Noncoherent Bit-Interleaved Coded OSTBC-OFDM with Maximum Spatial-Frequency Diversity

Ye Yang, Tsung-Hui Chang, *Member, IEEE,* Wing-Kin Ma, *Senior Member, IEEE,* Jianhua Ge, Chong-Yung Chi, *Senior Member, IEEE,* and P. C. Ching, *Fellow, IEEE*

***Abstract*—The combination of bit-interleaved coded modulation (BICM), orthogonal space-time block coding (OSTBC) and orthogonal frequency division multiplexing (OFDM) has been shown recently to be able to achieve maximum spatial-frequency diversity in frequency selective multi-path fading channels, provided that perfect channel state information (CSI) is available to the receiver. In view of the fact that perfect CSI can be obtained only if a sufficient amount of resource is allocated for training or pilot data, this paper investigates pilot-efficient noncoherent decoding methods for the BICM-OSTBC-OFDM system. In particular, we propose a noncoherent maximum-likelihood (ML) decoder that uses only one OSTBC-OFDM block. This block-wise decoder is suitable for relatively fast fading channels whose coherence time may be as short as one OSTBC-OFDM block. Our focus is mainly on noncoherent diversity analysis. We study a class of carefully designed transmission schemes, called perfect channel identifiability (PCI) achieving schemes, and show that they can exhibit good diversity performance. Specifically, we present a worst-case diversity analysis framework to show that PCI-achieving schemes can achieve the maximum noncoherent spatial-frequency diversity of BICM-OSTBC-OFDM. The developments are further extended to a distributed BICM-OSTBC-OFDM scenario in cooperative relay networks. Simulation results are presented to confirm our theoretical claims and show that the proposed noncoherent schemes can exhibit near-coherent performance.**

***Index Terms*—Bit-interleaved coded modulation (BICM), orthogonal space-time block coding (OSTBC), orthogonal frequency division multiplexing (OFDM), noncoherent decoding, maximum-likelihood (ML) decoding, diversity.**

Manuscript received October 30, 2011; revised April 19, 2012; accepted June 6, 2012. The associate editor coordinating the review of this paper and approving it for publication was N. C. Sagias.

This work of Y. Yang and J. Ge was supported by the National Basic Research Program of China (973 Program, Grant No. 2012CB316100), the "111" project (Grant No. B08038), the National Natural Science Foundation of China (Grant Nos. 61001207, 61101144 and 61101145), and the Fundamental Research Funds for the Central Universities (Grant No. K50510010017). The work of T.-H. Chang and C.-Y. Chi was supported by the National Science Council, Taiwan, under Grant NSC-99-2221-E-007-052-MY3. The work of W.-K. Ma and P. C. Ching was supported by a General Research Fund of the Hong Kong Research Grant Council (Project No. CUHK411709).

Y. Yang and J. Ge are with the State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an, Shaanxi 710071, China (e-mail: yeyang@mail.xidian.edu.cn, jhge@xidian.edu.cn).

T.-H. Chang is with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95616, USA (e-mail: tsunghui.chang@ieee.org). He was with National Tsing Hua University.

W.-K. Ma and P. C. Ching are with the Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong S.A.R., China (e-mail: wkma@ieee.org, pcching@ee.cuhk.edu.hk).

C.-Y. Chi is with the Institute of Communications Engineering and the Department of Electrical Engineering, National Tsing Hua University, Hsinchu, Taiwan 30013 (e-mail: cychi@ee.nthu.edu.tw).

Digital Object Identifier 10.1109/TWC.2012.071612.111936

I. Introduction

To approach the Shannon capacity, bit-interleaved coded modulation (BICM) [1] has been popularly used for channel-coded transmission, owing to its flexibility in the tradeoff between bit error performance and decoding complexity [2]. BICM has been concatenated with orthogonal space-time block codes (OSTBCs) to harvest temporal and spatial diversities in flat fading channels [3], [4], and has also been used in conjunction with orthogonal frequency division multiplexing (OFDM) to exploit frequency diversity [5], [6]. Recently, BICM is combined with OSTBC-OFDM and it was shown in [5] that the BICM-OSTBC-OFDM system is able to extract both the spatial and frequency diversities in independent and identically distributed (i.i.d.) Rayleigh fading channels. Further studies of BICM-OSTBC-OFDM can be found in [7] for generalized channel models. The designs mentioned above assume that the receiver has perfect channel state information (CSI). This can be achieved only if the receiver has been well trained through a sufficient amount of training or pilot data. While such (training) pilot-aided schemes generally work well under slow time-varying channels, and are popularly implemented in current wireless communication systems, they may no longer be efficient under fast time-varying channels. With shorter coherence times, which means more frequent training, both pilot overhead and CSI estimation error could become significant issues [8].

To improve the spectral efficiency and receiver performance, noncoherent techniques, including noncoherent data detection and blind/semiblind channel estimation methods, have been proposed. The noncoherent techniques are appealing because they use only a few number of pilots, or even no pilot, for data detection and channel estimation, and thus have great potential for deployment under fast time-varying channels. However, most of the existing works focus on two separate scenarios, namely, coded flat-fading space-time systems (no OFDM) [9]–[11] and uncoded OSTBC-OFDM systems (no channel coding) [12]–[16]. While it is possible to extend the works in [9], [10] to coded OSTBC-OFDM by considering each OFDM subcarrier as an individual flat-fading channel, such natural extension requires the channel to remain static over a large number of OFDM blocks. Similarly, many of the blind channel estimation methods in [12], [13] for uncoded OSTBC-OFDM also require the channel to remain unchanged for several OSTBC-OFDM blocks. These approaches are therefore suitable for slow or perhaps moderately fast fading channels. Another noncoherent approach is to employ differential space-

1536-1276/12$31.00 © 2012 IEEE

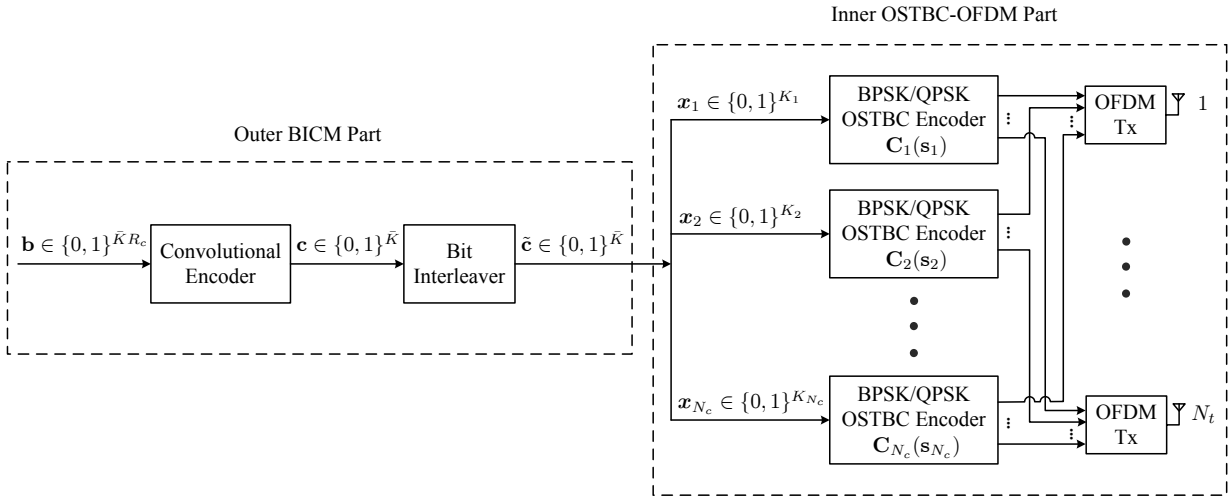

Fig. 1. Block diagram of the BICM-OSTBC-OFDM transmitter.

time coding schemes, e.g., [11]; however they suffer from a 3 dB signal-to-noise ratio (SNR) loss compared to the coherent receiver. In our previous works in [14]–[16], a noncoherent OSTBC-OFDM detection method based on the deterministic blind maximum-likelihood (ML) criterion was proposed. It was shown that this noncoherent detection method can exhibit near-coherent performance using only one OSTBC-OFDM block, thus appealing for fast time-varying channels. However, channel coding was not considered in [14]–[16].

In this paper, we consider the convolutional coded BICM-OSTBC-OFDM system, and aim to develop a noncoherent BICM-OSTBC-OFDM decoder that also uses one OSTBC-OFDM block. We assume that the time-domain multiple-input multiple-output (MIMO) multi-path channel coefficients are i.i.d. Rayleigh distributed. By exploiting the inter-subcarrier relationship of OFDM [14], we develop a block-wise noncoherent ML decoder and present a complexity-reduced implementation method. The primary focus of this paper is on the performance aspects, aiming to show the potential performance advantages of the proposed noncoherent approach. Firstly, like most of the noncoherent methods, the presented noncoherent ML decoder can be subject to the data ambiguity problem in the noise-free situation. We review some of the transmission schemes reported in [14]–[16], which can be directly used for the considered BICM-OSTBC-OFDM system for unique codeword decoding; e.g., the pilot-efficient *perfect channel identifiability (PCI)* achieving schemes [15], [16]. Secondly, we analyze the diversity order of the noncoherent ML decoder. To distinguish it from the diversity achieved by a coherent decoder, we refer to the diversity order achieved by the noncoherent decoder as *noncoherent diversity*. While the fundamental definitions of diversity are the same for both coherent and noncoherent systems, the characterization of the noncoherent diversity is much more difficult [17], [18]. In addition, the involvement of channel coding in this work further increases the challenge of the noncoherent diversity analysis. To overcome this issue, we present a worst-case diversity analysis framework for BICM-OSTBC-OFDM, and use it to show that PCI-achieving schemes can fully harvest both the maximum noncoherent spatial and frequency diversities. Finally, as a meaningful scenario variation, we extend the noncoherent ML decoder and diversity analysis framework to a distributed BICM-OSTBC-OFDM scenario arising from cooperative relay networks.

The remainder of this paper is organized as follows. In Section II, the system model is introduced. In Section III, the noncoherent ML decoder is derived and a complexity-reduced implementation method is presented. Section IV presents the unique data identification conditions and noncoherent diversity analysis. Extension to the distributed scenario is presented in Section V. Simulation results are shown in Section VI. Section VII draws the conclusions.

Notation: Throughout this paper, we use boldface lowercase letters and boldface uppercase letters to represent vectors and matrices, respectively. $\mathbf{I}_n$, $\mathbf{1}_n$ and $\mathbf{0}$ denote the $n \times n$ identity matrix, $n \times 1$ all-one vector and zero matrix or vector, respectively. $\mathbb{C}^{n \times m}$ denotes the set of all n by m complex matrices. Superscripts 'T' and 'H' respectively denote the transpose and conjugate transpose of a vector or a matrix. $\mathrm{Tr}(\cdot)$, $\mathrm{rank}(\cdot)$ and $\det(\cdot)$ stand for the trace, rank and determinant of a matrix, respectively. $\|\cdot\|_F$ and $\|\cdot\|_2$ respectively refer to the matrix Frobenius norm and 2-norm. $\mathbf{X} \succeq \mathbf{0}$ indicates that $\mathbf{X}$ is positive semidefinite. $\lambda_i(\mathbf{X})$ denotes the ith largest eigenvalue of $\mathbf{X}$; $\Re(\mathbf{X})$ represents the real part of $\mathbf{X}$. The operator $\otimes$ represents the Kronecker product. The cardinality of a set $\mathcal{N}$ is denoted by $|\mathcal{N}|$. Finally, $\mathbb{E}\{\cdot\}$, $\Pr(\cdot)$ and $p(\cdot)$ stand for the statistical expectation, probability (mass) function and probability density function, respectively.

II. System Model

We consider a point-to-point BICM-OSTBC-OFDM system with N_t transmit antennas and N_r receive antennas. Frequency selective multi-path channel fading between the transmitter and the receiver is assumed. As illustrated in Fig. 1, the transmitter is composed of two parts: (i) the outer convolutional coded BICM part which consists of a convolutional encoder and a bit interleaver [1]; (ii) the inner OSTBC-OFDM modulator [14], [19]. The convolutional encoder, which has a code

rate R_c, encodes the information bit vector $\boldsymbol{b} \in \{0,1\}^{\bar{K}R_c}$ into a length-$\bar{K}$ codeword $\mathbf{c} \in \boldsymbol{\mathcal{C}} \subset \{0,1\}^{\bar{K}}$, where $\boldsymbol{\mathcal{C}}$ is the convolutional code set. The codeword $\mathbf{c}$ is then processed by the interleaver, which outputs a scrambled codeword $\tilde{\mathbf{c}} \in \{0,1\}^{\bar{K}}$. The interleaver permutates the order of the codeword bits over the frequency domain, and, as will be shown later, can help the receiver to harvest frequency diversity. For OSTBC-OFDM modulation, the interleaved codeword $\tilde{\mathbf{c}}$ is segmented into N_c bit vectors, say, $\boldsymbol{x}_n \in \{0,1\}^{K_n}$, $n = 1, \ldots, N_c$, where N_c is the number of subcarriers of OFDM and K_n is the number of bits transmitted over subcarrier n, which satisfies $\sum_{n=1}^{N_c} K_n = \bar{K}$. Let $\mathbf{s}_n = 2\boldsymbol{x}_n - 1 \in \{\pm 1\}^{K_n}$ be the corresponding binary bits of $\boldsymbol{x}_n$ for all $n = 1, \ldots, N_c$. For ease of presentation, let us assume BPSK/QPSK OSTBC mappings[1]. For such cases, the transmitted OSTBC code over subcarrier n is given by $\mathbf{C}_n(\mathbf{s}_n) = \frac{1}{\sqrt{K_n}} \sum_{i=1}^{K_n} \mathbf{X}_{n,i} s_{n,i} \in \mathbb{C}^{T \times N_t}$, where T is the OSTBC code length, $s_{n,i} \in \{\pm 1\}$ is the ith entry of $\mathbf{s}_n$, and $\mathbf{X}_{n,i} \in \mathbb{C}^{T \times N_t}$ are the basis matrices of $\mathbf{C}_n(\cdot)$ which are designed such that $\mathbf{C}_n^H(\mathbf{s}_n)\mathbf{C}_n(\mathbf{s}_n) = \mathbf{I}_{N_t}$, $\forall\ \mathbf{s}_n \in \{\pm 1\}^{K_n}$ [20].

Assuming that the channels between the transmitter and the receiver remain static for T OFDM symbols, i.e., one OSTBC-OFDM block, the received signal at the receiver is given by [14]

$$\mathbf{Y}_n = \mathbf{C}_n(\mathbf{s}_n)\mathbf{H}_n + \mathbf{W}_n,\ n = 1, \ldots, N_c, \tag{1}$$

where $\mathbf{H}_n \in \mathbb{C}^{N_t \times N_r}$ and $\mathbf{W}_n \in \mathbb{C}^{T \times N_r}$ are the MIMO channel frequency response matrix and the additive white Gaussian noise (AWGN) matrix for the nth subcarrier, respectively. Each entry of $\mathbf{W}_n$ is assumed to have zero mean and variance σ_w^2. While the convolutional codeword $\mathbf{c}$ is not explicitly shown in (1), it can be related to the data bits $\mathbf{s}_n$, $n = 1, \ldots, N_c$, through a one-to-one binary mapping function $\mu(\cdot) : \{0,1\} \to \{\pm 1\}$. Specifically, suppose that the interleaver and OSTBC mapping function $\mathbf{C}_n(\cdot)$ map the kth bit of the codeword vector $\mathbf{c}$, denoted by c_k, to the data bit $s_{n,i}$ in the nth subcarrier. The relation between $s_{n,i}$ and c_k can be expressed as

$$s_{n,i} = \mu(c_k). \tag{2}$$

Assume that the receiver has perfect knowledge of the CSI $\mathbf{H}_1, \ldots, \mathbf{H}_{N_c}$. The coherent ML decoder for the received signal model in (1) is shown in [5] to be

$$\hat{\mathbf{c}} = \arg \min_{\mathbf{c} \in \boldsymbol{\mathcal{C}}} \sum_{k=1}^{\bar{K}} \left\{ \min_{\substack{\mathbf{s}_n \in \{\pm 1\}^{K_n} \\ s_{n,i} = \mu(c_k)}} \|\mathbf{Y}_n - \mathbf{C}_n(\mathbf{s}_n)\mathbf{H}_n\|_F^2 \right\}. \tag{3}$$

The above coherent ML decoding problem can be efficiently implemented by employing the Viterbi decoder (VD) that exploits the trellis structure of the convolutional code [21].

III. Noncoherent ML Decoding

Our focus in this section is on noncoherent decoding, i.e., decoding the information codeword $\mathbf{c}$ without knowing the CSI *a priori*. We are specifically interested in a block-wise noncoherent decoding approach that uses only $\mathbf{Y}_1, \ldots, \mathbf{Y}_{N_c}$ in one OSTBC-OFDM block. This approach is particularly suitable for the mobile scenarios where the channel coherence time may be as short as one OSTBC-OFDM block. To this end, let us first express (1) in a more compact form. Let

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \mathbf{h}_{1,1} & \cdots & \mathbf{h}_{1,N_r} \\ \vdots & \ddots & \vdots \\ \mathbf{h}_{N_t,1} & \cdots & \mathbf{h}_{N_t,N_r} \end{bmatrix} \in \mathbb{C}^{LN_t \times N_r} \tag{4}$$

denote the time-domain MIMO channel, where $\mathbf{h}_{m,i} \in \mathbb{C}^L$ is the channel impulse response vector from the mth transmit antenna to the ith receive antenna, and L is the channel length in the time domain. Moreover, let $\mathbf{f}_n = \frac{1}{\sqrt{N_c}}[1, e^{-j\frac{2\pi}{N_c}(n-1)}, \ldots, e^{-j\frac{2\pi}{N_c}(n-1)(L-1)}]^T$ be a discrete Fourier transform (DFT) vector for the nth subcarrier, where $j = \sqrt{-1}$. By the fact that each $\mathbf{H}_n$ can be parameterized by $\boldsymbol{\mathcal{H}}$ as $\mathbf{H}_n = (\mathbf{I}_{N_t} \otimes \mathbf{f}_n^T)\boldsymbol{\mathcal{H}}$, equation (1) can be expressed compactly as [14]

$$\boldsymbol{\mathcal{Y}} \triangleq [\mathbf{Y}_1^T, \ldots, \mathbf{Y}_{N_c}^T]^T = \boldsymbol{\mathcal{G}}(\boldsymbol{s})\boldsymbol{\mathcal{H}} + \boldsymbol{\mathcal{W}}, \tag{5}$$

where $\boldsymbol{\mathcal{W}} = [\mathbf{W}_1^T, \ldots, \mathbf{W}_{N_c}^T]^T \in \mathbb{C}^{N_c T \times N_r}$, $\boldsymbol{s} = [\mathbf{s}_1^T, \ldots, \mathbf{s}_{N_c}^T]^T \in \{\pm 1\}^{\bar{K}}$, and

$$\boldsymbol{\mathcal{G}}(\boldsymbol{s}) = \begin{bmatrix} \mathbf{C}_1(\mathbf{s}_1)(\mathbf{I}_{N_t} \otimes \mathbf{f}_1^T) \\ \vdots \\ \mathbf{C}_{N_c}(\mathbf{s}_{N_c})(\mathbf{I}_{N_t} \otimes \mathbf{f}_{N_c}^T) \end{bmatrix} \in \mathbb{C}^{N_c T \times LN_t} \tag{6}$$

is a supercode that satisfies $\boldsymbol{\mathcal{G}}^H(\boldsymbol{s})\boldsymbol{\mathcal{G}}(\boldsymbol{s}) = \mathbf{I}_{LN_t}$, $\forall\ \boldsymbol{s} \in \{\pm 1\}^{\bar{K}}$.

A. Block-Wise Noncoherent ML Decoding

Our noncoherent decoding design is based on the noncoherent ML criterion [1], [2]:

$$\hat{\mathbf{c}} = \arg \max_{\mathbf{c} \in \boldsymbol{\mathcal{C}}} \sum_{k=1}^{\bar{K}} \log p\left(\boldsymbol{\mathcal{Y}} | c_k\right), \tag{7}$$

where c_k's are assumed to be uniformly distributed and $\log p\left(\boldsymbol{\mathcal{Y}} | c_k\right)$ is the ML bit metric for c_k. By applying the law of total probability and taking into account the relation of $s_{n,i} = \mu(c_k)$, the noncoherent ML decoder in (7) can be further reformulated as [1]

$$\hat{\mathbf{c}} = \arg \max_{\mathbf{c} \in \boldsymbol{\mathcal{C}}} \sum_{k=1}^{\bar{K}} \log \sum_{\substack{\boldsymbol{s} \in \{\pm 1\}^{\bar{K}} \\ s_{n,i} = \mu(c_k)}} p\left(\boldsymbol{\mathcal{Y}} | \boldsymbol{s}\right). \tag{8}$$

We assume that the elements of channel $\boldsymbol{\mathcal{H}}$ are i.i.d. complex Gaussian distributed with zero mean and variance σ_h^2. Under this channel model, $p\left(\boldsymbol{\mathcal{Y}} | \boldsymbol{s}\right)$ can be shown to be

$$p\left(\boldsymbol{\mathcal{Y}} | \boldsymbol{s}\right) = \alpha \exp\left(\frac{\sigma_h^2}{\sigma_w^2(\sigma_h^2 + \sigma_w^2)} \mathrm{Tr}\left(\boldsymbol{\mathcal{Y}}^H \boldsymbol{\mathcal{G}}(\boldsymbol{s})\boldsymbol{\mathcal{G}}^H(\boldsymbol{s})\boldsymbol{\mathcal{Y}}\right)\right), \tag{9}$$

where

$$\alpha = (\pi\sigma_w^2)^{-N_c T N_r} \left(1 + \frac{\sigma_h^2}{\sigma_w^2}\right)^{-LN_t N_r} \exp\left(\frac{-1}{\sigma_w^2} \mathrm{Tr}\left(\boldsymbol{\mathcal{Y}}^H \boldsymbol{\mathcal{Y}}\right)\right).$$

[1]We should note that fundamentally, the techniques to be developed are also applicable to M-ary phase shift keying (PSK) constellations, or any constant modulus constellations.

By (9) and by applying the max-log approximation $\log\sum_\ell e^{x_\ell} \simeq \max_\ell x_\ell$, we can represent (8) (approximately) as

$$\hat{\mathbf{c}} = \arg \max_{\mathbf{c}\in\mathcal{C}} \sum_{k=1}^{\bar{K}} \max_{\substack{\boldsymbol{s}\in\{\pm1\}^{\bar{K}} \\ s_{n,i}=\mu(c_k)}} \mathrm{Tr}\left(\boldsymbol{\mathcal{Y}}^H \boldsymbol{\mathcal{G}}(\boldsymbol{s})\boldsymbol{\mathcal{G}}^H(\boldsymbol{s})\boldsymbol{\mathcal{Y}}\right). \quad (10)$$

The above noncoherent ML decoding problem can be handled by the Viterbi decoder as in coherent ML [5]. However, the noncoherent ML case in (10) is much more complex to process. Specifically, in the trellis search of the noncoherent Viterbi decoder, we have to solve the following two discrete optimization problems in order to decide c_k:

$$\max_{\substack{\boldsymbol{s}\in\{\pm1\}^{\bar{K}} \\ s_{n,i}=\mu(c_k)}} \mathrm{Tr}\left(\boldsymbol{\mathcal{Y}}^H \boldsymbol{\mathcal{G}}(\boldsymbol{s})\boldsymbol{\mathcal{G}}^H(\boldsymbol{s})\boldsymbol{\mathcal{Y}}\right) \quad (11)$$
$$\text{s.t.} \quad c_k = 1 \ (\text{or } c_k = 0).$$

The above two optimization problems are in fact the noncoherent ML detection problem studied in the context of uncoded OSTBC-OFDM [14]. While it has been shown in [14] that (11) can be recast as a Boolean quadratic program (BQP), which can be efficiently handled by a convex approximation method, namely, semidefinite relaxation (SDR) [22], [23], the total complexity required for solving the noncoherent ML decoding problem (10) is still much higher compared to the coherent ML decoder in (3). To overcome this implementation issue, instead we present in the next subsection a heuristic complexity-reduced implementation method with good performance.

B. Complexity-Reduced Implementation

It is interesting to note that the inner maximization problem in (10) can be expressed as [14]

$$\max_{\substack{\boldsymbol{s}\in\{\pm1\}^{\bar{K}} \\ s_{n,i}=\mu(c_k)}} \mathrm{Tr}\left(\boldsymbol{\mathcal{Y}}^H \boldsymbol{\mathcal{G}}(\boldsymbol{s})\boldsymbol{\mathcal{G}}^H(\boldsymbol{s})\boldsymbol{\mathcal{Y}}\right)$$
$$= \|\boldsymbol{\mathcal{Y}}\|_F^2 - \min_{\substack{\boldsymbol{s}\in\{\pm1\}^{\bar{K}} \\ s_{n,i}=\mu(c_k)}} \left\{ \min_{\boldsymbol{\mathcal{H}}\in\mathbb{C}^{LN_t\times N_r}} \|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{G}}(\boldsymbol{s})\boldsymbol{\mathcal{H}}\|_F^2 \right\}, \quad (12)$$

where the minimization problem is known as the deterministic blind ML detector in the literature [14]. By substituting (12) into (10), we can rewrite problem (10) as

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}\in\mathcal{C}} \sum_{k=1}^{\bar{K}} \min_{\substack{\boldsymbol{s}\in\{\pm1\}^{\bar{K}} \\ s_{n,i}=\mu(c_k)}} \left\{ \min_{\boldsymbol{\mathcal{H}}\in\mathbb{C}^{LN_t\times N_r}} \|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{G}}(\boldsymbol{s})\boldsymbol{\mathcal{H}}\|_F^2 \right\}. \quad (13)$$

In contrast to the coherent ML decoder in (3) where the channel $\boldsymbol{\mathcal{H}}$ is given, the inner minimization term of the noncoherent decoder (13) implicitly performs joint (modulated) data detection and channel estimation. This simple observation motivates us to handle the noncoherent ML decoding problem (10) in a two-step approach — first solve the deterministic blind ML problem

$$\{\hat{\boldsymbol{s}}, \hat{\boldsymbol{\mathcal{H}}}\} = \arg \min_{\boldsymbol{s}\in\{\pm1\}^{\bar{K}}} \left\{ \min_{\boldsymbol{\mathcal{H}}\in\mathbb{C}^{LN_t\times N_r}} \|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{G}}(\boldsymbol{s})\boldsymbol{\mathcal{H}}\|_F^2 \right\} \quad (14)$$

to obtain a channel estimate $\hat{\boldsymbol{\mathcal{H}}}$, followed by coherently decoding the codeword $\mathbf{c}$ by (3) using $\hat{\boldsymbol{\mathcal{H}}}$ as the true channel. In accordance with (12), the optimal $(\boldsymbol{s}, \boldsymbol{\mathcal{H}})$ of (14) can be obtained by

$$\hat{\boldsymbol{s}} = \arg \max_{\boldsymbol{s}\in\{\pm1\}^{\bar{K}}} \mathrm{Tr}\left(\boldsymbol{\mathcal{Y}}^H \boldsymbol{\mathcal{G}}(\boldsymbol{s})\boldsymbol{\mathcal{G}}^H(\boldsymbol{s})\boldsymbol{\mathcal{Y}}\right), \quad (15a)$$

$$\hat{\boldsymbol{\mathcal{H}}} = \arg \min_{\boldsymbol{\mathcal{H}}\in\mathbb{C}^{LN_t\times N_r}} \|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{G}}(\hat{\boldsymbol{s}})\boldsymbol{\mathcal{H}}\|_F^2 = \boldsymbol{\mathcal{G}}^H(\hat{\boldsymbol{s}})\boldsymbol{\mathcal{Y}}. \quad (15b)$$

As mentioned above, problem (15a) can be efficiently handled by the convex SDR approximation method. As a result, the two-step approach only involves solving a convex SDR problem and the coherent ML decoding problem (3). This alternative is therefore computationally much cheaper than directly solving problem (10). While, for a large scale problem where N_c is large, (15a) is still more complex than conventional pilot-aided channel estimation methods (e.g., least squares (LS) and minimum mean-squared error (MMSE) channel estimators [24]), we should emphasize that divide-and-conquer strategies such as subchannel grouping [14] and low-complexity SDR solvers (e.g., [23]) can be further employed for efficient implementation. As we will show in the simulation section, the presented complexity-reduced approach performs very well.

IV. Unique Data Identifiability and Diversity Analysis

Having considered the implementation of noncoherent ML decoding in the last section, we now turn our attention to the fundamental performance aspects. In the first subsection, we review some transmission schemes reported in [14]–[16] that can guarantee unique data identification, i.e., guarantee the codeword $\mathbf{c}$ to be uniquely decoded by the noncoherent ML decoder (10) in the noise-free situation. In the second subsection, we present our main results on the diversity order of the noncoherent ML decoder (10).

A. Unique Data and Channel Identification

As a common issue in noncoherent approaches, the noncoherent ML decoder (10) is subject to data ambiguity in the noise-free situation. In particular, one can easily verify that the two problems in (11) have the same optimal objective value since if $\boldsymbol{s}$ is an optimal solution of the problem with $c_k = 1$, then $-\boldsymbol{s}$ is optimal to the other with $c_k = 0$, and vice versa. This implies that the noncoherent ML decoder (10) is not able to uniquely determine whether c_k is equal to one or zero. To ensure that $\mathbf{c}$ can be uniquely identified in the noise-free situation, we need to guarantee that the unconstrained problem (without the bit mapping constraint $s_{n,i} = \mu(c_k)$)

$$\max_{\boldsymbol{s}\in\{\pm1\}^{\bar{K}}} \mathrm{Tr}\left(\boldsymbol{\mathcal{Y}}^H \boldsymbol{\mathcal{G}}(\boldsymbol{s})\boldsymbol{\mathcal{G}}^H(\boldsymbol{s})\boldsymbol{\mathcal{Y}}\right) \quad (16)$$

can uniquely identify the true $\boldsymbol{s}$ in the absence of noise. Since, by (12), problem (16) is equivalent to (14), it is sufficient to guarantee that the following ambiguity condition

$$\boldsymbol{\mathcal{G}}(\boldsymbol{s})\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{G}}(\boldsymbol{s}')\boldsymbol{\mathcal{H}}' \quad (17)$$

holds only when $\boldsymbol{s} = \boldsymbol{s}'$ and $\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{H}}'$, i.e., unique data and channel identification.

To this end, we need to insert some pilots in $\boldsymbol{s}$. Consider a general pilot placement as follows:

$$\boldsymbol{s} \triangleq \boldsymbol{\Pi}[\mathbf{s}_{\rm p}^T, \mathbf{s}_{\rm d}^T]^T \in \{\pm 1\}^{\bar{K}}, \tag{18}$$

where $\mathbf{s}_{\rm d} \in \{\pm 1\}^{K_{\rm d}}$ denotes the data bit vector, and $\mathbf{s}_{\rm p} \in \{\pm 1\}^{K_{\rm p}}$ denotes the pilot bit vector, in which $K_{\rm d}$ and $K_{\rm p}$ ($\bar{K} = K_{\rm d} + K_{\rm p}$) are the numbers of data and pilot bits[2], respectively. The matrix $\boldsymbol{\Pi}$ in (18) is a $\bar{K}$ by $\bar{K}$ permutation matrix that describes how the pilots and data are assigned. We need to carefully design $\mathbf{s}_{\rm p}$ and $\boldsymbol{\Pi}$ such that unique data and channel identifiability can be achieved. This design problem has been studied by the authors in [14]–[16]. Here we summarize some of the key results:

1) One-pilot-code scheme: In this scheme, only one subcarrier is dedicated to transmitting pilot codes, e.g., $\mathbf{s}_{\rm p} = \mathbf{s}_1$. While the number of pilot codes used is far less than the channel length L, it is shown in [14] that this scheme can still ensure unique identification of $\boldsymbol{s}$ and $\boldsymbol{\mathcal{H}}$ *with probability one* in i.i.d. Rayleigh fading channels. Although this scheme can exhibit promising bit error performance in the uncoded system, we will show later via both analysis and simulations that this one-pilot-code scheme may not be able to fully harvest the coding and diversity gains provided by the outer channel code.

2) L-pilot-code scheme: It is intuitive that inserting more pilots will improve the data identifiability. The L-pilot-code scheme allows L out of N_c subcarriers for pilot transmission, e.g., $\mathbf{s}_{\rm p} = [\mathbf{s}_1^T, \ldots, \mathbf{s}_L^T]^T$. This scheme is stronger than the one-pilot-code scheme in the sense that, with this scheme, any (nonzero) channel can be uniquely identified, regardless of its statistical distribution; that is, the ambiguity condition (17) holds only if $\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{H}}'$ ($\neq \mathbf{0}$). This property is called *perfect channel identifiability (PCI)* [15], [16]. We will show in the next subsection that PCI-achieving schemes can fully harvest the spatial and frequency diversity gains provided by the channel coded OSTBC-OFDM system.

3) xL-pilot-bit scheme: The xL-pilot-bit scheme proposed in [15] is a scheme that can also achieve PCI, but consumes a smaller number of pilot bits than the L-pilot-code scheme. One of the key ingredients that makes the xL-pilot-bit scheme PCI achieving is the so called non-intersecting subspace (NIS) OSTBCs. Readers may refer to [25] for the detailed properties and construction method of NIS-OSTBCs[3]. Here we only emphasize how the xL-pilot-bit scheme can be constructed. Let $\mathbf{C}_{\rm NIS}(\cdot)$ denote the NIS-OSTBC and let $\mathbf{C}_{\rm O}(\cdot)$ denote an arbitrary OSTBC with the same dimension as $\mathbf{C}_{\rm NIS}(\cdot)$. The xL-pilot-bit scheme can be built as follows [15]:

xL-pilot-bit scheme: Let $\mathcal{N} \subset \{1, \ldots, N_c\}$, $|\mathcal{N}| = L$, be the subcarrier subset for which NIS-OSTBCs are allocated. Set

$$\mathbf{C}_n(\mathbf{s}_n) = \mathbf{C}_{\rm NIS}(\mathbf{s}_n)\ \forall n \in \mathcal{N}, \tag{19}$$

$$\mathbf{C}_n(\mathbf{s}_n) = \mathbf{C}_{\rm O}(\mathbf{s}_n)\ \forall n \in \{1, \ldots, N_c\} \setminus \mathcal{N}. \tag{20}$$

Moreover, for each $n \in \mathcal{N}$, assign x pilot bits to $s_{n,1}, \ldots, s_{n,x}$, where $1 \le x \le K_n$.

[2]When $K_{\rm p}$ pilot bits are transmitted, the total number of data bits reduced to $K_{\rm d}$, and thus the information bits transmitted in one OSTBC-OFDM block reduces to $K_{\rm d}R_c$.

[3]For NIS-OSTBC with M-ary PSK modulation, readers may refer to [26].

As will be shown in the simulation section, $x = 3$ is sufficient to achieve good decoding performance in general.

B. Noncoherent Diversity Analysis

Diversity order is an important performance measure for space-time-frequency coded systems. For the coherent BICM-OSTBC-OFDM system, the coherent diversity order has been analyzed in [5]. However, the analysis techniques used there is not applicable to the noncoherent diversity analysis. In fact, transmission schemes that achieve the full coherent diversity order do not necessarily achieve the full noncoherent diversity order. For the considered noncoherent BICM-OSTBC-OFDM, as will be seen below, the associated diversity order is difficult to characterize due to the involvement of outer channel coding. We will derive a *worst-case noncoherent diversity order* that can capture the worst diversity characteristic of the noncoherent ML decoder (10).

We simply assume that the receiver has only one antenna, i.e., $N_r = 1$. For $N_r > 1$, the diversity order is N_r times larger. In this case, the signal model in (5) reduces to

$$\boldsymbol{y} = \boldsymbol{\mathcal{G}}(\boldsymbol{s})\boldsymbol{h} + \boldsymbol{w}, \tag{21}$$

where $\boldsymbol{h} \in \mathbb{C}^{LN_t}$ and $\boldsymbol{w} \in \mathbb{C}^{N_c T}$ are respectively the channel vector and noise vector for $N_r = 1$. The noncoherent ML decoder in (10) (taking into account the pilot placement in (18)) reduces to

$$\hat{\mathbf{c}} = \arg \max_{\mathbf{c} \in \boldsymbol{\mathcal{C}}} \sum_{k=1}^{K_{\rm d}} \max_{\substack{\mathbf{s}_{\rm d} \in \{\pm 1\}^{K_{\rm d}} \\ s_{n,i} = \mu(c_k)}} \boldsymbol{y}^H \boldsymbol{\mathcal{G}}(\boldsymbol{s}) \boldsymbol{\mathcal{G}}^H(\boldsymbol{s}) \boldsymbol{y}. \tag{22}$$

We first analyze the pair-wise error probability (PEP) of the noncoherent ML decoder (22), that is, the probability that the noncoherent ML decoder mistakes the true transmitted codeword $\bar{\mathbf{c}} \in \boldsymbol{\mathcal{C}}$ for a distinct codeword $\hat{\mathbf{c}} \in \boldsymbol{\mathcal{C}}$. Let $d_{\rm free}$ be the free Hamming distance (i.e., the minimum Hamming distance) of the convolutional code $\boldsymbol{\mathcal{C}}$. Let $k_1, \ldots, k_d$ be the bit indices for which $\bar{c}_{k_\ell} \neq \hat{c}_{k_\ell}$ for all $\ell = 1, \ldots, d$, where $d \ge d_{\rm free}$. Moreover, let $\bar{s}_{n_\ell, i_\ell}$, the i_ℓth entry of $\bar{\mathbf{s}}_{n_\ell}$ in subcarrier $n_\ell \in \{1, \ldots, N_c\}$, be the binary bit mapped from the codeword bit $\bar{c}_{k_\ell}$, i.e., $\bar{s}_{n_\ell, i_\ell} = \mu(\bar{c}_{k_\ell})$, for all $\ell = 1, \ldots, d$. We prove in Appendix A the following proposition:

Proposition 1 *The PEP of the noncoherent ML decoder* (22) *is upper bounded as*

$$\Pr(\bar{\mathbf{c}} \to \hat{\mathbf{c}} | \bar{\mathbf{c}}) \le \beta_1 \sum_{\hat{\boldsymbol{s}} \in \boldsymbol{S}(\hat{\mathbf{c}})} \det{}^{-1}\left(\mathbf{I}_{LN_t} + \frac{\sigma_h^2 \beta_2}{\sigma_w^2} \boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}}) \right), \tag{23}$$

where $\beta_1 > \frac{1}{2}$, $0 < \beta_2 \le \frac{1}{16}$ *are constants,* $\bar{\boldsymbol{s}} = \boldsymbol{\Pi}[\mathbf{s}_{\rm p}^T, \bar{\mathbf{s}}_{\rm d}^T]^T$, $\hat{\boldsymbol{s}} = [(\hat{\boldsymbol{s}}^{(1)})^T, \ldots, (\hat{\boldsymbol{s}}^{(d)})^T]^T$, $\hat{\boldsymbol{s}}^{(\ell)} = \boldsymbol{\Pi}[\mathbf{s}_{\rm p}^T, (\hat{\mathbf{s}}_{\rm d}^{(\ell)})^T]^T$,

$$\boldsymbol{S}(\hat{\mathbf{c}}) = \{\hat{\boldsymbol{s}} | \hat{\mathbf{s}}_{\rm d}^{(\ell)} \in \{\pm 1\}^{K_{\rm d}}, \hat{s}_{n_\ell, i_\ell}^{(\ell)} = \mu(\hat{c}_{k_\ell}), \ell = 1, \ldots, d\}, \tag{24}$$

$$\boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}}) \triangleq \mathbf{I}_{LN_t} - \frac{1}{d} \sum_{\ell=1}^{d} \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\mathcal{G}}(\hat{\boldsymbol{s}}^{(\ell)}) \boldsymbol{\mathcal{G}}^H(\hat{\boldsymbol{s}}^{(\ell)}) \boldsymbol{\mathcal{G}}(\bar{\boldsymbol{s}}). \tag{25}$$

The PEP upper bound in (23) is considerably different from those for coherent BICM-OSTBC-OFDM systems [5]

and uncoded space-time systems [20]. The only resemblance one may roughly see is that each constituent term $\det^{-1}\left(\mathbf{I}_{LN_t}+\frac{\sigma_h^2\beta_2}{\sigma_w^2}\mathbf{\Omega}(\bar{s},\hat{s})\right)$ of (23) is somehow similar to the PEP upper bound for uncoded space-time systems [20]; the precise expression is nevertheless different upon close inspection, and we have to deal with a sum of such terms in (23).

Proposition 1 gives an important insight into analysis on the noncoherent BICM-OSTBC-OFDM diversity. As in the coherent scenario, the noncoherent diversity is defined as the high-SNR slope of the PEP in a log-log scale over all possible pairs of $\bar{\mathbf{c}}$ and $\hat{\mathbf{c}}$ [27]:

$$D_{\mathrm{NC}}=\min_{\bar{\mathbf{c}}\neq\hat{\mathbf{c}}}\left\{\lim_{\sigma_h^2/\sigma_w^2\to\infty}\frac{-\log\Pr(\bar{\mathbf{c}}\to\hat{\mathbf{c}}|\bar{\mathbf{c}})}{\log(\sigma_h^2/\sigma_w^2)}\right\}. \tag{26}$$

Substituting (23) into (26), we show in Appendix B that the noncoherent diversity order of the BICM-OSTBC-OFDM system is lower bounded as

$$D_{\mathrm{NC}}\geq D_{\mathrm{NC}}^{\star}\triangleq\min_{\bar{\mathbf{c}}\neq\hat{\mathbf{c}}}\left\{\min_{\hat{\boldsymbol{s}}\in\boldsymbol{S}(\hat{\mathbf{c}})}\mathrm{rank}(\mathbf{\Omega}(\bar{\boldsymbol{s}},\hat{\boldsymbol{s}}))\right\}. \tag{27}$$

The intuition behind (27) is that the PEP upper bound in (23) is asymptotically dominated by the constituent term $\det^{-1}\left(\mathbf{I}_{LN_t}+\frac{\sigma_h^2\beta_2}{\sigma_w^2}\mathbf{\Omega}(\bar{\boldsymbol{s}},\hat{\boldsymbol{s}})\right)$ that is least diminishing with the SNR, and $D_{\mathrm{NC}}^{\star}$ in (27) captures the worst diversity order of that term among all possible pairs of $\bar{\mathbf{c}}$ and $\hat{\mathbf{c}}$. We will henceforth call $D_{\mathrm{NC}}^{\star}$ the *worst-case noncoherent diversity order*.

Next we analyze the worst-case noncoherent diversity orders of the three transmission schemes presented in the previous subsection. We will need the following assumption on the bit interleaver:

A1) *For any codeword pair $\bar{\mathbf{c}}$ and $\hat{\mathbf{c}}$, there exists a subindex set $\{k_1',\ldots,k_{d_{\mathrm{free}}}'\}\subseteq\{k_1,\ldots,k_d\}$ such that $\{\bar{c}_{k_1'},\ldots,\bar{c}_{k_{d_{\mathrm{free}}}'}\}$ are mapped by the bit interleaver onto different subcarriers, say, $\{n_1',\ldots,n_{d_{\mathrm{free}}}'\}\subseteq\{n_1,\ldots,n_d\}$ where $n_i'\neq n_k'$ for all $i\neq k$.*

Assumption **A1)** basically says that the bit interleaver has to be 'random' enough. Under **A1)**, we prove in Appendix C the following theorem on the noncoherent diversity:

Theorem 1 *Assume that* **A1)** *holds. If the PCI-achieving schemes, e.g., the L-pilot-code scheme and the xL-pilot-bit scheme, are employed, then $D_{\mathrm{NC}}^{\star}=N_t\min(d_{\mathrm{free}},L)$.*

Since $N_t\min(d_{\mathrm{free}},L)$ is also the maximum diversity that can be achieved by the coherent ML decoder in (3) [5], Theorem 1 indicates that the L-pilot-code scheme and the xL-pilot-bit scheme can achieve the maximum spatial-frequency diversity offered by the system in a noncoherent manner. Interestingly, the one-pilot-code scheme, which is not PCI-achieving, may not benefit from the use of BICM to harvest the frequency diversity:

Corollary 1 *When the one-pilot-code scheme is employed, the worst-case diversity is no larger than N_t, i.e., $D_{\mathrm{NC}}^{\star}\leq N_t$.*

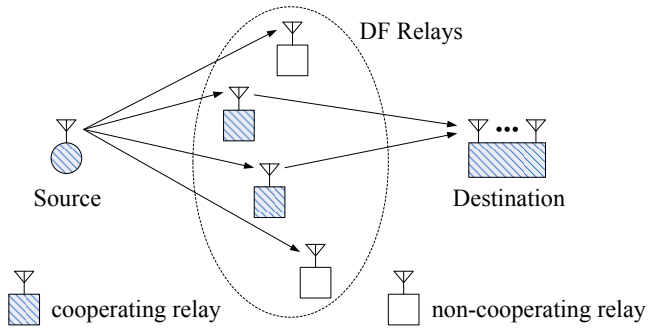

Fig. 2. Block diagram of the relay-based distributed BICM-OSTBC-OFDM system.

The proof is presented in Appendix D. Theorem 1 and Corollary 1 will be further corroborated by simulations in Section VI.

V. Extension to Distributed BICM-OSTBC-OFDM

In this section, we extend the noncoherent ML decoder in (10) to a relay-based distributed BICM-OSTBC-OFDM system. In this system, as illustrated in Fig. 2, a set of single-antenna relays collaborate to transmit the information bits sent from the source to a destination receiver. The direct link between the source and the destination is not considered. We assume that the relays employ the decode-and-forward (DF) strategy. Moreover, we assume that there is no central control and the relays may choose to join the cooperation or not, depending on whether they can successfully decode the information bits from the source. Suppose that there are N_s cooperating relays, and that the relays employ the distributed OSTBC scheme in [28]. Specifically, given the information bits $\boldsymbol{b}\in\{0,1\}^{\bar{K}R_c}$, the mth relay transmits the block sequence $\mathbf{C}_n(\mathbf{s}_n)\mathbf{q}_m\in\mathbb{C}^T$ over subcarrier n, for $n=1,\ldots,N_c$, where $\mathbf{C}_n(\mathbf{s}_n)$ is the OSTBC mapping function defined in Section II, and $\mathbf{q}_m\in\mathbb{C}^{N_t}$ is a preassigned signature sequence for the mth relay. The signature sequences $\mathbf{q}_1,\ldots,\mathbf{q}_{N_s}$ may be designed to enhance the receiver performance [28]; here we assume that $\mathbf{q}_m$ are given and fixed. Let $\boldsymbol{\mathcal{H}}\in\mathbb{C}^{LN_s\times N_r}$ be the channels from all the cooperating relays to the destination. Then the received signal for subcarrier n is given by

$$\begin{aligned}\mathbf{Y}_n&=\mathbf{C}_n(\mathbf{s}_n)\mathbf{Q}(\mathbf{I}_{N_s}\otimes\mathbf{f}_n^T)\boldsymbol{\mathcal{H}}+\mathbf{W}_n\\&=\mathbf{C}_n(\mathbf{s}_n)(\mathbf{I}_{N_t}\otimes\mathbf{f}_n^T)\boldsymbol{\mathcal{H}}_{\mathrm{v}}+\mathbf{W}_n,\end{aligned} \tag{28}$$

where $\mathbf{Q}=[\mathbf{q}_1,\ldots,\mathbf{q}_{N_s}]\in\mathbb{C}^{N_t\times N_s}$, $\boldsymbol{\mathcal{H}}_{\mathrm{v}}\triangleq(\mathbf{Q}\otimes\mathbf{I}_L)\boldsymbol{\mathcal{H}}\in\mathbb{C}^{LN_t\times N_r}$ is the virtual channel, and the second equality is obtained by using the Kronecker product property $(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D})=(\mathbf{AC}\otimes\mathbf{BD})$ [29]. The received OSTBC-OFDM block signal is thus given by

$$\boldsymbol{\mathcal{Y}}=\boldsymbol{\mathcal{G}}(\boldsymbol{s})\boldsymbol{\mathcal{H}}_{\mathrm{v}}+\boldsymbol{\mathcal{W}}, \tag{29}$$

where $\boldsymbol{\mathcal{Y}}$, $\boldsymbol{\mathcal{G}}(\boldsymbol{s})$ and $\boldsymbol{\mathcal{W}}$ are as defined in (5) and (6). For the signal model in (29), the corresponding noncoherent ML decoder can be shown to be

$$\hat{\mathbf{c}}=\arg\max_{\mathbf{c}\in\boldsymbol{\mathcal{C}}}\sum_{k=1}^{\bar{K}}\max_{\substack{\mathbf{s}_{\mathrm{d}}\in\{\pm1\}^{K_{\mathrm{d}}}\\ s_{n,i}=\mu(c_k)}}\mathrm{Tr}\left(\boldsymbol{\mathcal{Y}}^H\boldsymbol{\mathcal{G}}(\boldsymbol{s})\mathbf{R}^{-1}\boldsymbol{\mathcal{G}}^H(\boldsymbol{s})\boldsymbol{\mathcal{Y}}\right). \tag{30}$$

where $\mathbf{R} = \frac{1}{\sigma_h^2}(\mathbf{Q}\mathbf{Q}^H \otimes \mathbf{I}_L)^{-1} + \frac{1}{\sigma_w^2}\mathbf{I}_{LN_t}$. As one can observe from (30), the receiver requires to know exactly which relays join the cooperation since the signature matrix $\mathbf{Q}$ is required to be known. Owing to the analogy between (29) and (5), the noncoherent decoder in (10) can be applied to (29). The noncoherent decoder in (10) is able to decode $\mathbf{c}$ without the need of knowing $\mathbf{Q}$, which is hence particularly attractive for the relay-based decentralized scenario because the number of cooperative relays may vary over time and this can cause unknown changes of the signature matrix $\mathbf{Q}$. Interestingly, (10) can still achieve the maximum cooperative (spatial)-frequency diversity:

Theorem 2 *Assume that* **A1)** *holds, and that the matrix* $\mathbf{Q}$ *is of full rank. Suppose that* $N_r = 1$. *Then the worst-case diversity order achieved by the PCI-achieving schemes is given by*

$$D^{\star}_{\mathrm{DNC}} \triangleq \min_{\bar{\mathbf{c}} \neq \hat{\mathbf{c}}} \left\{ \min_{\hat{\boldsymbol{s}} \in \boldsymbol{S}(\hat{\mathbf{c}})} \mathrm{rank}((\mathbf{Q} \otimes \mathbf{I}_L)^H \boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}})(\mathbf{Q} \otimes \mathbf{I}_L)) \right\} = \min(N_s, N_t) \min(d_{\mathrm{free}}, L).$$

Proof: The worst-case diversity order $D^{\star}_{\mathrm{DNC}}$ can be proved following similar ideas as in Proposition 1 and in (26) and (27). To quickly see how the second equality can hold, let us consider the case of $d_{\mathrm{free}} \geq L$. Since, by Theorem 1, $\boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}})$ has the full rank for $d_{\mathrm{free}} \geq L$, $\mathrm{rank}((\mathbf{Q} \otimes \mathbf{I}_L)^H \boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}})(\mathbf{Q} \otimes \mathbf{I}_L)) = \mathrm{rank}(\mathbf{Q} \otimes \mathbf{I}_L) = \min(N_s, N_t)L$. The case of $d_{\mathrm{free}} < L$ can be proved following similar steps in Theorem 1; the details are omitted here. ■

Since $\min(N_s, N_t)$ is the maximum cooperative diversity order achieved by the distributed OSTBC [28] and $\min(d_{\mathrm{free}}, L)$ is the maximum frequency diversity order [5], Theorem 2 implies that the noncoherent BICM-OSTBC-OFDM decoder (10), using PCI-achieving schemes, achieves the maximum cooperative-frequency diversity order offered by the relay system.

VI. Simulation Results

In this section, we present some simulation results to examine the bit error performance of the noncoherent ML decoder (10) by considering the three transmission schemes in Section IV-A. We consider a BICM-OSTBC-OFDM system with 256 subcarriers ($N_c = 256$), and employ 1/2-rate convolutional codes ($R_c = 1/2$), QPSK modulation, and the Alamouti-concatenated code in [15, Eqn. (15)] ($N_t = 2$, $T = 4$ and 8 bits per code). If the xL-pilot-bit scheme in Section IV-A is employed, the NIS-OSTBC in [15, Eqn. (12)] is used as $\mathbf{C}_{\mathrm{NIS}}(\cdot)$ while the code in [15, Eqn. (15)] is adopted as $\mathbf{C}_{\mathrm{O}}(\cdot)$. The bit interleaver specified by the IEEE 802.11a standard [30] is employed [5]. In the simulations, we adopt packet transmissions where each packet contains 8 OSTBC-OFDM blocks. A total $8K_{\mathrm{d}}R_c$ information bits are encoded by the convolutional encoder and transmitted across 8 OSTBC-OFDM blocks, where the number K_{d} depends on which transmission scheme is employed. Unless specifically stated otherwise, the elements of $\boldsymbol{\mathcal{H}}$ are i.i.d. complex Gaussian distributed with zero mean and unit variance, and they remain invariant within one OSTBC-OFDM block but change independently from block to block. Note that the interleaver used in the IEEE 802.11a standard [30] is designed such that the code bits in one OSTBC-OFDM block are interleaved within the same block. Hence the noncoherent ML decoder (10) can still be implemented in a block-by-block manner. The complexity-reduced method in Section III-B is used for pragmatic implementation. We should emphasize here that, although the convolutional encoder described above encodes the information bits across multiple OSTBC-OFDM blocks, the diversity results presented in Section IV-B still hold true as long as assumption **A1**) is true. The E_b/N_0 per receive antenna is defined as

$$\frac{E_b}{N_0} = \frac{\mathbb{E}\left\{\|\boldsymbol{\mathcal{G}}(\boldsymbol{s})\boldsymbol{\mathcal{H}}\|_F^2\right\}}{N_r K_{\mathrm{d}} R_c \sigma_w^2} = \frac{LN_t}{K_{\mathrm{d}} R_c \sigma_w^2}. \tag{31}$$

All the simulation results to be presented are obtained by averaging over 10,000 packets.

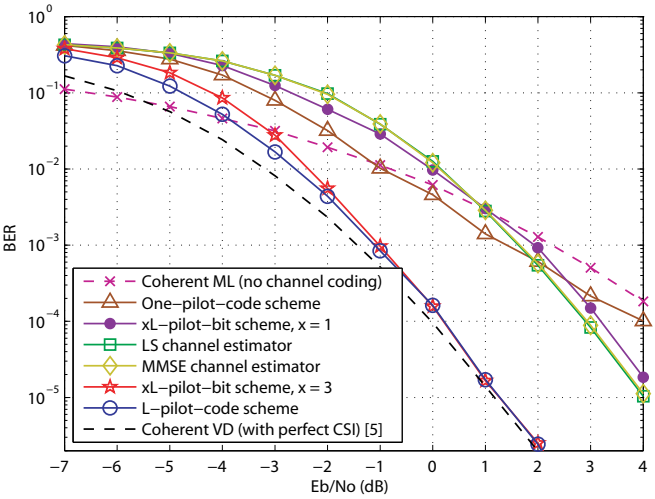

Fig. 3. BER performance comparison of point-to-point BICM-OSTBC-OFDM, using $(7,5)_8$ convolutional code with $d_{\mathrm{free}} = 5$, for $N_c = 256$, $L = 4$, $T = 4$, $N_t = 2$ and $N_r = 4$.

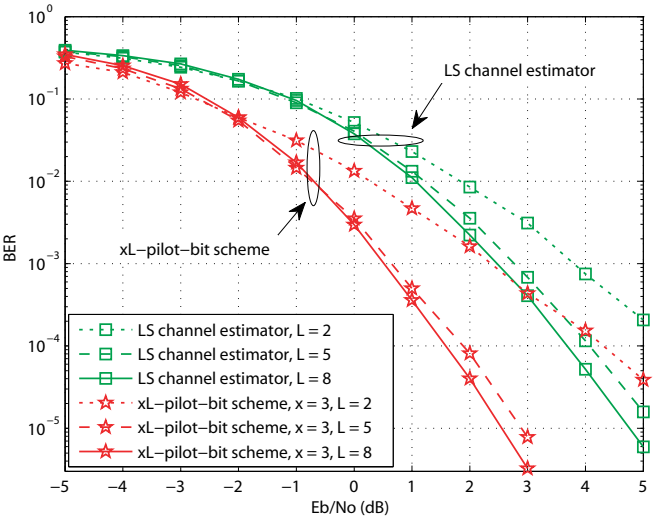

Fig. 4. BER performance of the xL-pilot-bit scheme and LS channel estimator in point-to-point BICM-OSTBC-OFDM, using $(7,5)_8$ convolutional code with $d_{\mathrm{free}} = 5$, for $N_c = 256$, $L = 2, 5, 8$, $T = 4$, $N_t = 2$ and $N_r = 3$.

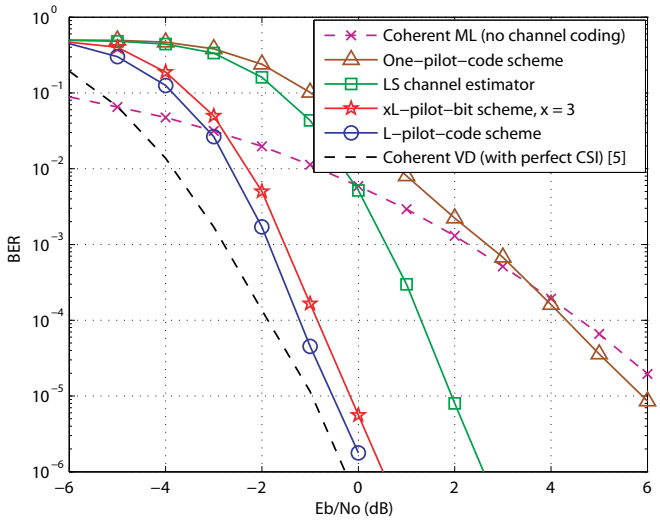

Fig. 5. BER performance comparison of point-to-point BICM-OSTBC-OFDM, using $(171, 133)_8$ convolutional code with $d_{\text{free}} = 10$, for $N_c = 256$, $L = 8$, $T = 4$, $N_t = 2$ and $N_r = 4$.

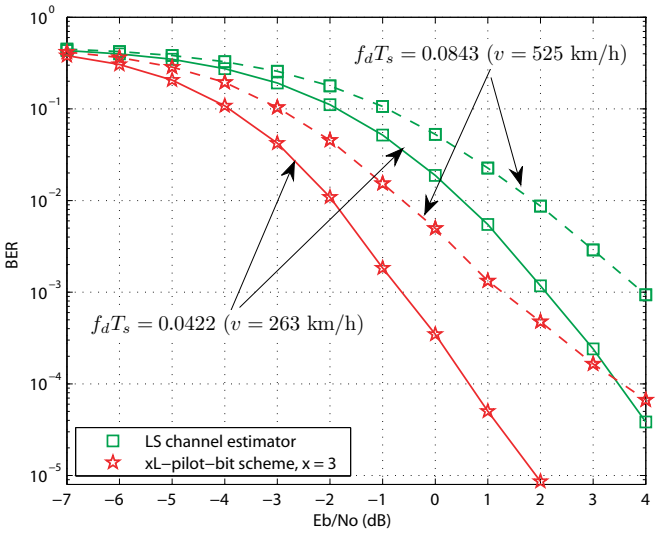

Fig. 6. BER performance comparison of point-to-point BICM-OSTBC-OFDM in mobile time-varying channels, using $(7, 5)_8$ convolutional code with $d_{\text{free}} = 5$, for $N_c = 256$, $L = 4$, $T = 4$, $N_t = 2$ and $N_r = 4$.

Example 1: Performance of point-to-point BICM-OSTBC-OFDM: Fig. 3 presents the bit error performance of the three schemes in Section IV-A for $L = 4$ and $N_r = 4$. The 1/2-rate convolutional code with generator polynomials $(7, 5)_8$ and $d_{\text{free}} = 5$ is used. In this figure, we also present the bit error rates (BERs) of the ideal coherent Viterbi decoder (VD) in (3) which has perfect CSI [5], the ideal uncoded coherent ML detector, and the pilot-aided LS and MMSE channel estimators [24]. The LS and MMSE channel estimators employ L equispaced pilot-code subcarriers for channel estimation and use the channel estimate for coherent ML decoding. There are two points to note from Fig. 3. Firstly, one can observe that the coherent VD, the LS and MMSE channel estimators, the L-pilot-code scheme, and the xL-pilot-bit scheme (with either $x = 1$ or $x = 3$) all exhibit the same diversity performance. Since the coherent VD can harvest all the diversity advantages offered by the system [5], this result implies that the L-pilot-code scheme and the xL-pilot-bit scheme also enjoy the same diversity advantages[4]. By contrast, we observe that the one-pilot-code scheme only achieves a similar diversity performance as the uncoded coherent ML detector. Since the uncoded coherent detector is known to be able to achieve the spatial diversity only, this result implies that the one-pilot-code scheme harvests at most the spatial diversity. All these results are consistent with our analyses in Theorem 1 and Corollary 1. Secondly, we observe that the xL-pilot-bit scheme with $x = 1$ has a slightly higher BER than the LS and MMSE channel estimators; whereas, when x increases to 3, the xL-pilot-bit scheme performs almost on a par with the L-pilot-code scheme and outperforms the LS and MMSE channel estimators. Specifically, as one can see, the xL-pilot-bit scheme with $x = 3$ has around 2.8 dB E_b/N_0 advantage over the LS and MMSE channel estimators at BER$= 10^{-5}$. We should note that with $x = 3$ the xL-pilot-bit scheme consumes $4L$ bits per OSTBC-OFDM block due to pilots and NIS-OSTBCs[5]; while all the L-pilot-code scheme, the LS and MMSE channel estimators have an $8L$-bit loss per OSTBC-OFDM block. Fig. 4 displays the performance comparison results between the xL-pilot-bit scheme (with $x = 3$) and the LS channel estimator, for $N_r = 3$ and for $L = 2, 5$ and 8. We can observe from the figure that both schemes have improved diversity performance as L increases from 2 to 5, but the diversity order remains the same as L increases from 5 to 8. The reason is that, according to Theorem 1, the maximum frequency diversity is $d_{\text{free}} = 5$ when $L \geq d_{\text{free}}$. One can also see from Fig. 4 that the xL-pilot-bit scheme outperforms the LS channel estimator for all L. In Fig. 5, we show the BER performance for $L = 8$ and $N_r = 4$ by using the 1/2-rate convolutional code with generator polynomials $(171, 133)_8$ and $d_{\text{free}} = 10$ (this code is popularly used in real systems). Similar comparison results as in Fig. 3 can be observed.

Example 2: Performance under mobile time-varying channels: Throughout this paper, we have used the block fading assumption to model fast fading channels. The block fading assumption is considered reasonable when the channel coherence time is no smaller than one OSTBC-OFDM block, although it is interesting to examine how our methods actually perform under a more realistic mobile time-varying fading channel. This example aims to do so. We assume that the elements of $\boldsymbol{\mathcal{H}}$ vary with time following Jakes' model [31]. The carrier frequency is set to 2.6 GHz and the sampling frequency is set to $1/T_s = 3.84$ MHz (e.g., LTE). We try two settings with the normalized Doppler frequency, namely, $N_c f_d T_s = 0.0843$ (corresponding to a moving speed 525 km/h) and $N_c f_d T_s = 0.0422$ (corresponding to a moving speed 263 km/h), where f_d is the maximum Doppler shift. One can see from Fig. 6 that the xL-pilot-bit scheme still yields consistent bit error performance and outperforms the

[4]While the corresponding log-log BER slopes are not as high as the theoretic value (which is 32), they will increase to the theoretic value as E_b/N_0 increases.

[5]The NIS-OSTBC in [15, Eqn. (15)] has one bit loss for achieving the NIS property; please see [25] for the details.

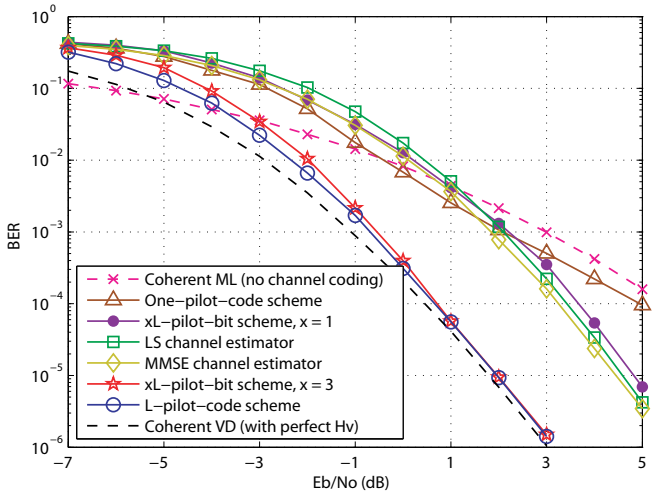

Fig. 7. BER performance comparison of distributed BICM-OSTBC-OFDM, using $(7,5)_8$ convolutional code with $d_{\text{free}} = 5$, for $N_c = 256$, $L = 4$, $N_s = 2$, $T = 4$, $N_t = 2$ and $N_r = 4$.

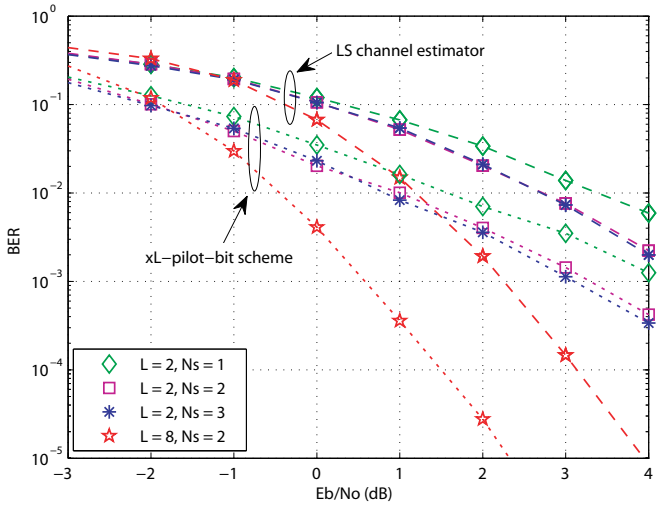

Fig. 8. BER performance of the xL-pilot-bit scheme ($x = 3$) and LS channel estimator in distributed BICM-OSTBC-OFDM, where $(171,133)_8$ convolutional code with $d_{\text{free}} = 10$ is used, $N_c = 256$, $T = 4$, $N_t = 2$ and $N_r = 3$.

LS channel estimator.

Example 3: Performance of Distributed BICM-OSTBC-OFDM: In this example, we examine the performance of the noncoherent ML decoder (10) for the distributed BICM-OSTBC-OFDM system in Section V. E_b/N_0 is defined as in (31) except that $\boldsymbol{\mathcal{G}}(\boldsymbol{s})$ is replaced by $\boldsymbol{\mathcal{G}}(\boldsymbol{s})(\mathbf{Q} \otimes \mathbf{I}_L)$. The elements of the signature matrix $\mathbf{Q}$ are randomly generated with unit modulus and fixed throughout the simulations. Fig. 7 shows the performance comparison results of different transmission schemes, for $N_s = 2$ and $N_r = 4$. It can be seen from this figure that the xL-pilot-bit scheme ($x = 3$) exhibits the same cooperative-frequency diversity order as the coherent VD and improves upon the LS and MMSE channel estimators. Fig. 8 shows the simulation results for various values of L and N_s. From this figure, one can see that the diversity order of both the xL-pilot-bit scheme and LS channel estimator improves as N_s increases from one to two or as L increases from two to eight. For $L = 2$, we see that the

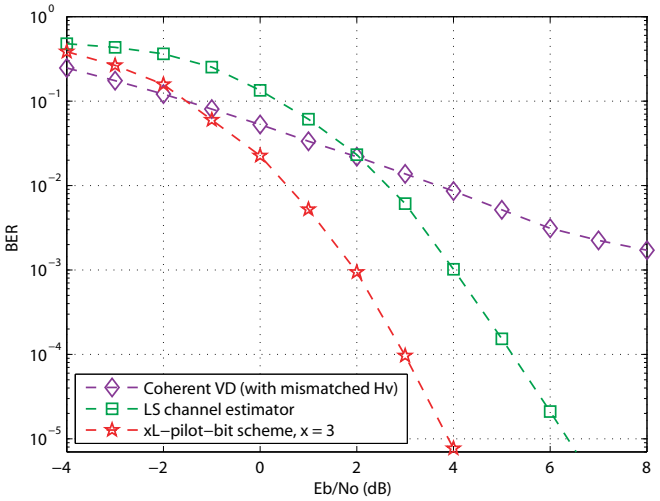

Fig. 9. BER performance comparison of distributed BICM-OSTBC-OFDM, where $(171,133)_8$ convolutional code with $d_{\text{free}} = 10$ is used, $N_c = 256$, $L = 4$, $T = 4$, $N_s = 8$, $N_t = 2$ and $N_r = 3$. There are ΔN_s (between zero and four) relays disconnecting from the network for the second half of the packet transmission.

diversity performance does not improve as N_s increases from two to three because, by Theorem 2, the maximum cooperative diversity is given by $\min\{N_t, N_s\}$. In Fig. 9, we consider a scenario where the physical channel $\boldsymbol{\mathcal{H}}$ remains unchanged during the whole packet transmission. The coherent VD is assumed to perfectly know the virtual channel $\boldsymbol{\mathcal{H}}_{\text{v}}$ through a dedicated training process before the packet transmission. However, we assume that there are ΔN_s relays out of the total N_s relays unexpectedly disconnecting from the network after the 4th OSTBC-OFDM block, and the number ΔN_s randomly changes (between 0 to 4) from block 5 to block 8. This simulates the scenario where there are ΔN_s relays failing to correctly decode the information bits from the source and thus fouling out the cooperation in that block. As a result, the virtual channel $\boldsymbol{\mathcal{H}}_{\text{v}}$ actually varies from block 5 to block 8, making the coherent VD (that uses $\boldsymbol{\mathcal{H}}_{\text{v}}$ known at the beginning for every block) have mismatched virtual channel information from block 5 to block 8. Fig. 9 presents the simulation results for $N_s = 8$ and $N_r = 3$. One can observe that the xL-pilot-bit scheme is more robust than the coherent VD and can still yield consistent BER performance even though $\boldsymbol{\mathcal{H}}_{\text{v}}$ changes from block to block in the second half transmission of the packet.

VII. Conclusions

In this paper, we have presented a block-wise noncoherent ML decoder in (10) for the BICM-OSTBC-OFDM system, and a complexity-reduced implementation method. In addition, we have analyzed the diversity order of the noncoherent ML decoder. We have presented a worst-case diversity analysis framework and have shown that the PCI-achieving schemes, such as the L-pilot-code scheme and the xL-pilot-bit scheme, can achieve the maximum spatial-frequency diversity of the BICM-OSTBC-OFDM system. The developed noncoherent ML decoder and diversity analysis are further extended to the relay-assisted distributed BICM-OSTBC-OFDM system. All

our theoretical claims have been corroborated by the presented simulation results.

APPENDIX A
PROOF OF PROPOSITION 1

By (22) and the fact that $\bar{c}_{k_\ell} \neq \hat{c}_{k_\ell}$ for all $\ell = 1, \ldots, d$, the PEP $\Pr(\bar{\mathbf{c}} \to \hat{\mathbf{c}}|\bar{\mathbf{c}})$ can be upper bounded in (32) as shown on the top of the next page, where $\boldsymbol{s_y} = [(\boldsymbol{s_y}^{(1)})^T, \ldots, (\boldsymbol{s_y}^{(d)})^T]^T$, $\boldsymbol{s_y}^{(\ell)} = \boldsymbol{\Pi}[\mathbf{s}_{\rm p}^T, (\mathbf{s}_{{\rm d},\boldsymbol{y}}^{(\ell)})^T]^T$,

$$\mathbf{s}_{{\rm d},\boldsymbol{y}}^{(\ell)} = \arg \max_{\substack{\mathbf{s}_{\rm d} \in \{\pm 1\}^{K_{\rm d}} \\ s_{n_\ell, i_\ell} = \mu(\hat{c}_{k_\ell})}} \boldsymbol{y}^H \boldsymbol{\mathcal{G}}(\boldsymbol{s}) \boldsymbol{\mathcal{G}}^H(\boldsymbol{s}) \boldsymbol{y}, \tag{33}$$

and $\boldsymbol{\Delta}(\boldsymbol{s_y}) \triangleq \sum_{\ell=1}^{d} \left(\boldsymbol{\mathcal{G}}(\bar{\boldsymbol{s}}) \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) - \boldsymbol{\mathcal{G}}(\boldsymbol{s_y}^{(\ell)}) \boldsymbol{\mathcal{G}}^H(\boldsymbol{s_y}^{(\ell)}) \right)$. By substituting (21) into (32) and by applying the fact that $\Pr(\alpha_1 + \alpha_2 \geq t) \leq \Pr(\alpha_1 \geq t/2 \text{ or } \alpha_2 \geq t/2) \leq \Pr(\alpha_1 \geq t/2) + \Pr(\alpha_2 \geq t/2)$ for any two random variables α_1 and α_2, we further obtain (34) on the next page.

To compute P_1 and P_2 in (34), let us define $\boldsymbol{S}(\hat{\mathbf{c}})$ shown in (35). Also, for each $\hat{\boldsymbol{s}} \in \boldsymbol{S}(\hat{\mathbf{c}})$, define a set $\boldsymbol{R}(\hat{\boldsymbol{s}})$ in (36), i.e., $\boldsymbol{s_y} = \hat{\boldsymbol{s}}$ for all $\boldsymbol{w} \in \boldsymbol{R}(\hat{\boldsymbol{s}})$. According to the law of total probability, we can compute P_1 in (37) as shown on the next page. To compute right-hand side (RHS) of (37), noting that by the fact that $\boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\mathcal{G}}(\bar{\boldsymbol{s}}) = \mathbf{I}_{LN_t}$, $\boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\Delta}(\hat{\boldsymbol{s}})$ can be expressed in (38) shown on the next page, where $\boldsymbol{\Psi}_\ell^\perp \triangleq \mathbf{I}_{N_c T} - \boldsymbol{\mathcal{G}}(\hat{\boldsymbol{s}}^{(\ell)}) \boldsymbol{\mathcal{G}}^H(\hat{\boldsymbol{s}}^{(\ell)})$ is a projection matrix, i.e., $(\boldsymbol{\Psi}_\ell^\perp)^2 = \boldsymbol{\Psi}_\ell^\perp$. By (38), we can obtain $\boldsymbol{h}^H \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\Delta}(\hat{\boldsymbol{s}}) \boldsymbol{w} = \boldsymbol{h}^H \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \left(\sum_{\ell=1}^{d} \boldsymbol{\Psi}_\ell^\perp \right) \boldsymbol{w}$. Then $\Re\{\boldsymbol{h}^H \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\Delta}(\hat{\boldsymbol{s}}) \boldsymbol{w}\}$ in (37) is real Gaussian distributed with zero mean and variance $(\sigma_w^2/2) \boldsymbol{h}^H \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \left(\sum_{\ell=1}^{d} \boldsymbol{\Psi}_\ell^\perp \right)^2 \boldsymbol{\mathcal{G}}(\bar{\boldsymbol{s}}) \boldsymbol{h}$. Hence, P_1 can be upper bounded in (39) as shown on the next page, where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-\frac{u^2}{2}) du$. By the fact that $\lambda_i((\frac{1}{d} \sum_{\ell=1}^{d} \boldsymbol{\Psi}_\ell^\perp)^2) \leq \lambda_i(\frac{1}{d} \sum_{\ell=1}^{d} \boldsymbol{\Psi}_\ell^\perp)$ for all i and the inequality $Q(x) \leq \frac{1}{2} \exp(-\frac{x^2}{2})$, one can further upper bound P_1 as

$$P_1 \leq \sum_{\hat{\boldsymbol{s}} \in \boldsymbol{S}(\hat{\mathbf{c}})} \frac{1}{2} \exp\left(-\frac{1}{16 \sigma_w^2} \boldsymbol{h}^H \boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}}) \boldsymbol{h} \right), \tag{40}$$

where $\boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}})$ is defined in (25). On the other hand, using the similar idea as in (37) for handling P_1 and by applying [32, Lemma 2], one can show that P_2 can be upper bounded as

$$P_2 \leq \sum_{\hat{\boldsymbol{s}} \in \boldsymbol{S}(\hat{\mathbf{c}})} \xi_1 \exp\left(-\frac{\xi_2 d}{2 \sigma_w^2} \boldsymbol{h}^H \boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}}) \boldsymbol{h} \right), \tag{41}$$

where $\xi_1 > 0$ and $\xi_2 > 0$ are constants. By substituting (40), (41) into (34), we then obtain

$$\Pr(\bar{\mathbf{c}} \to \hat{\mathbf{c}}|\bar{\mathbf{c}}) \leq \sum_{\hat{\boldsymbol{s}} \in \boldsymbol{S}(\hat{\mathbf{c}})} \mathbb{E}_{\boldsymbol{h}} \left\{ \beta_1 \exp\left(-\frac{\beta_2}{\sigma_w^2} \boldsymbol{h}^H \boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}}) \boldsymbol{h} \right) \right\}, \tag{42}$$

where $\beta_1 = \xi_1 + \frac{1}{2} > \frac{1}{2}$, $0 < \beta_2 = \min\{\frac{1}{16}, \frac{\xi_2 d}{2}\} \leq \frac{1}{16}$. Finally, the end result in (23) is obtained by deriving the expectation in (42) with respect to the i.i.d. complex Gaussian $\boldsymbol{h}$ [20]. ■

APPENDIX B
DERIVATION OF (27)

The RHS of (23) can be bounded as

$$\begin{aligned} \Pr(\bar{\mathbf{c}} \to \hat{\mathbf{c}}|\bar{\mathbf{c}}) &\leq \beta_1 \sum_{\hat{\boldsymbol{s}} \in \boldsymbol{S}(\hat{\mathbf{c}})} \det{}^{-1} \left(\mathbf{I}_{LN_t} + \frac{\sigma_h^2 \beta_2}{\sigma_w^2} \boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}}) \right) \\ &\leq \beta_1 \beta_3 \sum_{\hat{\boldsymbol{s}} \in \boldsymbol{S}(\hat{\mathbf{c}})} \left(\frac{\beta_2 \sigma_h^2}{\sigma_w^2} \right)^{-r(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}})}, \end{aligned} \tag{43}$$

where $\beta_3 = \max_{\hat{\boldsymbol{s}} \in \boldsymbol{S}(\hat{\mathbf{c}})} \prod_{i=1}^{\operatorname{rank}(\boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}}))} \frac{1}{\lambda_i(\boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}}))}$, and $r(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}}) = \operatorname{rank}(\boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}}))$. From the RHS of the above inequality, we further obtain, for $\beta_2 \sigma_h^2 / \sigma_w^2 > 1$, that

$$\begin{aligned} \Pr(\bar{\mathbf{c}} \to \hat{\mathbf{c}}|\bar{\mathbf{c}}) &\leq \beta_1 \beta_3 |\boldsymbol{S}(\hat{\mathbf{c}})| \max_{\hat{\boldsymbol{s}} \in \boldsymbol{S}(\hat{\mathbf{c}})} \left(\frac{\beta_2 \sigma_h^2}{\sigma_w^2} \right)^{-r(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}})} \\ &= \beta_1 \beta_3 |\boldsymbol{S}(\hat{\mathbf{c}})| \left(\frac{\beta_2 \sigma_h^2}{\sigma_w^2} \right)^{-\min\limits_{\hat{\boldsymbol{s}} \in \boldsymbol{S}(\hat{\mathbf{c}})} r(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}})}. \end{aligned} \tag{44}$$

Plugging the upper bound in (44) into (26), we obtain the diversity lower bound in (27). ■

APPENDIX C
PROOF OF THEOREM 1

By (25), $\boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}})$ can be expressed as

$$\boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}}) = \mathbf{I}_{LN_t} - \mathbf{Z}^H \mathbf{W} \mathbf{W}^H \mathbf{Z} \succeq \mathbf{0}, \tag{45}$$

where $\mathbf{W} = \operatorname{blkdiag}\left(\boldsymbol{\mathcal{G}}(\hat{\boldsymbol{s}}^{(1)}), \boldsymbol{\mathcal{G}}(\hat{\boldsymbol{s}}^{(2)}), \ldots, \boldsymbol{\mathcal{G}}(\hat{\boldsymbol{s}}^{(d)}) \right) \in \mathbb{C}^{dN_c T \times dLN_t}$ (which is a block diagonal matrix) and $\mathbf{Z} = \frac{1}{\sqrt{d}} (\mathbf{1}_d \otimes \boldsymbol{\mathcal{G}}(\bar{\boldsymbol{s}})) \in \mathbb{C}^{dN_c T \times LN_t}$. Since $\mathbf{W}^H \mathbf{W} = \mathbf{I}_{dLN_t}$ and $\mathbf{Z}^H \mathbf{Z} = \mathbf{I}_{LN_t}$, we have $\|\mathbf{W}^H \mathbf{Z}\|_2 \leq \|\mathbf{W}\|_2 \|\mathbf{Z}\|_2 = 1$, which implies that all the singular values of $\mathbf{W}^H \mathbf{Z}$ are no larger than one. Hence determining the rank of $\boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}})$ is equivalent to identifying the number of the singular values of $\mathbf{W}^H \mathbf{Z}$ equal to one. Let η be the number of singular values of $\mathbf{W}^H \mathbf{Z}$ equal to one. Then $\operatorname{rank}(\boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}})) = LN_t - \eta$. To proceed, we will need the following lemma which is proved in [16]:

Lemma 1 [16] *The matrix $\mathbf{W}^H \mathbf{Z} \in \mathbb{C}^{dLN_t \times LN_t}$ has $\eta \geq 1$ singular values equal to one if and only if there exist two linearly independent sets $\{\mathbf{x}_1, \ldots, \mathbf{x}_\eta\} \in \mathbb{C}^{dLN_t}$ and $\{\mathbf{y}_1, \ldots, \mathbf{y}_\eta\} \in \mathbb{C}^{LN_t}$ such that $\mathbf{W}\mathbf{x}_m = \mathbf{Z}\mathbf{y}_m$, $m = 1, \ldots, \eta$.*

According to the above lemma, we obtain that

$$\begin{aligned} \boldsymbol{\mathcal{G}}(\hat{\boldsymbol{s}}^{(\ell)}) \mathbf{x}_m^{(\ell)} &= \boldsymbol{\mathcal{G}}(\bar{\boldsymbol{s}}) \left(\mathbf{y}_m / \sqrt{d} \right), \\ \ell &= 1, \ldots, d, \ m = 1, \ldots, \eta, \end{aligned} \tag{46}$$

where $\mathbf{x}_m = [(\mathbf{x}_m^{(1)})^T, \ldots, (\mathbf{x}_m^{(d)})^T]^T$. Firstly, as mentioned in Section IV-A, for PCI-achieving schemes such as the L-pilot-code and xL-pilot-bit schemes, (46) holds only if $\mathbf{x}_m^{(\ell)} = \frac{1}{\sqrt{d}} \mathbf{y}_m$, $\ell = 1, \ldots, d$, $m = 1, \ldots, \eta$. Secondly, since $\bar{c}_{k'_\ell} \neq \hat{c}_{k'_\ell}$ for $\ell = 1, \ldots, d_{\rm free}$, it holds that $\hat{\mathbf{s}}_{n'_\ell}^{(\ell)} \neq \bar{\mathbf{s}}_{n'_\ell}$ for $\ell = 1, \ldots, d_{\rm free}$; moreover, under **A1)**, we have $n'_i \neq n'_k$ for all

$$\Pr(\bar{\mathbf{c}} \to \hat{\mathbf{c}}|\bar{\mathbf{c}}) = \Pr\left(\sum_{\ell=1}^{d} \max_{\substack{\mathbf{s}_{\mathrm{d}} \in \{\pm 1\}^{K_{\mathrm{d}}} \\ s_{n_\ell, i_\ell} = \mu(\bar{c}_{k_\ell})}} \boldsymbol{y}^H \boldsymbol{\mathcal{G}}(\boldsymbol{s}) \boldsymbol{\mathcal{G}}^H(\boldsymbol{s}) \boldsymbol{y} \le \sum_{\ell=1}^{d} \max_{\substack{\mathbf{s}_{\mathrm{d}} \in \{\pm 1\}^{K_{\mathrm{d}}} \\ s_{n_\ell, i_\ell} = \mu(\hat{c}_{k_\ell})}} \boldsymbol{y}^H \boldsymbol{\mathcal{G}}(\boldsymbol{s}) \boldsymbol{\mathcal{G}}^H(\boldsymbol{s}) \boldsymbol{y} \middle| \bar{\mathbf{c}}\right)$$
$$\le \Pr\left(\sum_{\ell=1}^{d} \left(\boldsymbol{y}^H \boldsymbol{\mathcal{G}}(\bar{\boldsymbol{s}}) \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{y} - \boldsymbol{y}^H \boldsymbol{\mathcal{G}}(\boldsymbol{s}_{\boldsymbol{y}}^{(\ell)}) \boldsymbol{\mathcal{G}}^H(\boldsymbol{s}_{\boldsymbol{y}}^{(\ell)}) \boldsymbol{y}\right) \le 0 \middle| \bar{\mathbf{c}}\right) = \Pr\left(\boldsymbol{y}^H \boldsymbol{\Delta}(\boldsymbol{s}_{\boldsymbol{y}}) \boldsymbol{y} \le 0 | \bar{\mathbf{c}}\right) \tag{32}$$

$$\Pr\left(\boldsymbol{y}^H \boldsymbol{\Delta}(\boldsymbol{s}_{\boldsymbol{y}}) \boldsymbol{y} \le 0 | \bar{\mathbf{c}}\right)$$
$$= \mathbb{E}_{\boldsymbol{h}} \left\{ \Pr\left(-2\Re\{\boldsymbol{h}^H \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\Delta}(\boldsymbol{s}_{\boldsymbol{y}}) \boldsymbol{w}\} - \boldsymbol{w}^H \boldsymbol{\Delta}(\boldsymbol{s}_{\boldsymbol{y}}) \boldsymbol{w} \ge \boldsymbol{h}^H \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\Delta}(\boldsymbol{s}_{\boldsymbol{y}}) \boldsymbol{\mathcal{G}}(\bar{\boldsymbol{s}}) \boldsymbol{h} | \bar{\mathbf{c}}, \boldsymbol{h}\right) \right\} \le \mathbb{E}_{\boldsymbol{h}} \{P_1 + P_2\} \tag{34}$$

where

$$P_1 \triangleq \Pr\left(-2\Re\{\boldsymbol{h}^H \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\Delta}(\boldsymbol{s}_{\boldsymbol{y}}) \boldsymbol{w}\} \ge (1/2) \boldsymbol{h}^H \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\Delta}(\boldsymbol{s}_{\boldsymbol{y}}) \boldsymbol{\mathcal{G}}(\bar{\boldsymbol{s}}) \boldsymbol{h} | \bar{\mathbf{c}}, \boldsymbol{h}\right)$$
$$P_2 \triangleq \Pr\left(|\boldsymbol{w}^H \boldsymbol{\Delta}(\boldsymbol{s}_{\boldsymbol{y}}) \boldsymbol{w}| \ge (1/2) \boldsymbol{h}^H \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\Delta}(\boldsymbol{s}_{\boldsymbol{y}}) \boldsymbol{\mathcal{G}}(\bar{\boldsymbol{s}}) \boldsymbol{h} | \bar{\mathbf{c}}, \boldsymbol{h}\right)$$

$$\boldsymbol{S}(\hat{\mathbf{c}}) = \{\hat{\boldsymbol{s}} = [(\hat{\boldsymbol{s}}^{(1)})^T, \ldots, (\hat{\boldsymbol{s}}^{(d)})^T]^T \mid \hat{\boldsymbol{s}}^{(\ell)} = \boldsymbol{\Pi}[\mathbf{s}_{\mathrm{p}}^T, (\hat{\mathbf{s}}_{\mathrm{d}}^{(\ell)})^T]^T,\ \hat{\mathbf{s}}_{\mathrm{d}}^{(\ell)} \in \{\pm 1\}^{K_{\mathrm{d}}}, \hat{s}_{n_\ell, i_\ell}^{(\ell)} = \mu(\hat{c}_{k_\ell})\} \tag{35}$$

$$\boldsymbol{R}(\hat{\boldsymbol{s}}) = \left\{ \boldsymbol{w} \mid\ \hat{\mathbf{s}}_{\mathrm{d}}^{(\ell)} = \arg \max_{\substack{\mathbf{s}_{\mathrm{d}} \in \{\pm 1\}^{K_{\mathrm{d}}} \\ s_{n_\ell, i_\ell} = \mu(\hat{c}_{k_\ell})}} \boldsymbol{y}^H \boldsymbol{\mathcal{G}}(\boldsymbol{s}) \boldsymbol{\mathcal{G}}^H(\boldsymbol{s}) \boldsymbol{y},\ \ell = 1, \ldots, d \right\} \tag{36}$$

$$P_1 = \sum_{\hat{\boldsymbol{s}} \in \boldsymbol{S}(\hat{\mathbf{c}})} \Pr\left(-2\Re\{\boldsymbol{h}^H \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\Delta}(\boldsymbol{s}_{\boldsymbol{y}}) \boldsymbol{w}\} \ge (1/2) \boldsymbol{h}^H \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\Delta}(\boldsymbol{s}_{\boldsymbol{y}}) \boldsymbol{\mathcal{G}}(\bar{\boldsymbol{s}}) \boldsymbol{h}, \boldsymbol{w} \in \boldsymbol{R}(\hat{\boldsymbol{s}}) | \bar{\mathbf{c}}, \boldsymbol{h}\right)$$
$$= \sum_{\hat{\boldsymbol{s}} \in \boldsymbol{S}(\hat{\mathbf{c}})} \Pr\left(-2\Re\{\boldsymbol{h}^H \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\Delta}(\hat{\boldsymbol{s}}) \boldsymbol{w}\} \ge (1/2) \boldsymbol{h}^H \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\Delta}(\hat{\boldsymbol{s}}) \boldsymbol{\mathcal{G}}(\bar{\boldsymbol{s}}) \boldsymbol{h}, \boldsymbol{w} \in \boldsymbol{R}(\hat{\boldsymbol{s}}) | \bar{\mathbf{c}}, \boldsymbol{h}\right)$$
$$\le \sum_{\hat{\boldsymbol{s}} \in \boldsymbol{S}(\hat{\mathbf{c}})} \Pr\left(-2\Re\{\boldsymbol{h}^H \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\Delta}(\hat{\boldsymbol{s}}) \boldsymbol{w}\} \ge (1/2) \boldsymbol{h}^H \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\Delta}(\hat{\boldsymbol{s}}) \boldsymbol{\mathcal{G}}(\bar{\boldsymbol{s}}) \boldsymbol{h} | \bar{\mathbf{c}}, \boldsymbol{h}\right) \tag{37}$$

$$\boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\Delta}(\hat{\boldsymbol{s}}) = \sum_{\ell=1}^{d} \left(\boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\mathcal{G}}(\bar{\boldsymbol{s}}) \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) - \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \boldsymbol{\mathcal{G}}(\hat{\boldsymbol{s}}^{(\ell)}) \boldsymbol{\mathcal{G}}^H(\hat{\boldsymbol{s}}^{(\ell)})\right) = \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \left(\sum_{\ell=1}^{d} \boldsymbol{\Psi}_\ell^\perp\right) \tag{38}$$

$$P_1 \le \sum_{\hat{\boldsymbol{s}} \in \boldsymbol{S}(\hat{\mathbf{c}})} Q\left(\frac{1}{4} \boldsymbol{h}^H \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \left(\frac{1}{d} \sum_{\ell=1}^{d} \boldsymbol{\Psi}_\ell^\perp\right) \boldsymbol{\mathcal{G}}(\bar{\boldsymbol{s}}) \boldsymbol{h} \Big/ \sqrt{(\frac{\sigma_w^2}{2}) \boldsymbol{h}^H \boldsymbol{\mathcal{G}}^H(\bar{\boldsymbol{s}}) \left(\frac{1}{d} \sum_{\ell=1}^{d} \boldsymbol{\Psi}_\ell^\perp\right)^2 \boldsymbol{\mathcal{G}}(\bar{\boldsymbol{s}}) \boldsymbol{h}}\right) \tag{39}$$

$i \ne k$. Since $\mathbf{C}_{n'_\ell}(\hat{\mathbf{s}}_{n'_\ell}^{(\ell)}) - \mathbf{C}_{n'_\ell}(\bar{\mathbf{s}}_{n'_\ell})$ is of full column rank [20], it follows from (46), (6) and the above two facts that

$$\boldsymbol{\Upsilon} \mathbf{y}_m \triangleq [(\mathbf{I}_{N_t} \otimes \mathbf{f}_{n'_1}^T)^T, \ldots, (\mathbf{I}_{N_t} \otimes \mathbf{f}_{n'_{d_{\mathrm{free}}}}^T)^T]^T \mathbf{y}_m = \mathbf{0},\ m = 1, \ldots, \eta. \tag{47}$$

Note that $\boldsymbol{\Upsilon} \in \mathbb{C}^{d_{\mathrm{free}} N_t \times L N_t}$ has full rank, a property implied by the Fourier structure of $\mathbf{f}_n$. If $d_{\mathrm{free}} \ge L$, then $\boldsymbol{\Upsilon}$ is tall and (47) holds only if $\eta = 0$. On the other hand, if $d_{\mathrm{free}} < L$, then the nullity of $\boldsymbol{\Upsilon}$ is $LN_t - d_{\mathrm{free}} N_t$ implying that $\eta = LN_t - d_{\mathrm{free}} N_t$. Therefore, we obtain $\mathrm{rank}(\boldsymbol{\Omega}(\bar{\boldsymbol{s}}, \hat{\boldsymbol{s}})) = N_t \min(d_{\mathrm{free}}, L)$. ■

APPENDIX D
PROOF OF COROLLARY 1

Suppose that $\mathbf{s}_{\mathrm{p}} = \mathbf{s}_1$ for the one-pilot-code scheme. Consider a special case of $\mathbf{C}_n(\hat{\mathbf{s}}_n^{(\ell)}) = \varphi_\ell \mathbf{C}_n(\bar{\mathbf{s}}_n)$, $n = 2, \ldots, N_c$, $\ell = 1, \ldots, d$, where $\varphi_\ell \in \mathbb{C}$, $|\varphi_\ell| = 1$, for all ℓ. Then (46) holds for $\mathbf{x}_m^{(\ell)} = \varphi_\ell^* \mathbf{y}_m / \sqrt{d}$, $\ell = 1, \ldots, d$, and

$(\mathbf{I}_{N_t}\otimes\mathbf{f}_1^T)\mathbf{y}_m=\mathbf{0}$, $m=1,\ldots,\eta$. Since the nullity of $\mathbf{I}_{N_t}\otimes\mathbf{f}_1^T$ is LN_t-N_t, $\mathbf{W}^H\mathbf{Z}$ can have $\eta=LN_t-N_t$ singular values equal to one, implying that $\mathrm{rank}(\mathbf{\Omega}(\bar{\boldsymbol{s}},\hat{\boldsymbol{s}}))=N_t$ for this special case. Hence, we have $D_{\mathrm{NC}}^{\star}\leq N_t$. ■

REFERENCES

[1] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[2] A. Guillén i Fàbregas, A. Martinez, and G. Caire, "Bit-interleaved coded modulation," *Found. Trends Commun. Inf. Theory*, vol. 5, no. 1-2, pp. 1–153, Jan. 2008.

[3] Z. Hong and B. L. Hughes, "Bit-interleaved space-time coded modulation with iterative decoding," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1912–1917, Nov. 2004.

[4] M. Teimouri, A. Hedayat, and M. Shiva, "Concatenated bit-interleaved coded modulation and orthogonal space-time block codes over fading channels," *IET Commun.*, vol. 3, no. 8, pp. 1354–1362, Aug. 2009.

[5] E. Akay and E. Ayanoglu, "Achieving full frequency and space diversity in wireless systems via BICM, OFDM, STBC, and Viterbi decoding," *IEEE Trans. Commun.*, vol. 54, no. 12, pp. 2164–2172, Dec. 2006.

[6] T. Islam, R. Schober, R. K. Mallik, and V. K. Bhargava, "Analysis and design of cooperative BICM-OFDM systems," *IEEE Trans. Commun.*, vol. 59, no. 6, pp. 1742–1751, June 2011.

[7] P. Loskot and N. C. Beaulieu, "Approximate performance analysis of coded OSTBC-OFDM systems over arbitrary correlated generalized Ricean fading channels," *IEEE Trans. Commun.*, vol. 57, no. 8, pp. 2235–2238, Aug. 2009.

[8] L. Hanzo, M. El-Hajjar, and O. Alamri, "Near-capacity wireless transceivers and cooperative communications in the MIMO era: evolution of standards, waveform design, and future perspectives," *Proc. IEEE*, vol. 99, no. 8, pp. 1343–1385, Aug. 2011.

[9] Y. Li and X.-G. Xia, "Constellation mapping for space-time matrix modulation with iterative decomdulation/decoding," *IEEE Trans. Commun.*, vol. 53, no. 5, pp. 764–768, May 2005.

[10] N. H. Tran, H. H. Nguyen, and T. Le-Ngoc, "Coded unitary space-time modulation with iterative decoding: error performance and mapping design," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 703–716, Apr. 2007.

[11] L. H.-J. Lampe and R. Schober, "Bit-interleaved coded differential space-time modulation," *IEEE Trans. Commun.*, vol. 50, no. 9, pp. 1429–1439, Sep. 2002.

[12] J. Via, I. Santamaria, and L. Vielva, "A new subspace method for blind estimation of selective MIMO-STBC channels," *Wireless Commun. Mob. Comput.*, vol. 10, no. 11, pp. 88–102, Nov. 2010.

[13] N. Sarmadi, S. Shahbazpanahi, and A. B. Gershman, "Blind channel estimation in orthogonally coded MIMO-OFDM systems: a semidefinite relaxation approach," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2354–2364, June 2009.

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

[15] T.-H. Chang, W.-K. Ma, C.-Y. Huang, and C.-Y. Chi, "On perfect channel identifiability of semiblind ML detection of orthogonal space-time block coded OFDM," in *Proc. 2009 IEEE ICASSP*, pp. 2713–2716.

[16] ——, "Noncoherent OSTBC-OFDM for MIMO and cooperative communications: perfect channel identifiability and achievable diversity order," to appear in *IEEE Trans. Signal Process.*, DOI: 10.1109/TSP.2012.2200476.

[17] M. Brehler and M. K. Varanasi, "Asymptotic error probability analysis of qudaratic receivers in Rayleigh-fading channels with application to a unified analysis of coherent and noncoherent space-time receivers," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2383–2399, Sep. 2001.

[18] H. E. Gamal, D. Aktas, and M. O. Damen, "Noncoherent space-time coding: an algebraic perspective," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2380–2390, July 2005.

[19] Z. Liu, G. B. Giannakis, S. Barbarossa, and A. Scaglione, "Transmit antennae space-time block coding for generalized OFDM in the presence of unknown multipath," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 7, pp. 1352–1364, July 2001.

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

[21] S. Lin and D. J. Costello, *Error Control Coding*, 2nd edition. Prentice-Hall, 2004.

[22] 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, no. 2, pp. 738–751, Feb. 2006.

[23] H.-T. Wai, W.-K. Ma, and A. M.-C. So, "Cheap semidefinite relaxation MIMO detection using row-by-row block coordinate descent," in *Proc. 2011 IEEE ICASSP*, pp. 3256–3259.

[24] M. K. Ozdemir and H. Arslan, "Channel estimation for wireless OFDM systems," *IEEE Commun. Surveys & Tutorials*, vol. 9, no. 2, pp. 18–48, Second Quarter 2007.

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

[26] J.-K. Zhang and W.-K. Ma, "Full diversity blind Alamouti space-time block codes for unique identification of flat-fading channels," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 635–644, Feb. 2009.

[27] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.

[28] S. Yiu, R. Schober, and L. Lampe, "Distributed space-time block coding," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1195–1206, July 2006.

[29] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.

[30] IEEE 802.11 Working Group on Broadband Wireless Access, IEEE Standard for Local and Metropolitan Area Networks- Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: High-speed physical layer in the 5 GHZ Band, Sep. 1999.

[31] W. C. Jakes and D. C. Cox, *Microwave Mobile Communications*. Wiley-IEEE, 1994.

[32] E. G. Larsson, "Diversity and channel estimation errors," *IEEE Trans. Commun.*, vol. 52, no. 2, pp. 205–208, Feb. 2004.

Ye Yang received the B.S. degree in communication engineering from Xidian University, Xi'an, China, in 2008. Currently, he is working toward the Ph.D. degree in communications and information systems at the same university. From January to July 2011, he was a visiting Ph.D. student at National Tsing Hua University, Hsinchu, Taiwan. From February to August 2012, he was a visiting scholar at The Chinese University of Hong Kong, Hong Kong. His research interests include cooperative communications, multiple-input multiple-output (MIMO)-OFDM systems, diversity techniques and physical-layer security.

Tsung-Hui Chang (S'07-M'08) received his B.S. degree in electrical engineering and his Ph.D. degree in communications engineering from the National Tsing Hua University (NTHU), Hsinchu, Taiwan, in 2003 and 2008, respectively. He was an exchange Ph.D. student of University of Minnesota, Minneapolis, MN, USA, a visiting scholar of The Chinese University of Hong Kong, Hong Kong, and a postdoctoral researcher in Institute of Communications Engineering, NTHU. He currently works as a postdoctoral researcher with the Department of Electrical and Computer Engineering, University of California, Davis. His research interests are widely in wireless communications, digital signal processing and convex optimization and its applications.

Wing-Kin Ma (M'01-SM'11) received the B.Eng. (first-class honors) degree in electrical and electronic engineering from the University of Portsmouth, Portsmouth, U.K., in 1995, and the M.Phil. and Ph.D. degrees, both in electronic engineering, from the Chinese University of Hong Kong (CUHK), Hong Kong, in 1997 and 2001, respectively. His Ph.D. dissertation was commended to be "of very high quality and well deserved honorary mentioning" by the Faculty of Engineering, CUHK, in 2001. He is currently an Assistant Professor with the Department of Electronic Engineering, CUHK. From 2005 to 2007, he was also an Assistant Professor with the Institute of Communications Engineering, National Tsing Hua University, Taiwan, R.O.C., where he is still holding an adjunct position. Prior to becoming a faculty member, he held various research positions with McMaster University, Canada; CUHK; and the University of Melbourne, Australia. His research interests are in signal processing and communications,with a recent emphasis on MIMO communication, convex optimization, and blind signal processing.

Dr. Ma currently serves as an Associate Editor of the IEEE Signal Processing Letters. From 2005 to 2009, he has served as an Associate Editor of the IEEE Transactions on Signal Processing. He is also serving or has served as a Guest Editor of the IEEE Journal of Selected Areas in Communications on the Special Issue titled "Signal Processing Techniques for Wireless Physical Layer Security," and the IEEE Signal Processing Magazine on the Special Issue titled "Convex Optimization for Signal Processing" in May 2010. He is currently a Member of the Signal Processing Theory and Methods (SPTM) Technical Committee. He was a Tutorial Speaker in EUSIPCO 2011 titled "Semidefinite Relaxation of Nonconvex Quadratic Optimization: A Key Technique in Signal Processing and Communications" (co-taught with Anthony M.-C. So). Dr. Ma's student has won an ICASSP 2011 Best Student Paper Award, and he is co-recipient of a WHISPERS 2011 Best Paper Award.

Jianhua Ge was born in Jiangsu Province, China, in 1961. He is currently a Professor and the Deputy Director of the State Key Laboratory of Integrated Services Networks, Department of Communication Engineering, Xidian University, Xi'an, China. He has worked on digital television (DTV) standardization as a DTV technical expert. His research interests include digital video broadcasting system, multiple-input multiple-output, and mobile communication techniques.

Chong-Yung Chi (S'83-M'83-SM'89) received the Ph.D. degree in Electrical Engineering from the University of Southern California, Los Angeles, California, in 1983. From 1983 to 1988, he was with the Jet Propulsion Laboratory, Pasadena, California. He has been a Professor with the Department of Electrical Engineering since 1989 and the Institute of Communications Engineering (ICE) since 1999 (also the Chairman of ICE during 2002-2005), National Tsing Hua University, Hsinchu, Taiwan. He has published more than 190 technical papers, including more than 65 journal papers (mostly in IEEE Trans. Signal Processing), 2 book chapters and more than 120 peer-reviewed conference papers, as well as a graduate-level textbook, Blind Equalization and System Identification, Springer-Verlag, 2006. His current research interests include signal processing for wireless communications, convex analysis and optimization for blind source separation, biomedical and hyperspectral image analysis.

Dr. Chi is a senior member of IEEE. He has been a Technical Program Committee member for many IEEE sponsored and cosponsored workshops, symposiums and conferences on signal processing and wireless communications, including Co-organizer and General Co-chairman of 2001 IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), and Co-Chair of Signal Processing for Communications (SPC) Symposium, ChinaCOM 2008 and Lead Co-Chair of SPC Symposium, ChinaCOM 2009. He was an Associate Editor of IEEE TRANSACTION ON SIGNAL PROCESSING (5/2001-4/2006), IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II (1/2006-12/2007), IEEE TRANSACTION ON CIRCUITS AND SYSTEMS I (1/2008-12/2009), Associate Editor of IEEE SIGNAL PROCESSING LETTERS (6/2006-5/2010), and a member of Editorial Board of *EURASIP Signal Processing Journal* (6/2005-5/2008), and an editor (7/2003-12/2005) as well as a Guest Editor (2006) of *EURASIP Journal on Applied Signal Processing*. He was a member of IEEE Signal Processing Committee on Signal Processing Theory and Methods (2005-2010). Currently, he is a member of IEEE Signal Processing Committee on Signal Processing for Communications and Networking.

Pak-Chung Ching (F'11) received the B. Eng. (1st Class Honors) and Ph.D. degrees from the University of Liverpool, UK, in 1977 and 1981 respectively. From 1981 to 1982 he was Research Officer at the University of Bath, UK. In 1982, Prof. Ching returned to Hong Kong and joined the Hong Kong Polytechnic University as a lecturer. Since 1984, he has been with the Department of Electronic Engineering of the Chinese University of Hong Kong (CUHK), where he is currently Professor of Electronic Engineering. He was Department Chairman from 1995 to 1997, Dean of Engineering from 1997 to 2003 and Head of Shaw College form 2004 to 2008. He became Director of the Shun Hing Institute of Advanced Engineering in 2004. Beginning August 1 2006, Prof Ching assumes his new responsibility as Pro-Vice-Chancellor of CUHK. His research interests include Adaptive digital signal processing, time delay estimation and target localization, blind signal estimation and separation, automatic speech recognition, speaker identification/verification and speech synthesis, and advanced signal processing techniques for wireless communications. Prof. Ching is very active in promoting professional activities, both in Hong Kong and overseas. He was a Council Member of the Institution of Electrical Engineers (IEE), past chairman of the IEEE Hong Kong Section, an associate editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1997 to 2000 and IEEE SIGNAL PROCESSING LETTERS from 2001 to 2003. He was also a member of the Technical Committee of the IEEE Signal Processing Society from 1996 to 2004. He was appointed Editor-in-Chief of the *HKIE Transactions* between 2001 and 2004. He is also an Honorary Member of the editorial committee for the *Journal of Data Acquisition and Processing*. Prof. Ching has been instrumental in organizing many international conferences in Hong Kong including the 1997 IEEE International Symposium on Circuits and Systems where he was the Vice-Chairman, and the IEEE International Conference on Acoustics, Speech and Signal Processing where he was the Technical Program Chair. He is a Fellow of IEEE, IEE, HKIE and HKAES. Prof. Ching was awarded the IEEE Third Millennium Medal in 2000 and the HKIE Hall of Fame in 2010. Prof. Ching is also a Guest Professor at the Beijing University of Posts & Telecommunications, the Huazhong University of Science & Technology, Southeast University and Institute of Acoustics, Chinese Academy of Sciences.