ML noncoherent detector. An interesting outcome of our analysis is that *polynomial-complexity ML noncoherent detection is always feasible for the static Rayleigh channel through the proposed algorithm* and -in contrast to Ricean fading- the complexity of the optimal receiver is reduced if the channel covariance rank is lower.

Case I: Block-fading Rayleigh channel

Due to Rayleigh fading, we have $\boldsymbol{\mu}=\mathbf{0}$, hence the bottom row

$$
\begin{aligned}
\hat{\mathbf{s}}_{\text{opt}} &= \arg\max_{\mathbf{s}\in\{\pm1\}^{NP}} \left\{\mathbf{y}^H \left(\mathbf{D}^T(\mathbf{s})\otimes \mathbf{I}_{M_\text{r}}\right)\left(\mathbf{1}_P\otimes \mathbf{I}_{M_\text{t}M_\text{r}}\right)\left(\mathbf{1}_P^T\otimes \mathbf{I}_{M_\text{t}M_\text{r}}\right)\left(\mathbf{D}^*(\mathbf{s})\otimes \mathbf{I}_{M_\text{r}}\right)\mathbf{y}\right\} \\
&= \arg\max_{\mathbf{s}\in\{\pm1\}^{NP}} \left\{\mathbf{y}^H \left(\mathbf{D}^T(\mathbf{s})\left(\mathbf{1}_P\otimes \mathbf{I}_{M_\text{t}}\right)\otimes \mathbf{I}_{M_\text{r}}\right)\left(\left(\mathbf{1}_P^T\otimes \mathbf{I}_{M_\text{t}}\right)\mathbf{D}^*(\mathbf{s})\otimes \mathbf{I}_{M_\text{r}}\right)\mathbf{y}\right\} \\
&= \arg\max_{\mathbf{s}\in\{\pm1\}^{NP}} \left\{\mathbf{y}^H \left(\mathbf{G}^T(\mathbf{s})\otimes \mathbf{I}_{M_\text{r}}\right)\left(\mathbf{G}^*(\mathbf{s})\otimes \mathbf{I}_{M_\text{r}}\right)\mathbf{y}\right\} = \arg\max_{\mathbf{s}\in\{\pm1\}^{NP}} \left\|\left(\mathbf{G}^*(\mathbf{s})\otimes \mathbf{I}_{M_\text{r}}\right)\text{vec}(\mathbf{Y})\right\|^2 \\
&\overset{(34)}{=} \arg\max_{\mathbf{s}\in\{\pm1\}^{NP}} \left\|\text{vec}\left(\mathbf{Y}\mathbf{G}^H(\mathbf{s})\right)\right\|^2 = \arg\max_{\mathbf{s}\in\{\pm1\}^{NP}} \left\|\mathbf{Y}\mathbf{G}^H(\mathbf{s})\right\|_\text{F}^2 = \arg\max_{\mathbf{s}\in\{\pm1\}^{NP}} \text{tr}\left\{\mathbf{Y}\mathbf{G}^H(\mathbf{s})\mathbf{G}(\mathbf{s})\mathbf{Y}^H\right\}. \qquad (29)
\end{aligned}
$$

of $\mathbf{A}$ in (22) becomes zero and the ML detector of (21) simplifies to $\hat{\mathbf{s}}_{\text{opt}} = \arg\max_{\mathbf{s}\in\{\pm1\}^{NP}} \left\|\mathbf{V}_1^T\mathbf{s}\right\|$ where $\mathbf{V}_1 \triangleq [\Re\{\mathbf{A}_1\}\ \Im\{\mathbf{A}_1\}]$ and $\mathbf{A}_1 \triangleq \mathbf{Z}\mathbf{E}^T\left(\mathbf{I}_{M_\text{t}P}\otimes \mathbf{Y}^H\right)\mathbf{Q}\left(\mathbf{I}_D + \frac{N}{\sigma_v^2}\mathbf{\Sigma}\right)^{-\frac{1}{2}}$.

Case II: Static Ricean channel

We have $\mathbf{h} = \mathbf{1}_P\otimes \underline{\mathbf{h}}$ where $\underline{\mathbf{h}} = \text{vec}\left(\mathbf{H}^{(1)}\right)$ is a circular complex Gaussian vector with mean $\underline{\boldsymbol{\mu}}$ and covariance matrix $\mathbf{C}_{\underline{h}} = \underline{\mathbf{Q}}\,\underline{\mathbf{Q}}^H$. Then, the covariance matrix of the channel vector $\mathbf{h}$ equals $\mathbf{C}_h = \mathbf{1}_P\mathbf{1}_P^T\otimes \mathbf{C}_{\underline{h}} = \mathbf{1}_P\mathbf{1}_P^T\otimes \underline{\mathbf{Q}}\,\underline{\mathbf{Q}}^H = (\mathbf{1}_P\otimes\underline{\mathbf{Q}})(\mathbf{1}_P\otimes\underline{\mathbf{Q}})^H$, therefore $\mathbf{Q} = \mathbf{1}_P\otimes\underline{\mathbf{Q}}$ and

$$
\begin{aligned}
\mathbf{Q}^H(\mathbf{I}_{M_\text{t}P}\otimes\mathbf{Y})\mathbf{E} &= \left(\mathbf{1}_P^T\otimes\underline{\mathbf{Q}}^H\right)(\mathbf{I}_P\otimes\mathbf{I}_{M_\text{t}}\otimes\mathbf{Y})\begin{bmatrix}\mathbf{I}_{M_\text{t}}\otimes\mathbf{e}_1\mathbf{e}_1^T\otimes\mathbf{I}_T\\ \vdots \\ \mathbf{I}_{M_\text{t}}\otimes\mathbf{e}_P\mathbf{e}_P^T\otimes\mathbf{I}_T\end{bmatrix} \\
&= \left(\mathbf{1}_P^T\otimes\underline{\mathbf{Q}}^H(\mathbf{I}_{M_\text{t}}\otimes\mathbf{Y})\right)\begin{bmatrix}\mathbf{I}_{M_\text{t}}\otimes\mathbf{e}_1\mathbf{e}_1^T\otimes\mathbf{I}_T\\ \vdots \\ \mathbf{I}_{M_\text{t}}\otimes\mathbf{e}_P\mathbf{e}_P^T\otimes\mathbf{I}_T\end{bmatrix} = \sum_{p=1}^{P}\underline{\mathbf{Q}}^H(\mathbf{I}_{M_\text{t}}\otimes\mathbf{Y}) \\
&\times\left(\mathbf{I}_{M_\text{t}}\otimes\mathbf{e}_p\mathbf{e}_p^T\otimes\mathbf{I}_T\right) = \underline{\mathbf{Q}}^H(\mathbf{I}_{M_\text{t}}\otimes\mathbf{Y})\left(\mathbf{I}_{M_\text{t}}\otimes\left(\sum_{p=1}^{P}\mathbf{e}_p\mathbf{e}_p^T\right)\otimes\mathbf{I}_T\right) \\
&= \underline{\mathbf{Q}}^H\left(\mathbf{I}_{M_\text{t}}\otimes\mathbf{Y}\right)\left(\mathbf{I}_{M_\text{t}}\otimes\mathbf{I}_P\otimes\mathbf{I}_T\right) = \underline{\mathbf{Q}}^H\left(\mathbf{I}_{M_\text{t}}\otimes\mathbf{Y}\right). \qquad (28)
\end{aligned}
$$

In addition, $\mathbf{\Sigma} = \mathbf{Q}^H\mathbf{Q} = (\mathbf{1}_P^T\otimes\underline{\mathbf{Q}}^H)(\mathbf{1}_P\otimes\underline{\mathbf{Q}}) = \mathbf{1}_P^T\mathbf{1}_P\otimes \underline{\mathbf{Q}}^H\underline{\mathbf{Q}} = P\otimes\underline{\mathbf{\Sigma}} = P\underline{\mathbf{\Sigma}}$, where $\underline{\mathbf{\Sigma}} \triangleq \underline{\mathbf{Q}}^H\underline{\mathbf{Q}}$, and $\boldsymbol{\mu}^H\mathbf{Q}\mathbf{\Sigma}^{-1} = \left(\mathbf{1}_P^T\otimes\underline{\boldsymbol{\mu}}^H\right)\left(\mathbf{1}_P\otimes\underline{\mathbf{Q}}\right)\frac{1}{P}\underline{\mathbf{\Sigma}}^{-1} = \frac{1}{P}\left(P\otimes\underline{\boldsymbol{\mu}}^H\underline{\mathbf{Q}}\right)\underline{\mathbf{\Sigma}}^{-1} = \underline{\boldsymbol{\mu}}^H\underline{\mathbf{Q}}\underline{\mathbf{\Sigma}}^{-1}$. Then, the ML detector in (21) simplifies to $\hat{\tilde{\mathbf{s}}}_{\text{opt}} = \arg\max_{\substack{\tilde{\mathbf{s}}\in\{\pm1\}^{NP+1}\\ \tilde{s}_{NP+1}=1}} \left\|\mathbf{V}_2^T\tilde{\mathbf{s}}\right\|$ where $\mathbf{V}_2 \triangleq [\Re\{\mathbf{A}_2\}\ \Im\{\mathbf{A}_2\}]$ and $\mathbf{A}_2 \triangleq \begin{bmatrix}\frac{1}{\sigma_v}\mathbf{Z}\left(\mathbf{I}_{M_\text{t}}\otimes\mathbf{Y}^H\right)\underline{\mathbf{Q}}\\ \sigma_v\underline{\boldsymbol{\mu}}^H\underline{\mathbf{Q}}\underline{\mathbf{\Sigma}}^{-1}\end{bmatrix}\left(\mathbf{I}_D + \frac{NP}{\sigma_v^2}\underline{\mathbf{\Sigma}}\right)^{-\frac{1}{2}}$. Of course, such a simplification is possible when $\underline{\boldsymbol{\mu}}$ is in the range of $\mathbf{C}_{\underline{h}}$.

Case III: Static Rayleigh channel

We have $\underline{\boldsymbol{\mu}} = \mathbf{0}$, hence the ML detector becomes $\hat{\mathbf{s}}_{\text{opt}} = \arg\max_{\mathbf{s}\in\{\pm1\}^{NP}} \left\|\mathbf{V}_3^T\mathbf{s}\right\|$ where $\mathbf{V}_3 \triangleq [\Re\{\mathbf{A}_3\}\ \Im\{\mathbf{A}_3\}]$ and $\mathbf{A}_3 \triangleq \mathbf{Z}\left(\mathbf{I}_{M_\text{t}}\otimes\mathbf{Y}^H\right)\underline{\mathbf{Q}}\left(\mathbf{I}_D + \frac{NP}{\sigma_v^2}\underline{\mathbf{\Sigma}}\right)^{-\frac{1}{2}}$. Note that the latter is always feasible, since $\boldsymbol{\mu} = \mathbf{0}$ always is in the range of $\mathbf{C}_h$.

It is interesting to mention that the ML noncoherent detector in the case of static Rayleigh fading (Case III) simplifies to the popular trace detector [5], [10] in the special cases of i.i.d. channel coefficients or joint channel estimation and data detection due to channel statistics uncertainty at the receiver. Indeed, if the channel coefficients are i.i.d., then $\mathbf{C}_{\underline{h}}$, $\underline{\mathbf{Q}}$, and $\underline{\mathbf{\Sigma}}$ are scaled versions of $\mathbf{I}_{M_\text{t}M_\text{r}}$, $D = M_\text{t}M_\text{r}$, and the ML detector in (10) becomes as in (29).

In the second special case, the receiver does not have knowledge of the channel statistics, hence joint ML estimation of $\underline{\mathbf{H}}$ and detection of $\mathbf{s}$ is performed according to $\widehat{\{\underline{\mathbf{H}},\mathbf{s}\}} = \arg\min_{\substack{\underline{\mathbf{H}}\in\mathbb{C}^{M_\text{r}\times M_\text{t}}\\ \mathbf{s}\in\{\pm1\}^{NP}}} \|\mathbf{Y}-\underline{\mathbf{H}}\mathbf{G}(\mathbf{s})\|_\text{F}^2$. Then,

$$
\hat{\mathbf{s}}_{\text{GLRT}} = \arg\min_{\mathbf{s}\in\{\pm1\}^{NP}}\left\{\min_{\underline{\mathbf{H}}\in\mathbb{C}^{M_\text{r}\times M_\text{t}}}\|\mathbf{Y}-\underline{\mathbf{H}}\mathbf{G}(\mathbf{s})\|_\text{F}^2\right\} \qquad (30)
$$

is the generalized-likelihood-ratio-test (GLRT) detection of $\mathbf{s}$ [17], [39]. For any $\mathbf{s}\in\{\pm1\}^{NP}$, the inner minimization in (30) results in

$$
\hat{\underline{\mathbf{H}}}(\mathbf{s}) = \arg\min_{\underline{\mathbf{H}}\in\mathbb{C}^{M_\text{r}\times M_\text{t}}}\|\mathbf{Y}-\underline{\mathbf{H}}\mathbf{G}(\mathbf{s})\|_\text{F}^2 = \frac{1}{NP}\mathbf{Y}\mathbf{G}^H(\mathbf{s}) \qquad (31)
$$

which is obtained by setting $\frac{\partial}{\partial\underline{\mathbf{H}}}\|\mathbf{Y}-\underline{\mathbf{H}}\mathbf{G}(\mathbf{s})\|_\text{F}^2 = 0$ and solving for $\underline{\mathbf{H}}$ using matrix differentiation [40] that gives $\frac{\partial}{\partial\underline{\mathbf{H}}}\|\mathbf{Y}-\underline{\mathbf{H}}\mathbf{G}(\mathbf{s})\|_\text{F}^2 = NP\underline{\mathbf{H}}^* - \mathbf{Y}^*\mathbf{G}^T(\mathbf{s})$. Using (31), the GLRT decision in (30) becomes

$$
\begin{aligned}
\hat{\mathbf{s}}_{\text{GLRT}} &= \arg\min_{\mathbf{s}\in\{\pm1\}^{NP}} \text{tr}\left\{\left(\mathbf{Y}-\frac{1}{NP}\mathbf{Y}\mathbf{G}^H(\mathbf{s})\mathbf{G}(\mathbf{s})\right)\right. \\
&\qquad\left.\times\left(\mathbf{Y}-\frac{1}{NP}\mathbf{Y}\mathbf{G}^H(\mathbf{s})\mathbf{G}(\mathbf{s})\right)^H\right\} \\
&= \arg\min_{\mathbf{s}\in\{\pm1\}^{NP}}\left\{-\frac{2}{NP}\text{tr}\left\{\mathbf{Y}\mathbf{G}^H(\mathbf{s})\mathbf{G}(\mathbf{s})\mathbf{Y}^H\right\}\right. \\
&\qquad\left.+\frac{1}{(NP)^2}\text{tr}\left\{\mathbf{Y}\mathbf{G}^H(\mathbf{s})\mathbf{G}(\mathbf{s})\mathbf{G}^H(\mathbf{s})\mathbf{G}(\mathbf{s})\mathbf{Y}^H\right\}\right\} \\
&= \arg\max_{\mathbf{s}\in\{\pm1\}^{NP}} \text{tr}\left\{\mathbf{Y}\mathbf{G}^H(\mathbf{s})\mathbf{G}(\mathbf{s})\mathbf{Y}^H\right\}. \qquad (32)
\end{aligned}
$$

Apparently, (29) and (32) are identical problems. Both constitute special cases of Case III and can be solved in polynomial time $\mathcal{O}(P^{2M_\text{t}M_\text{r}})$ if we follow the proposed approach.

We conclude that for static Rayleigh channels the ML noncoherent detector can always operate with polynomial complexity in the sequence length P, the order of which is determined strictly by the rank D of the channel covariance matrix. That is, the lower the Rayleigh channel covariance rank the lower the receiver complexity. Therefore, the worst-case complexity is $\mathcal{O}\left(P^{2M_\text{t}M_\text{r}}\right)$ and is met, for example, when the channel coefficients are i.i.d. Similar properties hold for block-fading Rayleigh channels, as long as the rank D of the correlation matrix is not a function of the sequence length P. Instead, for Ricean channel distribution, polynomial complexity is attained through the proposed receiver if the mean channel vector is in the range of the channel covariance

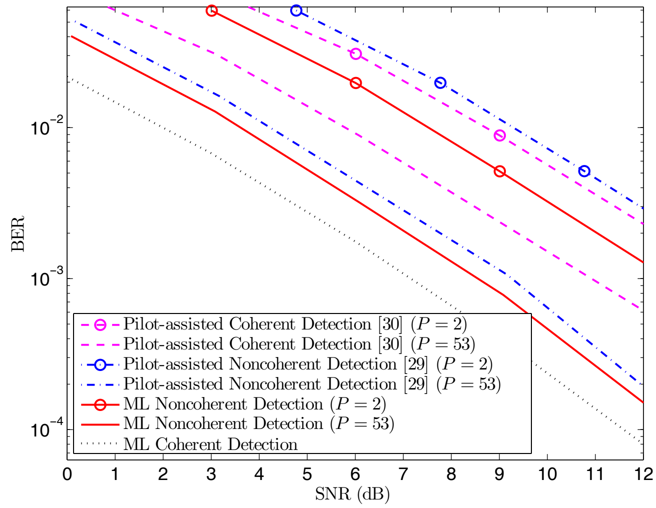

Fig. 2. BER versus SNR for ML and pilot-assisted [30] coherent OSTBC receivers and proposed ML and pilot-assisted [29] noncoherent OSTBC receivers with sequence length $P = 2$ (conventional receiver) and $P = 53$ (106 bits) upon Rayleigh fading.

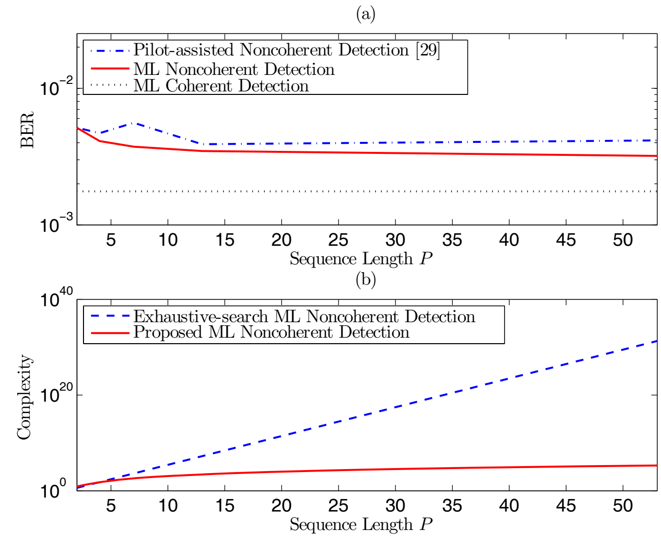

Fig. 3. (a) BER and (b) complexity versus sequence length P for SNR = 6dB and Rayleigh fading.

matrix. Hence, in contrast to Rayleigh fading where low-rank correlation is desired, full-rank correlation is sufficient to guarantee polynomial detection complexity in the case of static Ricean channels. In the following section, we illustrate our theoretic findings.

V. Simulation Studies

We consider a 2×2 MIMO 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 or Ricean fading channel environment. Space-time ambiguity induced by the rotatability of the Alamouti code [41] is resolved by employing differential space-time modulation [24] due to which the pth transmitted space-time block is $\mathbf{C}^{(p)} = \mathbf{C}^{(p-1)}\mathbf{F}\left(\mathbf{s}^{(p)}\right)$ where $\mathbf{F}\left(\mathbf{s}^{(p)}\right) = \begin{bmatrix} s_1^{(p)} & 0 \\ 0 & s_2^{(p)} \end{bmatrix}$ if $s_1^{(p)} s_2^{(p)} > 0$, $\mathbf{F}\left(\mathbf{s}^{(p)}\right) = \begin{bmatrix} 0 & s_2^{(p)} \\ -s_1^{(p)} & 0 \end{bmatrix}$ if $s_1^{(p)} s_2^{(p)} < 0$, and $\mathbf{C}^{(0)} = \left[\begin{smallmatrix} 1 & 1 \\ -1 & 1 \end{smallmatrix}\right]$ so that $\mathbf{C}^{(p)}$ follows the Alamouti code structure, for any $p = 0, 1, 2, \ldots$. For the covariance matrix of the vectorized channel matrix we adopt the model of [42], according to which $\mathbf{C}_{\underline{h}} = \begin{bmatrix} 1 & r & t & w_1 \\ r^* & 1 & w_2 & t \\ t^* & w_2^* & 1 & r \\ w_1^* & t^* & r^* & 1 \end{bmatrix}$. In our studies we set $t = r = 0$ and $w_1 = w_2 = 1$, a setup that exhibits higher ergodic capacity in comparison to the one with independent channel coefficients [42]. Observe that the rank of such a matrix is 2, therefore the overall complexity of the proposed ML receiver becomes $\mathcal{O}\left(P^4\right)$. We present results averaged over $1,000$ channel realizations.

In Fig. 2, we study the Rayleigh fading channel case and present the bit error rate (BER) of the one-shot coherent MRC receiver and the ML noncoherent receiver implemented with complexity $\mathcal{O}\left(P^4\right)$ by the proposed algorithm as a function of the information SNR for sequence lengths $P = 2$ and 53. For comparison purposes, we also present the BER of the pilot-assisted noncoherent receiver [29] implemented with complexity $\mathcal{O}\left(P^4\right)$ and the pilot-assisted coherent receiver [30] implemented with linear complexity. We observe that -as expected- the conventional one-lag ($P = 2$) noncoherent receiver exhibits a $2 - 3$dB loss in comparison with the coherent MRC receiver. The SNR loss is reduced to $1 - 1.5$dB by ML sequence detection for $P = 53$, i.e. block detection of 106 bits, which is implemented by the proposed algorithm with polynomial complexity while the conventional sequence detection implementation would require an exhaustive search among 2^{104} vectors of length 106. The pilot-assisted noncoherent receiver [29] operates with rate $\frac{P-1}{P+1}$ since it consumes one additional block for the initial pilot transmission and one additional block for differential encoding and exhibits similar performance with the proposed ML receiver for $P = 2$. However, for $P = 53$ the proposed ML noncoherent receiver outperforms the pilot-assisted noncoherent one when they operate with the same complexity $\mathcal{O}\left(P^4\right)$. The proposed receiver is also superior to the pilot-assisted coherent receiver [30]. For the latter, we used one pilot OSTBC matrix and $P-1$ information OSTBC matrices to maintain the information rate $\frac{P-1}{P}$ of the differential encoding scheme associated with the proposed ML noncoherent receiver.

In Fig. 3, we set the SNR to 6dB and present BER and computational complexity curves of the proposed ML and pilot-assisted [29] noncoherent receivers versus sequence length P. We observe that, if the pilot-assisted receiver operates with the same complexity with the proposed ML receiver, then its performance deteriorates as the sequence length P increases. In Fig. 3(a), the BER of the coherent MRC receiver is also presented as a performance lower bound. Fig. 3(b) demonstrates the significant complexity gain offered by the proposed algorithm. We recall that exhaustive search fails for a sequence length $P > 15$ while the proposed algorithm maintains ML performance with polynomial computational complexity. For example, if $P = 53$ ($NP = 106$), then

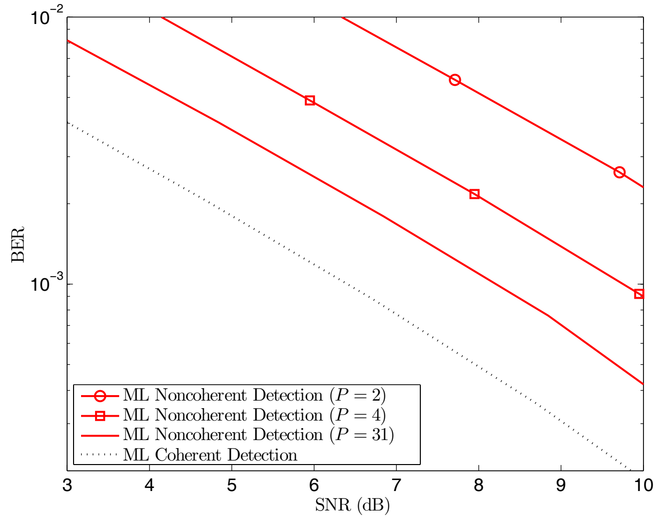

Fig. 4. BER versus SNR for ML coherent OSTBC receiver and ML noncoherent OSTBC receivers with sequence length $P=2$ (conventional receiver), $P=4$ (8 bits), and $P=31$ (62 bits) upon Ricean fading.

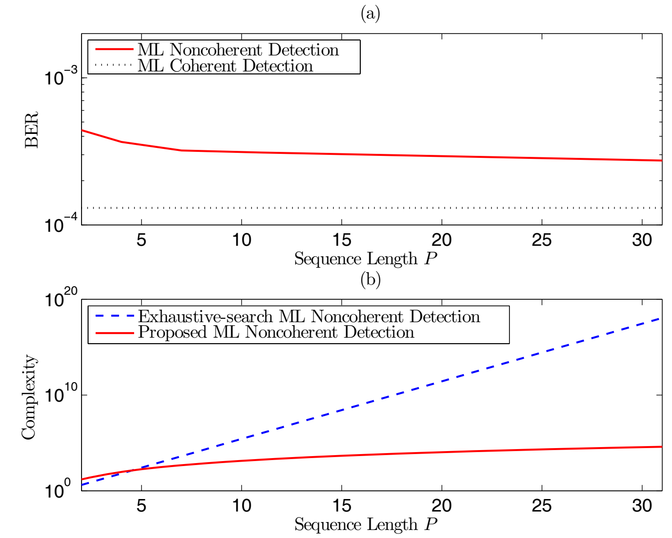

Fig. 5. (a) BER and (b) complexity versus sequence length P for SNR $= 10.7$dB and Ricean fading.

the conventional receiver implementation would require an exhaustive search among $2^{104} \approx 2 \cdot 10^{31}$ binary vectors of length 106 while the proposed implementation performs a search among $\binom{105}{0} + \binom{105}{1} + \binom{105}{2} + \binom{105}{3} \approx 2 \cdot 10^5$ binary vectors of length 106. Finally, to demonstrate the efficiency of our proposed algorithm for the Ricean fading channel case, relevant performance and complexity results are presented in Figs. 4 and 5 where the mean channel vector is selected to belong to the range of the channel covariance matrix.

VI. Conclusions

We proved that ML noncoherent sequence detection is always polynomially solvable with respect to the sequence length for OSTBC and correlated (in general) static Rayleigh channels and developed a novel algorithm that performs ML noncoherent OSTBC detection with polynomial complexity whose order is completely determined by the channel covariance matrix rank. Our proposed detector operates in polynomial time for static Ricean channels as well, if the mean channel vector is in the range of the channel covariance matrix. Hence, low-rank channel correlation is desired for Rayleigh channels while full-rank channel correlation is desired for Ricean channels to guarantee polynomial ML detection complexity. For the the cases of block-fading Rayleigh or Ricean channel matrices, similar conclusions were drawn.

Appendix

Proof of (8)

We note that $\mathbf{D}(\mathbf{s})$ satisfies the orthogonality property, since

$$\mathbf{D}(\mathbf{s})\mathbf{D}^H(\mathbf{s}) = \operatorname{diag}\left(\left[\mathbf{C}\left(\mathbf{s}^{(1)}\right)\mathbf{C}^H\left(\mathbf{s}^{(1)}\right), \ldots, \mathbf{C}\left(\mathbf{s}^{(P)}\right)\mathbf{C}^H\left(\mathbf{s}^{(P)}\right)\right]\right) = N\mathbf{I}_{M_t P}. \tag{33}$$

Then, the received matrix in (6) becomes $\mathbf{Y} = \mathbf{H}\mathbf{D}(\mathbf{s}) + \mathbf{V}$ where $\mathbf{V} \triangleq \left[\mathbf{V}^{(1)} \ldots \mathbf{V}^{(P)}\right] \in \mathbb{C}^{M_r \times TP}$. Due to [37]

$$\operatorname{vec}(\mathbf{ABC}) = \left(\mathbf{C}^T \otimes \mathbf{A}\right)\operatorname{vec}(\mathbf{B}), \tag{34}$$

we obtain

$$\mathbf{y} = \operatorname{vec}(\mathbf{H}\mathbf{D}(\mathbf{s}) + \mathbf{V}) = \left(\mathbf{D}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r}\right)\mathbf{h} + \mathbf{v} \tag{35}$$

where $\mathbf{v} = \operatorname{vec}(\mathbf{V}) \in \mathbb{C}^{M_r TP}$. Then, it can be proven that $\mathbf{y}$ given $\mathbf{s}$ is a complex Gaussian vector with mean $E\{\mathbf{y}|\mathbf{s}\} = \mathrm{E}\left\{\left(\mathbf{D}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r}\right)\mathbf{h} + \mathbf{v}\middle|\,\mathbf{s}\right\} = \left(\mathbf{D}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r}\right)E\{\mathbf{h}\} + E\{\mathbf{v}\} = \left(\mathbf{D}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r}\right)\boldsymbol{\mu}$ and covariance matrix

$$\begin{aligned}\mathbf{C}_y(\mathbf{s}) =& E\Big\{\left(\left(\mathbf{D}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r}\right)(\mathbf{h}-\boldsymbol{\mu}) + \mathbf{v}\right) \\ &\times \left(\left(\mathbf{D}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r}\right)(\mathbf{h}-\boldsymbol{\mu}) + \mathbf{v}\right)^H \Big|\, \mathbf{s}\Big\} \\ =& \left(\mathbf{D}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r}\right)\mathbf{Q}\,\mathbf{Q}^H\left(\mathbf{D}^*(\mathbf{s}) \otimes \mathbf{I}_{M_r}\right) + \sigma_v^2 \mathbf{I}_{M_r TP},\end{aligned} \tag{36}$$

hence (8) is obtained. □

Proof of (10)

Using identities for the determinant and inverse of a rank-deficient update [38], we compute

$$\begin{aligned}&|\mathbf{C}_y(\mathbf{s})| \\ &\overset{(36)}{=} \left|\sigma_v^2 \mathbf{I}_{M_r TP}\right| \left|\mathbf{I}_D + \frac{1}{\sigma_v^2}\mathbf{Q}^H\left(\mathbf{D}^*(\mathbf{s}) \otimes \mathbf{I}_{M_r}\right)\left(\mathbf{D}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r}\right)\mathbf{Q}\right| \\ &= \sigma_v^{2M_r TP}\left|\mathbf{I}_D + \frac{1}{\sigma_v^2}\mathbf{Q}^H\left(\mathbf{D}^*(\mathbf{s})\mathbf{D}^T(\mathbf{s}) \otimes \mathbf{I}_{M_r}\right)\mathbf{Q}\right| \\ &\overset{(33)}{=} \sigma_v^{2M_r TP}\left|\mathbf{I}_D + \frac{N}{\sigma_v^2}\mathbf{Q}^H\left(\mathbf{I}_{M_t P} \otimes \mathbf{I}_{M_r}\right)\mathbf{Q}\right| \\ &= \sigma_v^{2M_r TP}\left|\mathbf{I}_D + \frac{N}{\sigma_v^2}\mathbf{Q}^H\mathbf{Q}\right| = \sigma_v^{2M_r TP}\left|\mathbf{I}_D + \frac{N}{\sigma_v^2}\boldsymbol{\Sigma}\right|\end{aligned} \tag{37}$$

$$
\begin{aligned}
\hat{\mathbf{s}}_{\text{opt}} = \underset{\mathbf{s}\in\{\pm 1\}^{NP}}{\arg\max} &\Bigg\{ -\left(\mathbf{y}-\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\boldsymbol{\mu}\right)^H \left(\frac{1}{\sigma_v^2}\mathbf{I}_{M_{\mathrm{r}}TP}-\frac{1}{\sigma_v^4}\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\mathbf{U}\mathbf{U}^H\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\right)\left(\mathbf{y}-\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\boldsymbol{\mu}\right)\Bigg\} \\
= \underset{\mathbf{s}\in\{\pm 1\}^{NP}}{\arg\max} &\Bigg\{ -\frac{1}{\sigma_v^2}\mathbf{y}^H\mathbf{y}+\frac{1}{\sigma_v^2}\mathbf{y}^H\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\boldsymbol{\mu}+\frac{1}{\sigma_v^2}\boldsymbol{\mu}^H\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\mathbf{y}-\frac{1}{\sigma_v^2}\boldsymbol{\mu}^H\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\boldsymbol{\mu} \\
&+\frac{1}{\sigma_v^4}\mathbf{y}^H\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\mathbf{U}\mathbf{U}^H\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\mathbf{y} \\
&-\frac{1}{\sigma_v^4}\mathbf{y}^H\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\mathbf{U}\mathbf{U}^H\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\boldsymbol{\mu} \\
&-\frac{1}{\sigma_v^4}\boldsymbol{\mu}^H\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\mathbf{U}\mathbf{U}^H\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\mathbf{y} \\
&+\frac{1}{\sigma_v^4}\boldsymbol{\mu}^H\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\mathbf{U}\mathbf{U}^H\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\boldsymbol{\mu}\Bigg\}
\end{aligned} \tag{39}
$$

and

$$
\begin{aligned}
\mathbf{C}_y^{-1}(\mathbf{s}) &\overset{(36)}{=} \frac{1}{\sigma_v^2}\mathbf{I}_{M_{\mathrm{r}}TP}-\frac{1}{\sigma_v^2}\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\mathbf{Q} \\
&\quad\times\left(\mathbf{I}_D+\frac{1}{\sigma_v^2}\mathbf{Q}^H\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\mathbf{Q}\right)^{-1} \\
&\quad\times\mathbf{Q}^H\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\frac{1}{\sigma_v^2} \\
&\overset{(33)}{=}\frac{1}{\sigma_v^2}\mathbf{I}_{M_{\mathrm{r}}TP}-\frac{1}{\sigma_v^4}\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\mathbf{Q} \\
&\quad\times\left(\mathbf{I}_D+\frac{N}{\sigma_v^2}\mathbf{Q}^H\left(\mathbf{I}_{M_{\mathrm{t}}P}\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\mathbf{Q}\right)^{-1}\mathbf{Q}^H\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right) \\
&=\frac{1}{\sigma_v^2}\mathbf{I}_{M_{\mathrm{r}}TP}-\frac{1}{\sigma_v^4}\left(\mathbf{D}^T(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right) \\
&\quad\times\mathbf{Q}\left(\mathbf{I}_D+\frac{N}{\sigma_v^2}\boldsymbol{\Sigma}\right)^{-1}\mathbf{Q}^H\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right).
\end{aligned} \tag{38}
$$

We observe that $|\mathbf{C}_y(\mathbf{s})|$ is independent of the transmitted sequence $\mathbf{s}$, drop it from the maximization in (8), and substitute (9) in (38) and then back in (8) to obtain (39) which, due to (33), results in (10). □

Proof of (12)

We define the matrices $\mathbf{X} \overset{\triangle}{=} [\mathbf{X}_1 \ldots \mathbf{X}_N] \in \mathbb{C}^{M_{\mathrm{t}}\times TN}$ and $\mathbf{S} \overset{\triangle}{=} \left[\mathbf{s}^{(1)}\ldots\mathbf{s}^{(P)}\right] \in \{\pm 1\}^{N\times P}$ and observe that $\mathbf{s} = \mathrm{vec}(\mathbf{S})$,

$$
\mathbf{X}^H = [\mathbf{X}_1\ldots\mathbf{X}_N]^H = \left[\mathrm{vec}\left(\tilde{\mathbf{X}}_1^H\right)\ \ldots\ \mathrm{vec}\left(\tilde{\mathbf{X}}_{M_{\mathrm{t}}}^H\right)\right], \tag{40}
$$

and

$$
\begin{aligned}
\mathbf{G}(\mathbf{s}) &= \left[\sum_{n=1}^{N}\mathbf{X}_n s_n^{(1)}\ \ldots\ \sum_{n=1}^{N}\mathbf{X}_n s_n^{(P)}\right] \\
&= \left[[\mathbf{X}_1\ldots\mathbf{X}_N]\left(\mathbf{s}^{(1)}\otimes\mathbf{I}_T\right)\ldots[\mathbf{X}_1\ldots\mathbf{X}_N]\left(\mathbf{s}^{(P)}\otimes\mathbf{I}_T\right)\right] \\
&= \mathbf{X}\left[\left(\mathbf{s}^{(1)}\otimes\mathbf{I}_T\right)\ \ldots\ \left(\mathbf{s}^{(P)}\otimes\mathbf{I}_T\right)\right] \\
&= \mathbf{X}\left(\left[\mathbf{s}^{(1)}\ldots\mathbf{s}^{(P)}\right]\otimes\mathbf{I}_T\right) = \mathbf{X}\left(\mathbf{S}\otimes\mathbf{I}_T\right).
\end{aligned} \tag{41}
$$

We rewrite $\mathbf{D}(\mathbf{s})$ as

$$
\mathbf{D}(\mathbf{s}) = \mathrm{diag}\left(\left[\mathbf{C}\left(\mathbf{s}^{(1)}\right),\ldots,\mathbf{C}\left(\mathbf{s}^{(P)}\right)\right]\right) \tag{42}
$$

$$
\begin{aligned}
&= \begin{bmatrix}\mathbf{G}(\mathbf{s})\left(\mathbf{e}_1\mathbf{e}_1^T\otimes\mathbf{I}_T\right)\\ \vdots \\ \mathbf{G}(\mathbf{s})\left(\mathbf{e}_P\mathbf{e}_P^T\otimes\mathbf{I}_T\right)\end{bmatrix} = \left(\mathbf{I}_P\otimes\mathbf{G}(\mathbf{s})\right)\begin{bmatrix}\mathbf{e}_1\mathbf{e}_1^T\otimes\mathbf{I}_T\\ \vdots \\ \mathbf{e}_P\mathbf{e}_P^T\otimes\mathbf{I}_T\end{bmatrix} \\
&= \left(\mathbf{I}_P\otimes\mathbf{G}(\mathbf{s})\right)\left(\tilde{\mathbf{I}}_P\otimes\mathbf{I}_T\right) \\
&\overset{(41)}{=} \left(\mathbf{I}_P\otimes\mathbf{X}\left(\mathbf{S}\otimes\mathbf{I}_T\right)\right)\left(\tilde{\mathbf{I}}_P\otimes\mathbf{I}_T\right)
\end{aligned}
$$

where $\tilde{\mathbf{I}}_P \overset{\triangle}{=} \left[\mathbf{e}_1\mathbf{e}_1^T\ \ \ldots\ \ \mathbf{e}_P\mathbf{e}_P^T\right]^T \in \{0,1\}^{P^2\times P}$. Then,

$$
\begin{aligned}
&\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\mathbf{y} \\
&\overset{(42)}{=} \left(\left(\mathbf{I}_P\otimes\mathbf{X}^*\left(\mathbf{S}\otimes\mathbf{I}_T\right)\right)\left(\tilde{\mathbf{I}}_P\otimes\mathbf{I}_T\right)\otimes\mathbf{I}_{M_{\mathrm{r}}}\right)\mathrm{vec}(\mathbf{Y}) \\
&\overset{(34)}{=}\ \mathrm{vec}\left(\mathbf{Y}\left(\tilde{\mathbf{I}}_P^T\otimes\mathbf{I}_T\right)\left(\mathbf{I}_P\otimes\left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\mathbf{X}^H\right)\mathbf{I}_{M_{\mathrm{t}}P}\right) \\
&\overset{(34)}{=} \left(\mathbf{I}_{M_{\mathrm{t}}P}\otimes\mathbf{Y}\left(\tilde{\mathbf{I}}_P^T\otimes\mathbf{I}_T\right)\right)\mathrm{vec}\left(\mathbf{I}_P\otimes\left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\mathbf{X}^H\right).
\end{aligned} \tag{43}
$$

We observe that

$$
\begin{aligned}
&\mathrm{vec}\left(\mathbf{I}_P\otimes\left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\mathbf{X}^H\right) \\
&= \mathrm{vec}\left(\left[\mathbf{e}_1\ \ldots\ \mathbf{e}_P\right]\otimes\left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\mathbf{X}^H\right) \\
&= \mathrm{vec}\left(\left[\mathbf{e}_1\otimes\left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\mathbf{X}^H\ \ \ldots\ \mathbf{e}_P\otimes\left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\mathbf{X}^H\right]\right) \\
&= \begin{bmatrix}\mathrm{vec}\left(\mathbf{e}_1\otimes\left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\mathbf{X}^H\right)\\ \vdots \\ \mathrm{vec}\left(\mathbf{e}_P\otimes\left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\mathbf{X}^H\right)\end{bmatrix}
\end{aligned} \tag{44}
$$

where, for any $p = 1, 2, \ldots, P$,

$$
\begin{aligned}
&\mathrm{vec}\left(\mathbf{e}_p\otimes\left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\mathbf{X}^H\right) \\
&= \mathrm{vec}\left(\mathbf{e}_p\cdot 1\otimes\mathbf{I}_{PT}\left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\mathbf{X}^H\right) \\
&= \mathrm{vec}\left(\left(\mathbf{e}_p\otimes\mathbf{I}_{PT}\right)\left(1\otimes\left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\mathbf{X}^H\right)\right) \\
&= \mathrm{vec}\left(\left(\mathbf{e}_p\otimes\mathbf{I}_{PT}\right)\left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\mathbf{X}^H\mathbf{I}_{M_{\mathrm{t}}}\right) \\
&\overset{(34)}{=} \left(\mathbf{I}_{M_{\mathrm{t}}}\otimes\mathbf{e}_p\otimes\mathbf{I}_{PT}\right)\mathrm{vec}\left(\left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\mathbf{X}^H\right).
\end{aligned} \tag{45}
$$

Then,

$$
\begin{aligned}
&\mathrm{vec}\left(\left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\mathbf{X}^H\right) \\
&\overset{(40)}{=} \mathrm{vec}\left(\left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\left[\mathrm{vec}\left(\tilde{\mathbf{X}}_1^H\right)\ \ldots\ \mathrm{vec}\left(\tilde{\mathbf{X}}_{M_{\mathrm{t}}}^H\right)\right]\right) \\
&= \mathrm{vec}\left(\left[\left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\mathrm{vec}\left(\tilde{\mathbf{X}}_1^H\right)\ \ldots\ \left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\mathrm{vec}\left(\tilde{\mathbf{X}}_{M_{\mathrm{t}}}^H\right)\right]\right) \\
&\overset{(34)}{=} \mathrm{vec}\left(\left[\mathrm{vec}\left(\tilde{\mathbf{X}}_1^H\mathbf{S}\right)\ \ldots\ \mathrm{vec}\left(\tilde{\mathbf{X}}_{M_{\mathrm{t}}}^H\mathbf{S}\right)\right]\right) \\
&\overset{(34)}{=} \mathrm{vec}\left(\left[\left(\mathbf{I}_P\otimes\tilde{\mathbf{X}}_1^H\right)\mathrm{vec}(\mathbf{S})\ \ldots\ \left(\mathbf{I}_P\otimes\tilde{\mathbf{X}}_{M_{\mathrm{t}}}^H\right)\mathrm{vec}(\mathbf{S})\right]\right)
\end{aligned}
$$

$$\begin{aligned}
\hat{\tilde{\mathbf{s}}}_{\text{opt}} &= \underset{\substack{\tilde{\mathbf{s}}\in\{\pm1\}^{NP+1}\\ \tilde{s}_{NP+1}=1}}{\arg\max}\ \tilde{\mathbf{s}}^T \begin{bmatrix} \frac{1}{\sigma_v}\mathbf{Z}\mathbf{E}^T\left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}^H\right)\mathbf{U} \\ \sigma_v\boldsymbol{\mu}^H\mathbf{U}\left(\mathbf{U}^H\mathbf{U}\right)^{-1}\left(\mathbf{I}_D-\frac{N}{\sigma_v^2}\mathbf{U}^H\mathbf{U}\right)\end{bmatrix}\left[\frac{1}{\sigma_v}\mathbf{U}^H\left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}\right)\mathbf{E}\mathbf{Z}^H\ \ \sigma_v\left(\mathbf{I}_D-\frac{N}{\sigma_v^2}\mathbf{U}^H\mathbf{U}\right)\left(\mathbf{U}^H\mathbf{U}\right)^{-1}\mathbf{U}^H\boldsymbol{\mu}\right]\tilde{\mathbf{s}} \\
&= \underset{\substack{\tilde{\mathbf{s}}\in\{\pm1\}^{NP+1}\\ \tilde{s}_{NP+1}=1}}{\arg\max}\left\|\left[\frac{1}{\sigma_v}\mathbf{U}^H\left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}\right)\mathbf{E}\mathbf{Z}^H\ \ \sigma_v\left(\left(\mathbf{U}^H\mathbf{U}\right)^{-1}-\frac{N}{\sigma_v^2}\mathbf{I}_D\right)\mathbf{U}^H\boldsymbol{\mu}\right]\tilde{\mathbf{s}}\right\|^2 \\
&\overset{(9)}{\underset{(49)}{=}} \underset{\substack{\tilde{\mathbf{s}}\in\{\pm1\}^{NP+1}\\ \tilde{s}_{NP+1}=1}}{\arg\max}\left\|\underbrace{\left[\frac{1}{\sigma_v}\left(\mathbf{I}_D+\frac{N}{\sigma_v^2}\boldsymbol{\Sigma}\right)^{-\frac{1}{2}}\mathbf{Q}^H\left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}\right)\mathbf{E}\mathbf{Z}^H\quad\sigma_v\boldsymbol{\Sigma}^{-1}\left(\mathbf{I}_D+\frac{N}{\sigma_v^2}\boldsymbol{\Sigma}\right)^{-\frac{1}{2}}\mathbf{Q}^H\boldsymbol{\mu}\right]\tilde{\mathbf{s}}}_{\mathbf{A}^H}\right\|^2 \\
&= \underset{\substack{\tilde{\mathbf{s}}\in\{\pm1\}^{NP+1}\\ \tilde{s}_{NP+1}=1}}{\arg\max}\ \tilde{\mathbf{s}}^T\mathbf{A}\mathbf{A}^H\tilde{\mathbf{s}} = \underset{\substack{\tilde{\mathbf{s}}\in\{\pm1\}^{NP+1}\\ \tilde{s}_{NP+1}=1}}{\arg\max}\ \Re\left\{\tilde{\mathbf{s}}^T\mathbf{A}\mathbf{A}^H\tilde{\mathbf{s}}\right\} = \underset{\substack{\tilde{\mathbf{s}}\in\{\pm1\}^{NP+1}\\ \tilde{s}_{NP+1}=1}}{\arg\max}\ \tilde{\mathbf{s}}^T\left(\Re\{\mathbf{A}\}\Re\{\mathbf{A}\}^T+\Im\{\mathbf{A}\}\Im\{\mathbf{A}\}^T\right)\tilde{\mathbf{s}} \\
&= \underset{\substack{\tilde{\mathbf{s}}\in\{\pm1\}^{NP+1}\\ \tilde{s}_{NP+1}=1}}{\arg\max}\ \tilde{\mathbf{s}}^T\left[\Re\{\mathbf{A}\}\ \Im\{\mathbf{A}\}\right]\left[\Re\{\mathbf{A}\}\ \Im\{\mathbf{A}\}\right]^T\tilde{\mathbf{s}} = \underset{\substack{\tilde{\mathbf{s}}\in\{\pm1\}^{NP+1}\\ \tilde{s}_{NP+1}=1}}{\arg\max}\left\|\mathbf{V}^T\tilde{\mathbf{s}}\right\|. \qquad (50)
\end{aligned}$$

$$= \operatorname{vec}\left(\left[\mathbf{Z}_1^H\mathbf{s}\ \ldots\ \mathbf{Z}_{M_t}^H\mathbf{s}\right]\right) = \begin{bmatrix}\mathbf{Z}_1^H\mathbf{s}\\ \vdots \\ \mathbf{Z}_{M_t}^H\mathbf{s}\end{bmatrix} = \mathbf{Z}^H\mathbf{s}. \qquad (46)$$

Substituting (46) in (45) and then back in (44), we obtain

$$\begin{aligned}
&\operatorname{vec}\left(\mathbf{I}_P\otimes\left(\mathbf{S}^T\otimes\mathbf{I}_T\right)\mathbf{X}^H\right) \qquad (47)\\
&= \begin{bmatrix}(\mathbf{I}_{M_t}\otimes\mathbf{e}_1\otimes\mathbf{I}_{PT})\mathbf{Z}^H\mathbf{s}\\ \vdots \\ (\mathbf{I}_{M_t}\otimes\mathbf{e}_P\otimes\mathbf{I}_{PT})\mathbf{Z}^H\mathbf{s}\end{bmatrix} = \left(\begin{bmatrix}\mathbf{I}_{M_t}\otimes\mathbf{e}_1\otimes\mathbf{I}_P\\ \vdots \\ \mathbf{I}_{M_t}\otimes\mathbf{e}_P\otimes\mathbf{I}_P\end{bmatrix}\otimes\mathbf{I}_T\right)\mathbf{Z}^H\mathbf{s}.
\end{aligned}$$

Using (47), eq. (43) becomes

$$\begin{aligned}
&\left(\mathbf{D}^*(\mathbf{s})\otimes\mathbf{I}_{M_r}\right)\mathbf{y} = \left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}\right)\\
&\quad\times\left(\mathbf{I}_{M_tP}\otimes\tilde{\mathbf{I}}_P^T\otimes\mathbf{I}_T\right)\left(\begin{bmatrix}\mathbf{I}_{M_t}\otimes\mathbf{e}_1\otimes\mathbf{I}_P\\ \vdots \\ \mathbf{I}_{M_t}\otimes\mathbf{e}_P\otimes\mathbf{I}_P\end{bmatrix}\otimes\mathbf{I}_T\right)\mathbf{Z}^H\mathbf{s}\\
&= \left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}\right)\left(\begin{bmatrix}\left(\mathbf{I}_{M_t}\otimes\tilde{\mathbf{I}}_P^T\right)\left(\mathbf{I}_{M_t}\otimes\mathbf{e}_1\otimes\mathbf{I}_P\right)\\ \vdots \\ \left(\mathbf{I}_{M_t}\otimes\tilde{\mathbf{I}}_P^T\right)\left(\mathbf{I}_{M_t}\otimes\mathbf{e}_P\otimes\mathbf{I}_P\right)\end{bmatrix}\otimes\mathbf{I}_T\right)\mathbf{Z}^H\mathbf{s}\\
&= \left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}\right)\left(\begin{bmatrix}\mathbf{I}_{M_t}\otimes\tilde{\mathbf{I}}_P^T(\mathbf{e}_1\otimes\mathbf{I}_P)\\ \vdots \\ \mathbf{I}_{M_t}\otimes\tilde{\mathbf{I}}_P^T(\mathbf{e}_P\otimes\mathbf{I}_P)\end{bmatrix}\otimes\mathbf{I}_T\right)\mathbf{Z}^H\mathbf{s}\\
&= \left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}\right)\left(\begin{bmatrix}\mathbf{I}_{M_t}\otimes\mathbf{e}_1\mathbf{e}_1^T\\ \vdots \\ \mathbf{I}_{M_t}\otimes\mathbf{e}_P\mathbf{e}_P^T\end{bmatrix}\otimes\mathbf{I}_T\right)\mathbf{Z}^H\mathbf{s}\\
&= \left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}\right)\mathbf{E}\mathbf{Z}^H\mathbf{s} \qquad (48)
\end{aligned}$$

which is (12). □

Proof of (21)

From (9), we observe that $\mathbf{Q}$ and $\mathbf{U}$ have identical ranges and

$$\mathbf{U}^H\mathbf{U} = \left(\boldsymbol{\Sigma}^{-1}+\frac{N}{\sigma_v^2}\mathbf{I}_D\right)^{-1}. \qquad (49)$$

If $\boldsymbol{\mu}$ is in the range of $\mathbf{C}_h = \mathbf{Q}\,\mathbf{Q}^H$ (equivalently, the range of $\frac{1}{\sigma_v}\mathbf{U}$, i.e. $\boldsymbol{\mu} = \frac{1}{\sigma_v}\mathbf{U}\mathbf{a}$ for some $\mathbf{a}\in\mathbb{C}^D$), then we obtain $\mathbf{a} = \sigma_v\left(\mathbf{U}^H\mathbf{U}\right)^{-1}\mathbf{U}^H\boldsymbol{\mu}$, set $\mathbf{B}\triangleq\frac{1}{\sqrt{N}}\mathbf{Z}\mathbf{E}^T\left(\mathbf{I}_{M_tP}\otimes\mathbf{Y}^H\right)$, $\mathbf{C}\triangleq\frac{\sqrt{N}}{\sigma_v}\mathbf{U}$, and $\mathbf{x}\triangleq\sqrt{N}\boldsymbol{\mu}$ in Lemma 2, and rewrite (17) as (50). □

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. Sel. 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," *Proc. IEEE*, vol. 92, pp. 219–270, Feb. 2004.

[8] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. 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 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, Mar. 2000.

[14] 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, Sep. 2000.

[15] 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.

[16] W.-K. Ma, P. C. Ching, T. N. Davidson, and X.-G. Xia, "Blind maximum-likelihood decoding for orthogonal space-time block codes: a semidefinite relaxation approach," in *Proc. IEEE GLOBECOM*, Dec. 2003, vol. 4, pp. 2094–2098.

[17] 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.

[18] 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.

[19] 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.

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

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

[22] 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.

[23] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Sel. Areas Commun.*, vol. 18, pp. 1169–1174, July 2000.

[24] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. Inf. Theory*, vol. 46, pp. 2567–2578, Nov. 2000.

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

[26] R. Schober and L. H.-J. Lampe, "Noncoherent receivers for differential space-time modulation," *IEEE Trans. Commun.*, vol. 50, pp. 768–777, May 2002.

[27] A. Steiner, M. Peleg, and S. Shamai, "Iterative decoding of space-time differentially coded unitary matrix modulation," *IEEE Trans. Signal Process.*, vol. 50, pp. 2385–2395, Oct. 2002.

[28] X.-B. Liang and X.-G. Xia, "Fast differential unitary space-time demodulation via square orthogonal designs," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1331–1336, July 2005.

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

[30] 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, Sep. 2004.

[31] 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.

[32] 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.

[33] G. N. Karystinos and A. P. Liavas, "Efficient computation of the binary vector that maximizes a rank-3 quadratic form," in *Proc. 2006 Allerton Conf. Commun., Control, Computing*, Sep. 2006, pp. 1286–1291.

[34] G. N. Karystinos and D. A. Pados, "Rank-2-optimal adaptive design of binary spreading codes," *IEEE Trans. Inf. Theory*, vol. 53, pp. 3075–3080, Sep. 2007.

[35] G. N. Karystinos and A. P. Liavas, "Efficient computation of the binary vector that maximizes a rank-deficient quadratic form," *IEEE Trans. Inf. Theory*, vol. 56, July 2010.

[36] D. S. Papailiopoulos and G. N. Karystinos, "Polynomial-complexity maximum-likelihood block noncoherent MPSK detection," in *Proc. IEEE ICASSP*, Apr. 2008, pp. 2681–2684.

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

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

[39] 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.

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

[41] 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.

[42] C. Oestges, B. Clerckx, D. Vanhoenacker-Janvier, and A. J. Paulraj, "Impact of fading correlations on MIMO communication systems in geometry-based statistical channel models," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1112–1120, May 2005.

Dimitris S. Papailiopoulos (S'10) was born in Athens, Greece, on September 25, 1984. He received the Diploma (five-year program) and M.Sc. degrees in electronic and computer engineering from the Technical University of Crete, Chania, Greece, in 2007 and 2009, respectively. Since 2009, he has been a graduate student with the Communication Sciences Institute, Department of Electrical Engineering, University of Southern California, Los Angeles, where he is currently working toward the Ph.D. degree in electrical engineering. His research interests are in the areas of communications, signal processing, and networking with an emphasis on low-complexity signal detection algorithms, coding for distributed storage, and sparse principal component analysis.

Mr. Papailiopoulos received the University of Southern California Annenberg Graduate Fellowship in 2009. He is a student member of the IEEE Communications and Information Theory Societies.

George N. Karystinos (S'98-M'03) was born in Athens, Greece, on April 12, 1974. He received the Diploma degree in computer science and engineering (five-year program) from the University of Patras, Patras, Greece, in 1997, and the Ph.D. degree in electrical engineering from the State University of New York at Buffalo, Amherst, NY, in 2003. In August 2003, he joined the Department of Electrical Engineering, Wright State University, Dayton, OH as an assistant professor. Since September 2005, he has been an assistant professor with the Department of Electronic and Computer Engineering, Technical University of Crete, Chania, Greece. His current research interests are in the general areas of communication theory and adaptive signal processing with an emphasis on wireless and cooperative communications systems, low-complexity data detection, optimization with low complexity and limited data, spreading code and signal waveform design, and sparse principal component analysis.

Dr. Karystinos received a 2001 IEEE International Conference on Telecommunications best paper award and the 2003 IEEE TRANSACTIONS ON NEURAL NETWORKS Outstanding Paper Award. He is a member of the IEEE Communications, Signal Processing, Information Theory, and Computational Intelligence Societies and a member of Eta Kappa Nu.