Differential Space-Time Modulation Over Frequency-Selective Channels

Hongbin Li, *Member, IEEE*

***Abstract*—We present herein a new *differential space-time-frequency* (DSTF) modulation scheme for systems that are equipped with an arbitrary number of transmit antennas and operate in frequency-selective channels. The proposed DSTF modulator consists of a concatenating spectral encoder and differential encoder that offer full spatio-spectral diversity and significant coding gain. A unitary structure is imposed on the differential encoder to admit linear, decoupled maximum likelihood (ML) detection in space and time. Optimum criteria based on pairwise error probability analysis are developed for spectral encoder design. We introduce a class of spectral codes, namely, *linear constellation decimation* (LCD) codes, which are nonbinary block codes obtained by decimating a phase-shift-keying (PSK) constellation with a group of decimation factors that are co-prime with the constellation size. Since LCD codes encode across a minimally necessary set of subchannels for full diversity, they incur modest decoding complexity among all full-diversity codes. Numerical results are presented to illustrate the performance of the proposed DSTF modulation and coding scheme, which compares favorably with several existing differential space-time schemes in frequency-selective channels.**

***Index Terms*—Differential modulation, frequency-selective fading, linear constellation decimation codes, maximum likelihood detection, maximum spatio-spectral diversity, OFDM, space-time coding.**

I. Introduction

MULTIANTENNA assisted space-time (ST) coding, which offers diversity gain over single-antenna systems and coding gain over uncoded systems, has generated widespread interest in recent years [1], [2]. Coherent decoding of ST codes requires reliable estimation of the underlying multichannels at the receiver. This is a challenging and costly task, especially when the channel experiences high mobility induced fast-channel fading [3], [4]. Noncoherent or differential ST coding, which circumvents the need for channel estimation, is an attractive alternative in such environments.

A number of differential ST coding schemes relying on either (complex) orthogonal design [5]–[7] or group design [8]–[12] have been proposed. They offer full spatial diversity with various tradeoffs between transmission rate and encoding/decoding complexity. It is noted that the above differential ST coding schemes were designed primarily for narrowband (flat-fading) channels. To deal with frequency-selective fading in wideband channels, these narrowband schemes can be used along with some equalization techniques to eliminate the multipath-induced inter-symbol interference (ISI) or with orthogonal frequency division multiplexing (OFDM) to convert the frequency-selective channel into a set of frequency-flat subchannels [3]. Whenever there is a complexity constraint, frequency-domain-based schemes, such as OFDM, are usually more attractive and preferred to the computationally more involved equalization-based approaches. While it is well known that OFDM, in general, need to be used with some error control coding to seek spectral diversity against frequency-selective fading, it has been a relatively new subject as to how to encode across space, time, and frequency to ensure the jointly maximum spatio-spectral diversity in a differential ST system that is subject to frequency-selective fading.

For OFDM-based systems, redundant linear precoding (LP) operating in the *complex field*, which leads to the so-called LP-OFDM, provides an effective way to achieve full spectral diversity [13]. It is shown that LP-OFDM compares favorably (in terms of bit error rate and spectral efficiency) to coded OFDM with standard error-control codes. LP-OFDM with multiantenna transmission and full spatio-spectral diversity has been discussed in [14]. An alternative scheme reported in [15] utilizes constellation rotation (CR) [16] with OFDM. In fact, CR can be thought of as a special form of LP with a rotational precoding matrix. These schemes yield full diversity only when the maximum likelihood (ML) detector with exponential complexity (in block size) is used for decoding.

While the above LP- and CR-based schemes are specially tailored for coherent systems, the research on differential space-time modulation with full spatio-spectral diversity appears more scarce. For single-antenna systems, block differentially encoded OFDM with full spectral diversity was recently studied (e.g., [17] and [18]). Diggavi *et al.* examined both equalization and, respectively, OFDM-based differential space-time coding [3]; for the latter, however, they did not address how to encode across the subcarriers, and therefore, full spectral diversity is not guaranteed.

In this paper, we extend and provide a general treatment of the *differential space-time-frequency* (DSTF) modulation scheme that was recently introduced in [19] and [20] for systems equipped with an *arbitrary* number of transmit antennas. The proposed DSTF scheme guarantees the full spatial diversity provided by multiantenna transmission/reception, as well as the maximum spectral diversity offered by multipath propagation induced frequency-selective fading. It utilizes a

Manuscript received June 13, 2003; revised July 23, 2004. This work was supported in part by the Army Research Office under Grant DAAD19-03-1-0184, and by the New Jersey Commission on Science and Technology. Part of this work was presented at the 40th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, Oct. 2002. The associate editor coordinating the review of this paper and approving it for publication was Dr. Alexi Gorokhov.

The author is with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: hli@stevens.edu).

Digital Object Identifier 10.1109/TSP.2005.847853

unitary differential encoder, which decouples the ML detection in space and time and a novel *linear constellation decimation* (LCD) encoder for spectral encoding, which incurs modest decoding complexity among all full-diversity codes. An outline of this paper is as follows.

Section II provides an overview of the proposed DSTF system. The concatenating structure of the spectral encoder and differential encoder is highlighted. The differential encoder is first detailed in Section III. It is obtained through *block orthogonal designs*, which are extensions of the classical orthogonal designs [21], which were recently successfully used for ST block coding [2]. The differential encoder imposes a unitary structure on the transmitted code matrices. We discuss how to exploit the structure to render decoupled ML detection in space and time.

In Section IV, we derive the performance criteria for code construction based on pairwise error probability (PEP) analysis. We show that for an Lth-order frequency-selective multipath channel with rich scattering, the proposed DSTF scheme can produce a diversity order of $N_t(L+1)$ with N_t transmit antennas and $N_r = 1$ receive antenna; if $N_r > 1$ receive antennas are in place, the diversity order can be further increased N_r-fold. Optimum design criteria that maximize the coding gain, in general, require encoding across all subchannels, which would result in codes with prohibitive decoding complexity. To overcome the decoding difficulty, we introduce in Section V a class of short codes, which are referred to as the *linear constellation decimation (LCD) codes*, which encode across a minimum set of subchannels that are necessary for full diversity. As such, LCD codes are shortest and incur modest decoding complexity among all full-diversity codes. We present optimum LCD codes with the largest coding gain obtained through computer searches. In Section VI, numerical results are presented to illustrate the performance of the proposed DSTF modulation and coding scheme. We show that it compares favorably with several existing differential ST schemes in wideband frequency-selective channels.

During the review process of this paper, we learned of an independent work by Ma *et al.* [22], which is similar to our approach in that both are OFDM-based techniques and guarantee full spatio-spectral diversity. A notable difference is that in their approach, independent subcarriers of OFDM are treated as virtual antennas, and therefore, diagonal space-time codes of [9] are used to provide joint spatial and spectral coding. On the other hand, our scheme exploits a concatenating structure that separates spectral and spatial coding. As such, our scheme may enjoy advantages in both error rate performance and decoding complexity, since it utilizes a smaller constellation size than that of [22] for the same number of transmit antennas and spectral efficiency.

Notation: Vectors (matrices) are denoted by boldface lower (upper) case letters; all vectors are column vectors; superscripts $(\cdot)^T, (\cdot)^*, (\cdot)^H$ denote the transpose, conjugate, and conjugate transpose, respectively; $E\{\cdot\}$ takes the statistical expectation; $\mathbf{I}_M$ denotes the $M \times M$ identity matrix; $\mathbf{0}$ and $\mathbf{1}$ denote an all-zero and, respectively, an all-one vector/matrix; $\mathrm{diag}\{\cdot\}$ denotes a diagonal matrix; $[\mathbf{X}]_{i,j}$ denotes the ijth element of matrix $\mathbf{X}$; $\|\cdot\|$ denotes the vector 2-norm; $\mathrm{tr}\{\cdot\}$ denotes the matrix trace operator; $\det\{\cdot\}$ takes the matrix determinant; and finally, $\otimes$ denotes the matrix/vector Kronecker product [23].

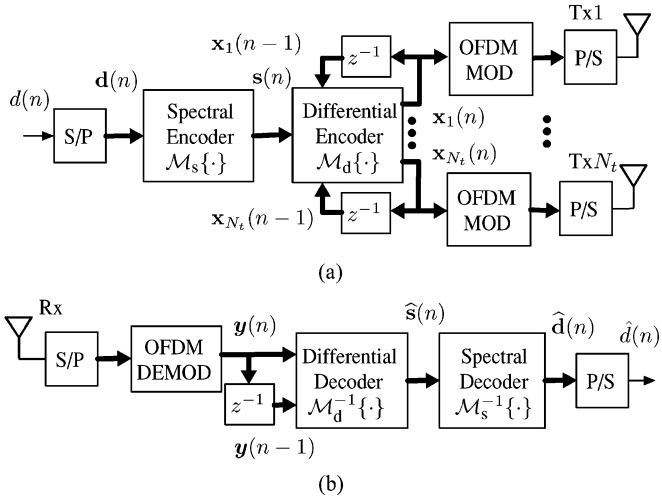

Fig. 1. Baseband differential DSTF system with N_t transmit antennas and one receive antenna. (a) Transmitter. (b) Receiver.

II. System Description

Fig. 1 depicts the block diagram of a baseband equivalent DSTF system with N_t ($N_t \geq 2$) transmit antennas and $N_r = 1$ receive antenna. The proposed techniques extend trivially to the case involving multiple receive antennas. The DSTF transmitter employs a concatenation of a *spectral encoder* $\mathcal{M}_s\{\cdot\}$ and a *differential encoder* $\mathcal{M}_d\{\cdot\}$, which encode across space, time and frequency to provide the maximum spatio-spectral diversity and coding gain.

At the transmitter, information stream $\{d(0), d(1), \ldots\}$ is first serial-to-parallel (S/P) converted to blocks of length P : $\mathbf{d}(n) \triangleq [d(nP), \ldots, d(nP+P-1)]^T$. The spectral encoder $\mathcal{M}_s\{\cdot\}$ maps $\mathbf{d}(n)$ to $P \times 1$ vectors $\mathbf{s}(n)$. As will become clear in Section V-C, the coded symbols of $\mathbf{s}(n)$ are in general drawn from a constellation with a larger size than that of $d(n)$ to provide the necessary redundancy. The differential encoder $\mathcal{M}_d\{\cdot\}$ takes as input N_s consecutive spectrally encoded vectors $\mathbf{s}(nN_s), \ldots, \mathbf{s}(nN_s+N_s-1)$ and outputs the following $N_tP \times N_d$ DSTF code matrix:[1]

$$\boldsymbol{\mathcal{X}}(n) \triangleq \begin{bmatrix} \mathbf{x}_1(nN_d) & \cdots & \mathbf{x}_1(nN_d+N_d-1) \\ \vdots & \vdots & \vdots \\ \mathbf{x}_{N_t}(nN_d) & \cdots & \mathbf{x}_{N_t}(nN_d+N_d-1) \end{bmatrix}. \quad (1)$$

Let $x_i(nN_d+l;p)$ be the pth element of the $P \times 1$ vector $\mathbf{x}_i(nN_d+l)$, for $i = 1, \ldots, N_t$ and $l = 0, \ldots, N_d-1$. This vector is next OFDM modulated by the inverse fast Fourier transform (IFFT), parallel-to-serial (P/S) converted, and then transmitted from the ith transmit antenna during the lth OFDM symbol interval. Clearly, $x_i(nN_d+l;p)$ modulates the pth subcarrier of the OFDM symbol $\mathbf{x}_i(nN_d+l)$, and each DSTF code matrix is transmitted using N_d OFDM symbol intervals.

[1]The choice of the parameters N_d, N_s, and N_t, as well as the encoder $\mathcal{M}_d\{\cdot\}$, will be discussed in Section III.

At the receiver, the received signal is S/P converted and OFDM demodulated by the fast Fourier transform (FFT), yielding $P \times 1$ vectors $\mathbf{y}(n) \triangleq [y(n;0), \ldots, y(n;P-1)]^T$, where $y(n;p)$ denotes the sample corresponding to the pth subcarrier of the nth OFDM symbol. The differential decoder $\mathcal{M}_{\rm d}^{-1}\{\cdot\}$ takes as inputs $2N_d$ data vectors $\mathbf{y}((n-1)N_d), \ldots, \mathbf{y}(nN_d + N_d - 1)$, i.e., the received data over the transmission of two adjacent DSTF code matrices $\boldsymbol{\mathcal{X}}(n-1)$ and $\boldsymbol{\mathcal{X}}(n)$, performs differential decoding, and outputs estimates $\hat{\mathbf{s}}(nN_s), \ldots, \hat{\mathbf{s}}(nN_s + N_s - 1)$. Finally, the spectral decoder $\mathcal{M}_{\rm s}^{-1}\{\cdot\}$ performs decoding and outputs estimates of the information blocks $\hat{\mathbf{d}}(n)$.

The baseband equivalent channel between the ith transmit antenna and the receive antenna is modeled as a finite impulse response (FIR) filter with coefficients $\{h_i(l)\}_{l=0}^{L}$ (e.g., [24, Sec. 14.5]), where L denotes the channel order. Due to OFDM modulation, the frequency-selective channel is equivalent to P parallel frequency-flat subchannels, with the frequency response for the pth subchannel given by $H_i(p) = \sum_{l=0}^{L} h_i(l) \exp(-j2\pi lp/P)$. The received signal can be expressed as

$$y(n;p) = \sum_{i=1}^{N_t} H_i(p) x_i(n;p) + w(n;p) \qquad p = 0, 1, \ldots, P-1 \quad (2)$$

where $w(n;p)$ denotes the zero-mean complex white Gaussian noise with variance $N_0/2$ per dimension.

The problem is to design the differential encoder $\mathcal{M}_{\rm d}\{\cdot\}$ and the spectral encoder $\mathcal{M}_{\rm s}\{\cdot\}$ for differential space-time transmission over wideband fading channels. Of primary interest is to seek the maximum spatio-spectral diversity as well as coding gain without incurring prohibitive decoding complexity.

Before we delve into the technical details, the proposed concatenating structure merits additional discussions. To ensure joint spatio-spectral diversity, it is necessary to introduce correlation across space, time, and frequency in the transmitted signal. While one could consider using a *single* space-time-frequency encoder to provide the needed correlation, design of such an encoder is, in general, intricately complicated, and so are the resulting encoding and decoding schemes (see [25] for a similar dilemma involving coherent space-time modulation). A remarkable advantage offered by concatenating two independent encoders (viz., $\mathcal{M}_{\rm d}\{\cdot\}$ and $\mathcal{M}_{\rm s}\{\cdot\}$) is that the design task is greatly simplified, and so is the encoding and decoding complexity. In particular, we will show that a nonbinary block code encoder can be used to encode *across frequency* (different subcarriers) to provide full spectral diversity and coding gain, and a unitary differential encoder can be used to encode *across space and time* at each subcarrier to provide maximum spatial diversity. The corresponding differential space-time decoding is linear (see Theorem 1), whereas the complexity of spectral decoding is controlled by utilizing spectral codes with minimum code length that is necessary to provide full spectral diversity (see Section V-C).

III. Differential Modulation

The proposed differential scheme makes use of *block orthogonal designs*. Classical orthogonal designs [21] have been recently exploited to construct space-time block codes for narrowband systems (e.g., [2] and [26]). Block extensions of orthogonal designs have been discussed in [27] for coherent communications in frequency-selective channels. Our interest here is to utilize them for differential space-time transmission. We consider complex block orthogonal designs in our case.

A. Square and Nonsquare Complex Orthogonal Designs

To introduce necessary notation, we briefly review complex orthogonal designs [2]. A *(generalized) complex orthogonal design* of size N_t in variables $s_0, \ldots, s_{N_s-1}$ is an $N_t \times N_d$ matrix $\mathbf{C}$, formed by entries $0, \pm s_0, \pm s_0^*, \ldots, \pm s_{N_s-1}, \pm s_{N_s-1}^*$ and their linear combinations, that satisfies $\mathbf{C}\mathbf{C}^H = \alpha(|s_0|^2 + \cdots + |s_{N_s-1}|^2)\mathbf{I}_{N_t}$ for some positive constant α. A complex orthogonal design $\mathbf{C}$ can be represented as [2]

$$\mathbf{C} = \sum_{k=0}^{N_s-1} [\mathbf{A}_k s_k + \mathbf{B}_k s_k^*] \quad (3)$$

where $\mathbf{A}_k$ and $\mathbf{B}_k$ are $N_t \times N_d$ real matrices that satisfy [2]:

$$\mathbf{A}_{k_1}\mathbf{A}_{k_2}^T + \mathbf{B}_{k_2}\mathbf{B}_{k_1}^T = \delta(k_1 - k_2)\mathbf{I}_{N_t} \quad (4)$$

$$\mathbf{A}_{k_1}\mathbf{B}_{k_2}^T + \mathbf{A}_{k_2}\mathbf{B}_{k_1}^T = \mathbf{0}_{N_t \times N_t}, \quad k_1 \neq k_2 \quad (5)$$

$$\mathbf{A}_k\mathbf{B}_k^T = \mathbf{0}_{N_t \times N_t} \quad (6)$$

where $\delta(k_1 - k_2)$ denotes the Kronecker delta.

To facilitate differential modulation, we classify complex orthogonal designs into two categories: *square* and *nonsquare* complex orthogonal designs. Square designs are those with $N_t = N_d$, which exist only for powers of two, i.e., $N_t = 2, 4, 8, \ldots$, and which also form the base of nonsquare designs with $N_d \neq N_t$ (and, necessarily, $N_d > N_t$ [2]). Specifically, any nonsquare complex orthogonal design can be formed by the first N_t rows of the corresponding base square design [2]. For example, a nonsquare design for $N_t = 7$ is formed by the first seven rows of the base square design with $N_t = 8$. The ratio

$$R_{\rm d} \triangleq N_s/N_d.$$

is called the *rate* of the orthogonal design. It is well known that full-rate (i.e., $R_{\rm d} = 1$) complex orthogonal designs exist only for $N_t = 2$. The best rate known for $N_t = 3$ and 4 is $R_{\rm d} = 3/4$, whereas for $N_t > 4$, it is $R_{\rm d} \leq 1/2$ [2], [26], [28]. For easy reference, we provide below the square designs for $N_t = 2$ and 4 transmit antennas, which are used in Section VI:

- $N_t = 2$ transmit antennas [2], [29]:

$$N_d = N_s = 2, \quad \text{and} \quad R_{\rm d} = 1$$
$$\mathbf{A}_0^{(2)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{A}_1^{(2)} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$
$$\mathbf{B}_0^{(2)} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{B}_1^{(2)} = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}. \quad (7)$$

- $N_t = 4$ transmit antennas [26]:

$$N_d = 4, N_s = 3, \quad \text{and} \quad R_{\rm d} = 3/4$$

$$\mathbf{A}_0^{(4)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{A}_1^{(4)} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

$$\mathbf{A}_2^{(4)} = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{B}_0^{(4)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{B}_1^{(4)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{B}_2^{(4)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}. \quad (8)$$

Other square designs can be found in [28].

B. Proposed Differential Encoder

Back to the DSTF system, the differential encoder $\mathcal{M}_{\rm d}\{\cdot\}$ in Fig. 1(a) takes as input N_s consecutive vectors: $\mathbf{s}(nN_s), \ldots, \mathbf{s}(nN_s + N_s - 1)$. Let $\boldsymbol{\mathcal{S}}(n)$ be an $\bar{N}_t P \times N_d$ matrix formed from these vectors through the following *block complex orthogonal design* [cf. (3)]:

$$\boldsymbol{\mathcal{S}}(n) \triangleq \frac{1}{\sqrt{N_s}} \sum_{k=0}^{N_s-1} [\mathbf{A}_k \otimes \mathbf{s}(nN_s + k) + \mathbf{B}_k \otimes \mathbf{s}^*(nN_s + k)] \quad (9)$$

where $\mathbf{A}_k$ and $\mathbf{B}_k$ are $\bar{N}_t \times N_d$ matrices associated with a complex orthogonal design of size $\bar{N}_t$, which is identical to the size of a base square complex orthogonal design:

$$\bar{N}_t = \begin{cases} N_t, & \text{if } N_t = 2, 4, 8, \ldots \\ 4, & \text{if } N_t = 3 \\ 8, & \text{if } N_t = 5, 6, \text{ and } 7 \\ \cdots. \end{cases} \quad (10)$$

We reiterate that a square design implies $N_d = \bar{N}_t$. Let the first transmitted DSTF code matrix be

$$\bar{\boldsymbol{\mathcal{X}}}(-1) = \sqrt{E_s}\mathbf{I}_{\bar{N}_t} \otimes \mathbf{1}_{P\times 1} \quad (11)$$

where E_s denotes the total energy emitted from all transmit antennas per subcarrier, and [cf. (1)]

$$\bar{\boldsymbol{\mathcal{X}}}(n) \triangleq \begin{bmatrix} \mathbf{x}_1(nN_d) & \cdots & \mathbf{x}_1(nN_d + N_d - 1) \\ \vdots & \vdots & \vdots \\ \mathbf{x}_{\bar{N}_t}(nN_d) & \cdots & \mathbf{x}_{\bar{N}_t}(nN_d + N_d - 1) \end{bmatrix}_{\bar{N}_t P \times N_d}.$$

The proposed differential encoding scheme proceeds *as if there were $\bar{N}_t$ transmit antennas*:

$$\bar{\boldsymbol{\mathcal{X}}}(n) = \bar{\boldsymbol{\mathcal{D}}}_x(n-1)\boldsymbol{\mathcal{S}}(n), \quad n = 0, 1, \ldots. \quad (12)$$

The matrix sequence $\bar{\boldsymbol{\mathcal{D}}}_x(n)$ is defined as

$$\bar{\boldsymbol{\mathcal{D}}}_x(n) \triangleq \begin{bmatrix} \mathbf{D}_{x_1}(nN_d) & \cdots & \mathbf{D}_{x_1}(nN_d + N_d - 1) \\ \vdots & \vdots & \vdots \\ \mathbf{D}_{x_{\bar{N}_t}}(nN_d) & \cdots & \mathbf{D}_{x_{\bar{N}_t}}(nN_d + N_d - 1) \end{bmatrix} \quad (13)$$

where $\mathbf{D}_{x_i}(n) \triangleq \text{diag}\{\mathbf{x}_i(n)\}$. Throughout the paper, we assume that the coded symbols $\mathbf{s}(n)$ are drawn from a constant-modulus constellation $\mathcal{A}_s$ (e.g., PSK) with unit-energy elements. This assumption, along with the orthogonal design (9), suggests that $\bar{\boldsymbol{\mathcal{D}}}_x(n)$ is a (scaled) unitary matrix with

$$\bar{\boldsymbol{\mathcal{D}}}_x(n)\bar{\boldsymbol{\mathcal{D}}}_x^H(n) = E_s\mathbf{I}_{N_d P}. \quad (14)$$

Since we have N_t rather than $\bar{N}_t$ antennas, we cannot proceed to transmit $\bar{\boldsymbol{\mathcal{X}}}(n)$. Instead, the following matrix is transmitted [cf. (1)]:

$$\boldsymbol{\mathcal{X}}(n) = \mathbf{T}\bar{\boldsymbol{\mathcal{X}}}(n) \quad (15)$$

where $\mathbf{T} \triangleq [\mathbf{I}_{N_t P}, \mathbf{0}_{N_t P \times (\bar{N}_t - N_t)P}]$. That is, the last $(\bar{N}_t - N_t)P$ rows of $\bar{\boldsymbol{\mathcal{X}}}(n)$ are discarded. It is noted that due to the unique structure of $\bar{\boldsymbol{\mathcal{D}}}_x(n)$ in (13) (formed by diagonal matrices), the differential encoding (12) *is performed at each subcarrier*.

Let $\mathbf{h}_{{\rm f},i} \triangleq [H_i(0), \ldots, H_i(P-1)]^T$ and $\mathbf{D}_{h_{{\rm f},i}} \triangleq \text{diag}\{\mathbf{h}_{{\rm f},i}\}$. In vector form, the received signal can be written as [cf. (2)]

$$\mathbf{y}(nN_d + l) = \sum_{i=1}^{\bar{N}_t} \mathbf{D}_{x_i}(nN_d + l)\mathbf{h}_{{\rm f},i} + \mathbf{w}(nN_d + l) \quad l = 0, \ldots, N_d - 1 \quad (16)$$

where $\mathbf{w}(nN_d + l) \triangleq [w(nN_d + l; 0), \ldots, w(nN_d + l; P-1)]^T$, and we note that $\mathbf{h}_{{\rm f},N_t+1} = \cdots = \mathbf{h}_{{\rm f},\bar{N}_t} = \mathbf{0}$. Let $\boldsymbol{y}(n) \triangleq [\mathbf{y}^T(nN_d), \ldots, \mathbf{y}^T(nN_d + N_d - 1)]^T$. We can write (16) collectively as

$$\boldsymbol{y}(n) = \bar{\boldsymbol{\mathcal{D}}}_x^T(n)\bar{\boldsymbol{h}}_{\rm f} + \boldsymbol{w}(n) \quad (17)$$

where $\bar{\boldsymbol{h}}_{\rm f} \triangleq [\mathbf{h}_{{\rm f},1}^T, \ldots, \mathbf{h}_{{\rm f},\bar{N}_t}^T]^T$, and $\boldsymbol{w}(n) \triangleq [\mathbf{w}^T(nN_d), \ldots, \mathbf{w}^T(nN_d + N_d - 1)]^T$. An equivalent form of (12) is

$$\bar{\boldsymbol{\mathcal{D}}}_x(n) = \bar{\boldsymbol{\mathcal{D}}}_x(n-1)\boldsymbol{\mathcal{D}}_s(n) \quad (18)$$

where [see (9)]

$$\boldsymbol{\mathcal{D}}_s(n) = \frac{1}{\sqrt{N_s}} \sum_{k=0}^{N_s-1} [\mathbf{A}_k \otimes \mathbf{D}_s(nN_s + k) + \mathbf{B}_k \otimes \mathbf{D}_s^*(nN_s + k)] \quad (19)$$

with $\mathbf{D}_s(n) \triangleq \text{diag}\{\mathbf{s}(n)\}$. Note that $\boldsymbol{\mathcal{D}}_s(n)$ is unitary by construction. Substituting (18) into (17) yields

$$\begin{aligned} \boldsymbol{y}(n) &= \boldsymbol{\mathcal{D}}_s^T(n)\bar{\boldsymbol{\mathcal{D}}}_x^T(n-1)\bar{\boldsymbol{h}}_{\rm f} + \boldsymbol{w}(n) \\ &= \boldsymbol{\mathcal{D}}_s^T(n)\boldsymbol{y}(n-1) + \boldsymbol{v}(n) \end{aligned} \quad (20)$$

where $\boldsymbol{v}(n) \triangleq \boldsymbol{w}(n) - \boldsymbol{\mathcal{D}}_s^T(n)\boldsymbol{w}(n-1)$. Due to the unitary transformation $\boldsymbol{\mathcal{D}}_s^T(n)$, $\boldsymbol{v}(n)$ consists of independently and identically distributed (i.i.d.) complex Gaussian entries with zero-mean and variance N_0 per dimension. Substituting (19) into (20) followed by simplification gives

$$\boldsymbol{y}(n) = \frac{1}{\sqrt{N_s}} \left\{ \sum_{k=0}^{N_s-1} \sum_{l=0}^{N_d-1} [\mathbf{a}_{k,l} \otimes \mathbf{D}_y((n-1)N_d + l)] \times \mathbf{s}(nN_s + k) + [\mathbf{b}_{k,l} \otimes \mathbf{D}_y((n-1)N_d + l)] \times \mathbf{s}^*(nN_s + k) \right\} + \boldsymbol{v}(n) \quad (21)$$

where $\mathbf{a}_{k,l}$ and $\mathbf{b}_{k,l}$ are $N_d \times 1$ vectors formed from the lth row of $\mathbf{A}_k$ and $\mathbf{B}_k$, respectively, and $\mathbf{D}_y((n-1)N_d+l) \triangleq \mathrm{diag}\{\mathbf{y}((n-1)N_d+l)\}$. Equation (21) is referred to as the *fundamental differential receiver equation*. Remarkably, the unknown channel vectors $\mathbf{h}_{\mathrm{f},i}$ are absent from this equation. The variance of the virtual noise $\boldsymbol{v}(n)$ is twice that of the real channel noise $\mathbf{w}(n)$. This translates to a 3-dB loss of signal-to-noise ratio (SNR), which is well-known for standard single-antenna differential modulation.

Due to the unitary structure of the proposed differential encoder, the ML detection of the space-time multiplexed vectors $\{\mathbf{s}(nN_s+k)\}_{k=0}^{N_s-1}$ is decoupled. In particular, we have the following result.

Theorem 1: The ML detection of the N_s coded vectors $\{\mathbf{s}(nN_s+k)\}_{k=0}^{N_s-1}$ based on two adjacent vectors $\boldsymbol{y}(n-1)$ and $\boldsymbol{y}(n)$ decouples into N_s individual detections:

$$\hat{\mathbf{s}}^{\mathrm{ML}}(nN_s+k) = \arg\max_{\mathbf{s}_k \in \boldsymbol{\mathcal{B}}_s} \Re\left\{\mathbf{z}^H(nN_s+k)\boldsymbol{\Omega}_y^{1/2}(n-1)\mathbf{s}_k\right\} \quad k=0,\ldots,N_s-1 \tag{22}$$

where $\boldsymbol{\mathcal{B}}_s \subseteq \mathcal{A}_s^{P\times 1}$ denotes a valid codebook (of the spectral encoder $\mathcal{M}_{\mathrm{s}}\{\cdot\}$), and

$$\mathbf{z}(nN_s+k) \triangleq \boldsymbol{\Omega}_y^{-1/2}(n-1)\sum_{l=0}^{N_d-1}\left\{\left[\mathbf{a}_{k,l}^T \otimes \mathbf{D}_y^H((n-1)N_d+l)\right] \times \boldsymbol{y}(n) + \left[\mathbf{b}_{k,l}^T \otimes \mathbf{D}_y((n-1)N_d+l)\right]\boldsymbol{y}^*(n)\right\} \tag{23}$$

$$\boldsymbol{\Omega}_y(n-1) \triangleq \sum_{l=0}^{N_d-1}\mathbf{D}_y((n-1)N_d+l)\mathbf{D}_y^H((n-1)N_d+l). \tag{24}$$

Proof: See Appendix A. ∎

Hence, the ML detector amounts to a linear filtering (in space and time), as described in (23), followed by spectral decoding in (22). The complexity can be further simplified a little bit by noticing that the premultiplication by $\boldsymbol{\Omega}_y^{-1/2}(n-1)$ in (23) is not needed since it cancels the $\boldsymbol{\Omega}_y^{1/2}(n-1)$ in (22). Next, we discuss how to design the spectral encoder $\mathcal{M}_{\mathrm{s}}\{\cdot\}$.

IV. Performance Criteria

We consider the PEP, which is defined as the probability of the event that the ML detector erroneously chooses codeword $\mathbf{s}'(nN_s+k)$ as $\mathbf{s}(nN_s+k)$ for code design. The following assumption is made through our analysis.

A1) *(Correlated) Rayleigh Fading:* The channel vectors $\mathbf{h}_i \triangleq [h_i(0), h_i(1), \ldots, h_i(L)]^T, i=1,\ldots,N_t$ are zero-mean complex Gaussian random vectors with nonsingular covariance matrix $\mathbf{R}_h \triangleq E\{\mathbf{h}\mathbf{h}^H\}$, where $\mathbf{h} \triangleq [\mathbf{h}_1^T,\ldots,\mathbf{h}_{N_t}^T]^T$.

Extensions to other channel models can be made in a manner as in [1]. By (57) and Theorem 1, the conditional PEP at high SNR can be approximated by (dropping the indices nN_s+k and $n-1$ for notational brevity) [24]:

$$\mathrm{P}(\mathbf{s}\to\mathbf{s}' \,|\, \boldsymbol{\Omega}_y) = \exp\left(-\frac{d^2(\mathbf{s},\mathbf{s}')E_s}{8N_0}\right) \tag{25}$$

where $d^2(\mathbf{s},\mathbf{s}') \triangleq (N_sE_s)^{-1}\mathbf{e}^H\boldsymbol{\Omega}_y\mathbf{e}$, and $\mathbf{e} \triangleq \mathbf{s}-\mathbf{s}'$. Using (16) in (24) and ignoring the noise at high SNR, we have

$$\boldsymbol{\Omega}_y \approx \bar{\mathbf{D}}_{h_{\mathrm{f}}}\bar{\boldsymbol{\mathcal{D}}}_x(n)\bar{\boldsymbol{\mathcal{D}}}_x^H(n)\bar{\mathbf{D}}_{h_{\mathrm{f}}}^H = E_s\bar{\mathbf{D}}_{h_{\mathrm{f}}}\bar{\mathbf{D}}_{h_{\mathrm{f}}}^H = E_s\sum_{i=1}^{N_t}\mathbf{D}_{h_{\mathrm{f},i}}\mathbf{D}_{h_{\mathrm{f},i}}^H$$

where $\bar{\boldsymbol{\mathcal{D}}}_x(n)$ is defined in (13), $\bar{\mathbf{D}}_{h_{\mathrm{f}}} \triangleq [\mathbf{D}_{h_{\mathrm{f},1}},\ldots,\mathbf{D}_{h_{\mathrm{f},N_t}}]$, and the second equality is due to (14). It follows that

$$\begin{aligned} d^2(\mathbf{s},\mathbf{s}') &\approx N_s^{-1}\mathbf{e}^H\left(\sum_{i=1}^{N_t}\mathbf{D}_{h_{\mathrm{f},i}}\mathbf{D}_{h_{\mathrm{f},i}}^H\right)\mathbf{e} \\ &= N_s^{-1}\sum_{i=1}^{N_t}\mathbf{h}_{\mathrm{f},i}^H\mathbf{D}_e^H\mathbf{D}_e\mathbf{h}_{\mathrm{f},i} \\ &= N_s^{-1}\sum_{i=1}^{N_t}\mathbf{h}_i^H\boldsymbol{\mathcal{F}}^H\mathbf{D}_e^H\mathbf{D}_e\boldsymbol{\mathcal{F}}\mathbf{h}_i \triangleq \mathbf{h}^H\boldsymbol{\Phi}_e\mathbf{h} \end{aligned}$$

where $\mathbf{D}_e \triangleq \mathrm{diag}\{\mathbf{e}\}$, the third equality uses the fact that $\mathbf{h}_{\mathrm{f},i} = \boldsymbol{\mathcal{F}}\mathbf{h}_i$, with $\boldsymbol{\mathcal{F}} \in \mathbb{C}^{P\times(L+1)}$ denoting the P-point FFT matrix: $[\boldsymbol{\mathcal{F}}]_{p,q} = \exp(-\jmath 2\pi(p-1)(q-1)/P)$, and finally

$$\boldsymbol{\Phi}_e \triangleq N_s^{-1}\mathbf{I}_{N_t} \otimes (\boldsymbol{\mathcal{F}}^H\mathbf{D}_e^*\mathbf{D}_e\boldsymbol{\mathcal{F}}). \tag{26}$$

Denote the eigenvalue decomposition (EVD) of $\mathbf{R}_h$ by $\mathbf{R}_h = \mathbf{U}_h\boldsymbol{\Lambda}_h\mathbf{U}_h^H$, where $\boldsymbol{\Lambda}_h$ is diagonal formed by the eigenvalues of $\mathbf{R}_h$, and $\mathbf{U}_h$ contains the eigenvectors. Furthermore, let

$$\boldsymbol{\Psi}_e \triangleq \boldsymbol{\Lambda}_h^{1/2}\mathbf{U}_h^H\boldsymbol{\Phi}_e\mathbf{U}_h\boldsymbol{\Lambda}_h^{1/2} \tag{27}$$

and denote its EVD by $\boldsymbol{\Psi}_e = \mathbf{U}_\psi\boldsymbol{\Lambda}_\psi\mathbf{U}_\psi^H$, where $\boldsymbol{\Lambda}_\psi \triangleq \mathrm{diag}\{\lambda_1,\ldots,\lambda_{N_t(L+1)}\}$, with $\lambda_1 \geq \cdots \geq \lambda_{N_t(L+1)} \geq 0$, and the eigenvectors are contained in $\mathbf{U}_\psi$. Then

$$\begin{aligned} d^2(\mathbf{s},\mathbf{s}') &\approx \mathbf{h}^H\mathbf{U}_h\boldsymbol{\Lambda}_h^{-1/2}\boldsymbol{\Psi}_e\boldsymbol{\Lambda}_h^{-1/2}\mathbf{U}_h^H\mathbf{h} \\ &= \mathbf{h}^H\mathbf{U}_h\boldsymbol{\Lambda}_h^{-1/2}\mathbf{U}_\psi\boldsymbol{\Lambda}_\psi\mathbf{U}_\psi^H\boldsymbol{\Lambda}_h^{-1/2}\mathbf{U}_h^H\mathbf{h} \\ &\triangleq \tilde{\mathbf{h}}^H\boldsymbol{\Lambda}_\psi\tilde{\mathbf{h}} \end{aligned} \tag{28}$$

where $\tilde{\mathbf{h}} \triangleq \mathbf{U}_\psi^H\boldsymbol{\Lambda}_h^{-1/2}\mathbf{U}_h^H\mathbf{h}$. Clearly, $\tilde{\mathbf{h}}$ is complex Gaussian with zero-mean and identity covariance matrix. Substituting (28) into (25) and taking the expectation with respect to $\tilde{\mathbf{h}}$, we have [1], [13]

$$\begin{aligned} \mathrm{P}(\mathbf{s}\to\mathbf{s}') &= \prod_{m=1}^{N_t(L+1)}\frac{1}{1+\lambda_m E_s/(8N_0)} \\ &\leq \left(\frac{E_s}{8N_0}\right)^{-r_e}\left(\prod_{m=1}^{r_e}\lambda_m\right)^{-1} \end{aligned} \tag{29}$$

where $r_e \triangleq \mathrm{rank}(\boldsymbol{\Psi}_e) = \mathrm{rank}(\boldsymbol{\Phi}_e) \leq N_t(L+1)$.

Let $G_{\mathrm{d},e} \triangleq r_e$ and $G_{\mathrm{c},e} \triangleq (\prod_{m=1}^{r_e}\lambda_m)^{1/r_e}$. Following [1], $G_{\mathrm{d}} \triangleq \min_{\forall \mathbf{e}\neq\mathbf{0}} G_{\mathrm{d},e}$ and $G_{\mathrm{c}} \triangleq \min_{\forall \mathbf{e}\neq\mathbf{0}} G_{\mathrm{c},e}$ are referred to as the *diversity gain* and *coding gain*, respectively. The design criteria are as follows. [1]

- *Rank criterion*: Maximize the diversity gain $G_{\rm d}$ over all possible coding schemes.
- *Product criterion*: Maximize $G_{\rm c}$ over all possible coding schemes.

Next, we summarize the maximum diversity and coding gain of the proposed DSTF system.

Theorem 2: Under the channel condition **A1)**, the maximum diversity gain of the DSTF system equipped with N_t transmit antennas and one receive antenna is $G_{\rm d,max} = N_t(L+1)$, which is achieved if and only if the minimum Hamming distance of the code $\mathbf{s}$ is no less than $L+1$. Any maximum-diversity achieving code has a coding gain that is upper bounded by

$$G_{\rm c,max} = N_s^{-1} d_{\min}^2 [\det(\mathbf{R}_h)]^{1/[N_t(L+1)]} \tag{30}$$

where $d_{\min}$ denotes the *minimum Euclidean distance* of the code [see (33)].

Proof: First note that the minimum rank of $\boldsymbol{\mathcal{F}}^H \mathbf{D}_e^* \mathbf{D}_e \boldsymbol{\mathcal{F}}$ is one. This occurs when the minimum Hamming distance of any two different codewords $\mathbf{s}$ and $\mathbf{s}'$ is one or, equivalently, when the spectral encoder is trivial, i.e., $\mathcal{M}_{\rm s}\{\cdot\}$ in Fig. 1(a) is simply an identity. Hence, the minimum diversity of the system is N_t (due to the block diagonal structure of $\boldsymbol{\Phi}_e$), which corresponds to the spatial diversity. The maximum possible rank of $\boldsymbol{\mathcal{F}}^H \mathbf{D}_e^* \mathbf{D}_e \boldsymbol{\mathcal{F}}$ is $L+1$ since its dimension is as well. Matrix $\boldsymbol{\mathcal{F}}^H \mathbf{D}_e^* \mathbf{D}_e \boldsymbol{\mathcal{F}}$ has rank exactly equal to $L+1$ for all error events if and only if the minimum Hamming distance is no less than $L+1$. This is because $\boldsymbol{\mathcal{F}}$ is a $P \times (L+1)$ Vandermonde matrix, and thus, any $L+1$ rows of $\boldsymbol{\mathcal{F}}$ are linearly independent, provided that $P \geq L+1$. Therefore, the maximum diversity is $N_t(L+1)$, due to, again, the block diagonal structure of $\boldsymbol{\Phi}_e$. This is the maximum diversity order offered by the system.

Now, consider an arbitrary full diversity code (i.e., $\boldsymbol{\Phi}_e$ has full rank). The coding gain is maximized if and only if $\det(\boldsymbol{\Psi}_e)$ is maximized. Using (26) and (27), we have

$$\begin{aligned} \det(\boldsymbol{\Psi}_e) &= \det\left(\boldsymbol{\Lambda}_h^{1/2} \mathbf{U}_h^H \boldsymbol{\Phi}_e \mathbf{U}_h \boldsymbol{\Lambda}_h^{1/2}\right) \\ &= \det(\boldsymbol{\Phi}_e)\det(\mathbf{R}_h) \\ &= N_s^{-N_t(L+1)} [\det(\boldsymbol{\mathcal{F}}^H \mathbf{D}_e^* \mathbf{D}_e \boldsymbol{\mathcal{F}})]^{N_t} \det(\mathbf{R}_h). \end{aligned} \tag{31}$$

We next observe that $\boldsymbol{\mathcal{F}}^H \mathbf{D}_e^* \mathbf{D}_e \boldsymbol{\mathcal{F}}$ is a Hermitian and Toeplitz matrix, with identical diagonal elements $\mathrm{tr}(\mathbf{D}_e^* \mathbf{D}_e)$. By the Hadamard inequality [30 , p. 477]

$$[\det(\boldsymbol{\mathcal{F}}^H \mathbf{D}_e^* \mathbf{D}_e \boldsymbol{\mathcal{F}})]^2 \leq [\mathrm{tr}(\mathbf{D}_e^* \mathbf{D}_e)]^{2(L+1)} = d_e^{4(L+1)} \tag{32}$$

where $d_e \triangleq [\mathrm{tr}(\mathbf{D}_e^* \mathbf{D}_e)]^{1/2} = \|\mathbf{e}\|$ is the *pairwise Euclidean distance* between the codewords $\mathbf{s}$ and $\mathbf{s}'$. The equality in (32) holds if and only if $\boldsymbol{\mathcal{F}}^H \mathbf{D}_e^* \mathbf{D}_e \boldsymbol{\mathcal{F}}$ is diagonal. Note that the above bound is asymptotically (for large P) tight since $\boldsymbol{\mathcal{F}}^H \mathbf{D}_e^* \mathbf{D}_e \boldsymbol{\mathcal{F}}$ is asymptotically diagonal (e.g., [31]). It follows that the coding gain is upper bounded by (30), where the *minimum Euclidean distance* of the code is defined as[2]

$$d_{\min} \triangleq \min_{\forall \mathbf{e} \neq \mathbf{0}} d_e \tag{33}$$

and the proof is complete. ■

[2] $d_{\min}^2$ should not be confused with the minimum Euclidean distance of the constellation points of $\mathcal{A}_s$.

To summarize, optimum codes based on the rank criterion should have a minimum Hamming distance of $L+1$, which yields the maximum spatial-spectral diversity gain. Meanwhile, to achieve the maximum coding gain, $\det(\boldsymbol{\mathcal{F}}^H \mathbf{D}_e^* \mathbf{D}_e \boldsymbol{\mathcal{F}})$ has to be maximized. The upper bound (30) indicates that it is beneficial to maximize the minimum Euclidean distance of the code, provided that the constraint on the minimum Hamming distance is met. These rules can be utilized to construct optimum codes. In general, full-diversity codes with optimum coding gain require encoding across all subchannels, which leads to exponential (in P) decoding complexity. For practical applications, it is more interesting to develop *short* codes that encode across a minimum set of subchannels but still achieve full diversity and significant coding gain. In the next section, we develop such short codes.

V. Spectral Encoding: LCD Codes

In this section, we present a class of short codes, referred to as the *linear constellation decimation* (LCD) codes, that encode across a minimum set of subcarriers to ensure the full spatio-spectral diversity and significant coding gain. LCD codes have the *minimum* code length and, therefore, incur modest decoding complexity among all full-diversity codes. LCD codes are used in conjunction with a *subcarrier interleaver* that interleaves coded symbols in frequency. In the following, we first briefly review the idea of subcarrier interleaving. Next, we develop the corresponding code design criteria that incorporate subcarrier interleaving. These criteria are employed to construct LCD codes. Finally, we revisit the issue of ML detection when LCD codes are used.

A. Subcarrier Interleaving

Our notion for subcarrier interleaving follows [15]. Let $\mathcal{I} \triangleq \{0, 1, \ldots, P-1\}$ collect the indices of all subcarriers. Subcarrier interleaving is a partition of $\mathcal{I}$ into M nonoverlapping subsets $\mathcal{I}_m \triangleq \{p_{m,0}, p_{m,1}, \ldots, p_{m,J_m-1}\}$, where J_m denotes the number of subcarriers grouped in the mth subset. Subcarrier interleaving allows transmitting multiple short codewords over different subcarrier sets. Since ML detection of these codewords is independent, using short codes along with subcarrier interleaving can lead to significant computational saving compared with using long codes. For an Lth-order Rayleigh fading channel satisfying **A1)**, it is necessary to have $J_m \geq L+1$ to ensure full spectral diversity [15]. On the other hand, since the ML decoding complexity is exponential in J_m, J_m should be as small as possible. Hence, we choose $J_m = L+1$. Among various alternatives, the following subcarrier interleaving scheme is conceptually simple, by which the subcarriers within a subset are maximally separated in the frequency domain:

$$\mathcal{I}_m = \{m, \quad M+m, \ldots, \quad LM+m\} \tag{34}$$

where $M \triangleq P/(L+1)$, and P is assumed to be an integer multiple of $L+1$. For simplicity, we will hereafter assume this subcarrier interleaving scheme.

B. Design Criteria

The input–output relation for the mth subcarrier subset is given by [cf. (57)]

$$\begin{aligned}\mathbf{z}_m(nN_s+k) &= N_s^{-1/2}\boldsymbol{\Omega}_{y,m}^{1/2}(n-1)\mathbf{s}_m(nN_s+k)+\boldsymbol{\nu}_{k,m}(nN_s+k)\\ &\qquad k=0,\ldots,N_s-1;\quad m=0,\ldots,M-1 \end{aligned} \tag{35}$$

where $\mathbf{z}_m(nN_s+k)\in\mathbb{C}^{(L+1)\times 1}$, $\boldsymbol{\Omega}_{y,m}(n-1)\in\mathbb{C}^{(L+1)\times(L+1)}$, $\mathbf{s}_m(nN_s+k)\in\boldsymbol{\mathcal{B}}_{s,m}$, and $\boldsymbol{\nu}_{k,m}(nN_s+k)\in\mathbb{C}^{(L+1)\times 1}$ are quantities associated with the mth subcarrier subset, which are similarly defined as their counterparts in (57). Note that $\boldsymbol{\mathcal{B}}_{s,m}\subseteq\mathcal{A}_s^{(L+1)\times 1}$ denotes the codebook used for the mth subset. The maximum diversity and the associated coding gains of the proposed DSTF system with subcarrier interleaving are summarized in the following result.

Theorem 3: Under the channel condition **A1)**, the maximum diversity gain of the DSTF system with N_t transmit antennas, one receive antenna and subcarrier interleaving (34) is $G_{\mathrm{d,max}}^{(m)}=N_t(L+1)$, which is achieved if and only if the codebook $\boldsymbol{\mathcal{B}}_{s,m}$ has a uniform Hamming distance of $L+1$. Any maximum-diversity achieving code has a coding gain given by

$$G_{\mathrm{c,max}}^{(m)}=N_s^{-1}(L+1)\left[\delta_{\min}^{2N_t}\det(\mathbf{R}_h)\right]^{1/[N_t(L+1)]} \tag{36}$$

where $\delta_{\min}$ denotes the *minimum product distance* of the code [see (41)].

Proof: Following a PEP analysis similar to the one in Section IV, we can show that the pairwise diversity gain with subcarrier interleaving is $G_{\mathrm{d},e}^{(m)}\triangleq\mathrm{rank}(\boldsymbol{\Phi}_{e,m})$, where [cf. (26)]

$$\boldsymbol{\Phi}_{e,m}\triangleq N_s^{-1}\mathbf{I}_{N_t}\otimes\left[\boldsymbol{\mathcal{F}}_m^H\mathbf{D}_{e,m}^*\mathbf{D}_{e,m}\boldsymbol{\mathcal{F}}_m\right]. \tag{37}$$

The diagonal matrix $\mathbf{D}_{e,m}$, which is similar to $\mathbf{D}_e$ in (26), is formed by the difference of two different codewords in $\boldsymbol{\mathcal{B}}_{s,m}$, and $\boldsymbol{\mathcal{F}}_m\in\mathbb{C}^{(L+1)\times(L+1)}$ is formed from rows $m,m+M,\ldots,m+LM$ of the $P\times(L+1)$ FFT matrix $\boldsymbol{\mathcal{F}}$. A direct calculation shows that $\boldsymbol{\mathcal{F}}_m$ is orthogonal:

$$\boldsymbol{\mathcal{F}}_m^H\boldsymbol{\mathcal{F}}_m=(L+1)\mathbf{I}_{L+1}. \tag{38}$$

It follows that $\mathrm{rank}(\boldsymbol{\Phi}_{e,m})=N_t\mathrm{rank}(\mathbf{D}_{e,m}^*\mathbf{D}_{e,m})$. Hence, the maximum diversity gain is $N_t(L+1)$, which is identical to that in Theorem 2, and is achieved if and only if the code $\boldsymbol{\mathcal{B}}_{s,m}$ has a uniform Hamming distance $L+1$.

Likewise, the pairwise coding gain of any maximum-diversity achieving code with subcarrier interleaving is [cf. (31)]

$$G_{\mathrm{c},e}^{(m)}\triangleq\left[\det(\boldsymbol{\Phi}_{e,m})\det(\mathbf{R}_h)\right]^{1/[N_t(L+1)]}. \tag{39}$$

Note that

$$\begin{aligned}\det(\boldsymbol{\Phi}_{e,m}) &= N_s^{-N_t(L+1)}\left[\det(\boldsymbol{\mathcal{F}}_m^H\mathbf{D}_{e,m}^*\mathbf{D}_{e,m}\boldsymbol{\mathcal{F}}_m)\right]^{N_t}\\ &= N_s^{-N_t(L+1)}\left[\det(\mathbf{D}_{e,m}^*\mathbf{D}_{e,m})\det\left(\boldsymbol{\mathcal{F}}_m^H\boldsymbol{\mathcal{F}}_m\right)\right]^{N_t}\\ &\triangleq\left[N_s^{-1}(L+1)\right]^{N_t(L+1)}\delta_e^{2N_t}\end{aligned} \tag{40}$$

where the third equality is due to (38), and $\delta_e\triangleq\left|\det(\mathbf{D}_{e,m})\right|$ denotes the *pairwise product distance* of the code. It follows that the maximum coding gain is (36), where the *minimum product distance* of the codebook is defined as

$$\delta_{\min}=\min_{\forall\mathbf{e}_m\neq\mathbf{0}}\left|\det(\mathbf{D}_{e,m})\right| \tag{41}$$

which concludes the proof. ■

Hence, subcarrier interleaving incurs no loss of diversity. Note that a Euclidean distance upper bound similar to (30) is less meaningful here since it requires a large L (i.e., the channel order) to be tight, as opposed to the requirement for a large P (the number of subcarriers) there. Hence, for maximum coding gain, the product distance δ_e, rather than the Euclidean distance of the code, has to be maximized in the current case with subcarrier interleaving.

C. LCD Codes

Even with subcarrier interleaving, exact maximization of the product distance $\delta_{\min}$ is still highly involved; in general, there is no closed-form solution. In what follows, we introduce one way of building *linear* codes by means of *constellation decimation*. We show that these LCD codes guarantee full spatio-spectral diversity, as well as substantial coding gain compared to other heuristic codes.

The subcarrier subset index m is dropped for notational brevity. To reduce decoding complexity, we consider codes with a length of $L+1$, which is the minimum code length to achieve the maximum diversity. To achieve a code rate of R_{s} bits per coded symbol using a set of $L+1$ interleaved subcarriers, we need a codebook with $N_c\triangleq 2^{R_{\mathrm{s}}(L+1)}$ distinct codewords (of length $L+1$), drawn from an M_c-ary PSK constellation $\mathcal{A}_s$. The need for a PSK constellation is due to the unitary differential mapping described in Section III. Let $\mathbf{s}_i\triangleq[s_{i,0},\ldots,s_{i,L}]^T\in\mathcal{A}_s^{(L+1)\times 1}, i=0,\ldots,N_c-1$ denote the ith codeword. We first present a result about the choice of the constellation size M_c.

Proposition 1: A necessary condition for achieving the maximum diversity gain is $M_c\geq N_c$.

Proof: Suppose $M_c<N_c$. Then, there exist $\mathbf{s}_i$ and $\mathbf{s}_j$ for $i\neq j$ with $s_{i,k}=s_{j,k}$, which implies that the minimum Hamming distance of the code is less than $L+1$. According to Theorem 3, the code cannot achieve the maximum diversity gain. ■

Note that for codes with $M_c<N_c$, the minimum product distance is zero since different codewords must share at least one identical coded symbol at one place. We choose the minimum constellation size $M_c=N_c$ to seek the minimum decoding complexity. This choice, however, has the following implication.

Proposition 2: If the minimum constellation size $M_c=N_c$ is adopted, then all symbols in $\mathcal{A}_s$ must appear in each row of the codebook $\boldsymbol{\mathcal{B}}_s\triangleq[\mathbf{s}_0,\ldots,\mathbf{s}_{N_c-1}]_{(L+1)\times N_c}$ once and only once, in order to achieve the maximum diversity gain.

Proof: First, if any symbol in $\mathcal{A}_s$ appears in a row of $\boldsymbol{\mathcal{B}}_s$ for more than once, the minimum Hamming distance of the resulting code is less than $L+1$, which loses the maximum di-

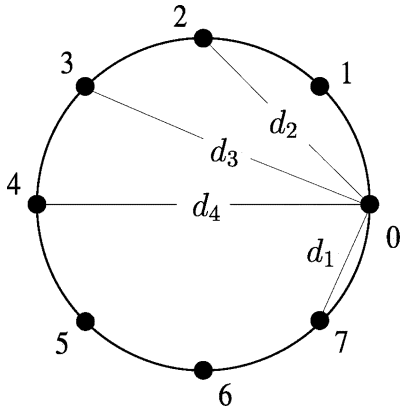

Fig. 2. The 8-PSK constellation.

versity gain. On the other hand, the fact that any symbol in $\mathcal{A}_s$ cannot appear in a row of $\boldsymbol{\mathcal{B}}_s$ more than once implies that each symbol must appear at least once, in order to have a codebook of $N_c = M_c$ distinct codewords. ■

If we label the constellation points in $\mathcal{A}_s$ as $0, 1, \ldots, M_c - 1$ (see, e.g., the 8-PSK constellation shown in Fig. 2) and form the sequence $\vec{c} \triangleq [0, 1, \ldots, M_c - 1]$, Proposition 2 mandates that each row of $\boldsymbol{\mathcal{B}}_s$ be a permutation of $\vec{c}$. It might be asked if any permutation may lead to loss in diversity. The answer is no.

Theorem 4: Assuming the minimum constellation size $M_c = N_c$, any code $\boldsymbol{\mathcal{B}}_s$ with rows formed by any permutation of the sequence $\vec{c}$ attains the maximum diversity gain; conversely, any maximum-diversity achieving code $\boldsymbol{\mathcal{B}}_s$ has rows that are permutations of $\vec{c}$.

Proof: The jth row $j = 0, 1, \ldots, L$ of codebook $\boldsymbol{\mathcal{B}}_s$, being a permutation of $\vec{c}$, indicates that all N_c codewords in $\boldsymbol{\mathcal{B}}_s$ differ at the jth coded symbol, which increases the Hamming distance by one. Since there are a total of $L + 1$ rows, $\boldsymbol{\mathcal{B}}_s$ has a uniform Hamming distance, thus achieving the maximum diversity order. The converse statement follows directly from Proposition 2. ■

Following Theorem 4, a simple calculation indicates that there are a total of $(N_c!)^L$ distinct codes of code length $L + 1$, all achieving the maximum diversity. An exhaustive search for codes with the largest product distance quickly becomes infeasible, even for relatively small N_c and L. To facilitate code construction, we introduce the idea of *constellation decimation*, which effectively imposes a linear structure on the code. The linear structure makes the analysis of distance property and search for good codes significantly easier.

Definition 1 (Constellation Decimation): Let $\vec{c} \triangleq \{0, 1, \ldots, M_c - 1\}$, and $\vec{c}[k]$ denote the kth element of $\vec{c}$. Denote by $\vec{c}_q \triangleq \{\vec{c}_q[0], \vec{c}_q[1], \ldots, \vec{c}_q[M_c - 1]\}$ the *qth decimation* of $\vec{c}, q = 1, 2, \ldots, M_c$, where $\vec{c}_q[k] \triangleq \vec{c}[qk \pmod{M_c}], k = 0, 1, \ldots, M_c - 1$.

Consider the 8-PSK constellation shown in Fig. 2 with $\vec{c} = \{0, 1, 2, 3, 4, 5, 6, 7\}$. It is straightforward to verify that $\vec{c}_1 = \{0, 1, 2, 3, 4, 5, 6, 7\}, \vec{c}_2 = \{0, 2, 4, 6, 0, 2, 4, 6\}$, $\vec{c}_3 = \{0, 3, 6, 1, 4, 7, 2, 5\}, \ldots$ One purpose of applying constellation decimation is to implement symbol permutation in a structured way. One can see that some decimation factors (e.g., $q = 2$) are *improper* in the sense that the decimated sequence does not include all symbols in $\mathcal{A}_s$. To avoid such improper decimation factors, we recall the following result.

Proposition 3: [32] Sequence $\vec{c}_q$ has period $M_c/\gcd(M_c, q)$, where $\gcd(\cdot)$ denotes the greatest common divisor.

Hence, *proper decimation* requires that M_c and q be relatively prime. The codebook $\boldsymbol{\mathcal{B}}_s$ is an $(L + 1) \times M_c$ matrix, each row of which is obtained by a proper decimation of $\vec{c}$. We use the notation

$$\boldsymbol{\mathcal{B}}_s = \langle q_0, \quad q_1, \ldots, \quad q_L \rangle \tag{42}$$

to signify that codebook $\boldsymbol{\mathcal{B}}_s$ is obtained by using decimation factors $q_j, j = 0, 1, \ldots, L$ for the jth row of $\boldsymbol{\mathcal{B}}_s$. Some examples of using constellation decimation for code construction are in order.

Example 1 (Repetition Coding): Consider the 8-PSK constellation shown in Fig. 2 with unit-energy symbols and code rate $R_s = 1$ bit per coded symbol. This implies that the minimum constellation size $M_c = N_c = 8$. Assume $L + 1 = 3$ (three-ray channel). Using all-one decimation factors, we obtain $\boldsymbol{\mathcal{B}}_s = \langle 1, 1, 1 \rangle$, or

$$\boldsymbol{\mathcal{B}}_s^{\langle 1,1,1 \rangle} = \begin{bmatrix} 0, & 1, & 2, & 3, & 4, & 5, & 6, & 7 \\ 0, & 1, & 2, & 3, & 4, & 5, & 6, & 7 \\ 0, & 1, & 2, & 3, & 4, & 5, & 6, & 7 \end{bmatrix} \tag{43}$$

which is effectively a repetition code. The fact that repetition codes achieve full spectral diversity can be explained in an intuitive manner. Specifically, since an Lth-order channel can have at most L channel zeros, one of the $L+1$ repetitive transmissions is guaranteed to survive. The minimum product distance for this code is $\delta_{\min}^{\langle 1,1,1 \rangle} = d_1^3$ (cf. Fig. 2), where $d_1 = 2\sin(\pi/8)$, which turns out to be the smallest among all full-diversity codes. Codes with a larger minimum product distance are appealing since they translate to a larger coding gain.

Example 2 (Optimum LCD Coding): Consider again the 8-PSK constellation with unit-energy symbols, $R = 1$ bit per coded symbol, $M_c = N_c = 8$, and $L + 1 = 3$. New LCD codes can be constructed by changing the decimation factors. Let us examine $\boldsymbol{\mathcal{B}}_s = \langle 1, 3, 5 \rangle$, or

$$\boldsymbol{\mathcal{B}}_s^{\langle 1,3,5 \rangle} = \begin{bmatrix} 0, & 1, & 2, & 3, & 4, & 5, & 6, & 7 \\ 0, & 3, & 6, & 1, & 4, & 7, & 2, & 5 \\ 0, & 5, & 2, & 7, & 4, & 1, & 6, & 3 \end{bmatrix}. \tag{44}$$

The minimum product distance for this code is $\delta_{\min}^{\langle 1,3,5 \rangle} = d_3 d_1^2$, where $d_1 = 2\sin(\pi/8)$, and $d_3 = 2\sin(3\pi/8)$. At high SNR, this translates to a coding gain [cf. (36)] of $10\log_{10}(\delta_{\min}^{\langle 1,3,5 \rangle}/\delta_{\min}^{\langle 1,1,1 \rangle})^{2/(L+1)} \approx 2.55$ dB relative to the previous repetition code. The $\langle 1, 3, 5 \rangle$ code is in fact the optimum LCD code in the sense that it achieves the largest coding gain among all LCD codes with the same constellation size and code length. It is, however, not the only optimum LCD code. A quick computer search indicates that there are a total of seven optimum LCD codes, namely, $\langle 1, 1, 3 \rangle, \langle 1, 1, 5 \rangle, \langle 1, 3, 3 \rangle, \langle 1, 3, 5 \rangle, \langle 1, 3, 7 \rangle, \langle 1, 5, 5 \rangle, \langle 1, 5, 7 \rangle$, all yielding the same minimum product distance. Note that LCD codes differing by a permutation of the decimation factors are considered equivalent, and only one of them is listed. For example, the following codes are considered equivalent: $\langle 1, 1, 3 \rangle, \langle 1, 3, 1 \rangle, \langle 3, 1, 1 \rangle$.

TABLE I
OPTIMUM LCD CODES, WHERE R_s DENOTES THE CODE RATE IN BITS PER CODED SYMBOL, $L+1$ DENOTES THE BLOCK LENGTH (DIVERSITY ORDER), $\delta_{\min}$ DENOTES THE MINIMUM PRODUCT DISTANCE OF THE CODE, AND M_c SIGNIFIES THE USE OF THE M_c-ARY PSK CONSTELLATION

	$L+1$	M_c	$\delta_{\min}$	$\langle q_1, \cdots, q_L \rangle$
$R_s = 1$	1	2	2	$\langle \emptyset \rangle$
	2	4	2	$\langle 1 \rangle$, $\langle 3 \rangle$
	3	8	1.0824	$\langle 1,3 \rangle$, $\langle 1,5 \rangle$, $\langle 3,3 \rangle$, $\langle 3,5 \rangle$, $\langle 3,7 \rangle$, $\langle 5,5 \rangle$, $\langle 5,7 \rangle$
	4	16	1.4142	$\langle 3,5,7 \rangle$, $\langle 3,5,9 \rangle$, $\langle 3,7,11 \rangle$, $\langle 3,9,11 \rangle$, $\langle 5,7,13 \rangle$, $\langle 5,9,13 \rangle$, $\langle 7,11,13 \rangle$, $\langle 9,11,13 \rangle$
$R_s = 2$	1	4	1.4142	$\langle \emptyset \rangle$
	2	16	0.5858	$\langle 7 \rangle$, $\langle 9 \rangle$
	3	64	0.1692	$\langle 11,27 \rangle$, $\langle 11,37 \rangle$, $\langle 15,29 \rangle$, $\langle 15,35 \rangle$, $\langle 17,19 \rangle$, $\langle 17,45 \rangle$, $\langle 19,47 \rangle$, $\langle 27,53 \rangle$, $\langle 29,49 \rangle$, $\langle 35,49 \rangle$, $\langle 37,53 \rangle$, $\langle 45,47 \rangle$
	4	256	0.0381	$\langle 25,97,107 \rangle$, $\langle 25,97,149 \rangle$, $\langle 25,107,159 \rangle$, $\langle 25,149,159 \rangle$, $\langle 35,41,119 \rangle$, $\langle 35,41,137 \rangle$, $\langle 35,119,215 \rangle$, $\langle 35,137,215 \rangle$, $\langle 41,119,221 \rangle$, $\langle 41,137,221 \rangle$, $\langle 67,99,117 \rangle$, $\langle 67,99,139 \rangle$, $\langle 67,117,157 \rangle$, $\langle 67,139,157 \rangle$, $\langle 71,75,95 \rangle$, $\langle 71,75,161 \rangle$, $\langle 71,95,181 \rangle$, $\langle 71,161,181 \rangle$, $\langle 75,95,185 \rangle$, $\langle 75,161,185 \rangle$, $\langle 95,181,185 \rangle$, $\langle 97,107,231 \rangle$, $\langle 97,149,231 \rangle$, $\langle 99,117,189 \rangle$, $\langle 99,139,189 \rangle$, $\langle 107,159,231 \rangle$, $\langle 117,157,189 \rangle$, $\langle 119,215,221 \rangle$, $\langle 137,215,221 \rangle$, $\langle 139,157,189 \rangle$, $\langle 149,159,231 \rangle$, $\langle 161,181,185 \rangle$

The above examples show that optimum LCD codes attain an impressive coding gain over heuristic codes such as repetition codes. In fact, the relative coding gain becomes even more prominent as the code rate and/or diversity order increases. For example, at a code rate of $R_s = 2$ bits per coded symbol with $M_c = N_c = 64$, a comparison of the minimum product distance of the optimum LCD code $\langle 1, 15, 35 \rangle$ (see Table I) and the repetition code $\langle 1, 1, 1 \rangle$ reveals that the former yields a coding gain of about 15 dB over the latter!

Remark 1: The LCD codes are *linear*, since the scaling (modulo-M_c) of a codeword or the sum (modulo-M_c) of any two codewords results in another valid codeword. In fact, all LCD codes can be thought of as nonbinary linear $(L+1, 1)$ block codes, where each codeword contains one message symbol, and the code length $L+1$ is the minimum length required to achieve the maximum diversity. LCD codes are also *maximum distance separable* since their minimum Hamming distance attains the *Singleton bound* $L+1$, which is an upper bound on the minimum Hamming distance of all linear block codes [24 , p. 461]. *Systematic* LCD codes are easy to implement, simply by adopting the decimation factor $q = 1$ at least once (usually, it is $q_0 = 1$ by convention).

Since LCD codes are linear, the set of Hamming distances between any pair of distinct LCD codewords is identical to the set of Hamming weights of the nonzero codewords in the code [24 , p. 419]. This simplifies remarkably the search for LCD codes with good distance structure. In particular, we only need to examine the weight structure of the N_c codewords, rather than the distance between N_c^2 codeword pairs. In Table I, we list optimum LCD codes resulting from our computer searches for several values of the code rate and diversity order. Systematic encoding is assumed with $q_0 = 1$. Hence, only $\langle q_1, \ldots, q_L \rangle$ are shown for brevity. Note that for frequency-flat channels (i.e., $L = 0$), the need for spectral encoding vanishes; the decimation factor is shown as $\emptyset$ in this case. The minimum product distances $\delta_{\min}$ of the various LCD codes are also shown in Table I. It is clear that as the constellation size increases, optimum LCD codes are abundant, and all achieve identically the same diversity and coding gain. Note that diversity gain and coding gain are only asymptotic (high SNR) performance measures. Hence, optimum LCD codes listed within the same category may have (slightly) different finite-SNR performance.

Remark 2: We see from Table I that as L (channel order) increases, the constellation size M_c increases rapidly, and so does the decoding complexity. In general, for systems with a large channel order, it suffices to use LCD codes based on a hypothesized channel order that is (usually much) smaller than the true channel order, since it is known that the diversity-induced performance gain diminishes as the diversity order increases (e.g., [24]). For systems rich in both spatial and spectral diversity, it is recommended to maximize the spatial diversity first, before seeking additional spectral diversity. We will further elaborate on this point with a numerical example in Section VI.

Remark 3: LCD spectral codes are also closely related to the cyclic space-time diagonal codes of [9]. In particular, the latter can be obtained by forming diagonal code matrices using LCD

code vectors. Our LCD interpretation establishes an interesting connection to permutation codes (via Theorem 4) and offers a new perspective for code construction. Specifically, it is possible to employ other linear or nonlinear constellation operations that may result in new codes with not only full diversity but higher coding gain as well. This remains an interesting future research subject.

Remark 4: Finally, we comment on the transmission rate of the proposed DSTF system. For a DSTF system using a spectral encoder with code rate $R_{\rm s}$ bits per coded symbol, the overall transmission rate R is defined as $R \triangleq R_{\rm d}R_{\rm s}$, where we recall that $R_{\rm d}$ is the rate of the orthogonal design used for differential encoding (see Section III).

D. Maximum Likelihood Detection

With subcarrier interleaving, the ML detector based on (35) reduces to M parallel detectors, each operating on signals received on the mth subcarrier subset [cf. (22)]:

$$\hat{\mathbf{s}}_m^{\rm ML}(nN_s+k) = \arg \max_{\mathbf{s}_{m,k}\in\boldsymbol{\mathcal{B}}_{s,m}} \Re\left\{\mathbf{z}_m^H(nN_s+k)\boldsymbol{\Omega}_{y,m}^{1/2}(n-1)\mathbf{s}_{m,k}\right\} \quad m=0,\ldots,M-1; \quad k=0,\ldots,N_s-1 \quad (45)$$

where $\boldsymbol{\mathcal{B}}_{s,m}$ denotes the codebook for the mth subcarrier subset.

When repetition coding is utilized with subcarrier interleaving, the ML detector (45) can be simplified. In particular, let $\boldsymbol{\zeta}_m(nN_s+k) \triangleq \boldsymbol{\Omega}_{y,m}^{1/2}(n-1)\mathbf{z}_m(nN_s+k)$, with $\zeta_{m,j}(nN_s+k)$ denoting its jth element. In addition, let $s_m(nN_s+k)$ denote the message symbol carried by the repetition codeword $\mathbf{s}_m(nN_s+k)$. The ML detector reduces to

$$\hat{s}_m^{\rm ML}(nN_s+k) = \arg \max_{s_{m,k}\in\mathcal{A}_s} \Re\left\{\left[\sum_{j=0}^{L}\zeta_{m,j}(nN_s+k)\right]^* s_{m,k}\right\} \quad m=0,\ldots,M-1; \quad k=0,\ldots,N_s-1, \quad (46)$$

which reduces to rounding the phase angle of $\sum_{j=0}^{L}\zeta_{m,j}(nN_s+k)$ to the nearest multiple of $2\pi/M_c$ (assuming M_c-ary PSK).

VI. Numerical Simulations

We consider an OFDM system with $P=48$ subcarriers, and the following schemes are utilized for differential modulation. The performance measure is the bit-error rate (BER) as a function of E_s/N_o, which is the SNR per receive antenna per subcarrier [cf. (2) and (11)].

- **DT**: Differential OFDM with *one* transmit (Tx) antenna and scalar differential PSK. It is acronymed as DT since differential modulation is performed across two adjacent OFDM symbol intervals (i.e., time) and not across frequency or space. This system yields no diversity.
- **DST**: Differential space-time coded OFDM with *multiple* Tx antennas. Differential modulation is performed on each subcarrier using the cyclic group differential space-time codes [9], which encode across space and time but not frequency.

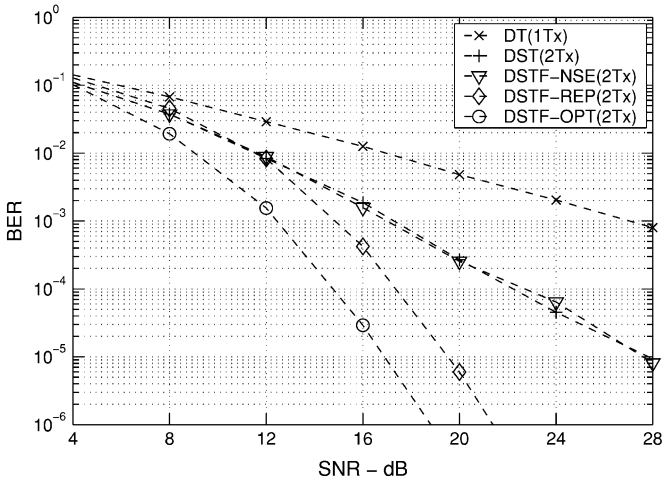

Fig. 3. BER versus E_s/N_0 with $N_t=2$ Tx antennas, $N_r=1$ Rx antenna, and transmission rate $R=1$ b/s/Hz in independent Rayleigh block-fading channels of channel order $L=2$.

- **DSTF-NSE**: The proposed DSTF modulation scheme with *multiple* Tx antennas but no spectral encoding (NSE). Hence, similar to DST, coding is performed across space and time but not frequency.
- **DSTF-REP**: DSTF with *multiple* Tx antennas and *repetition codes* for spectral encoding.
- **DSTF-OPT**: DSTF with *multiple* Tx antennas and *optimum LCD codes* for spectral encoding.

A. Independent Rayleigh Block-Fading Channels With Known Channel Order

In this case, the channel coefficients are generated as i.i.d. complex Gaussian variables with zero-mean and variance $1/(L+1)$, where $L=2$. They are fixed within one block consisting of multiple DSTF matrices and change independently from one block to another. The following examples also assume $N_r=1$ receive (Rx) antenna for all methods.

Fig. 3 depicts the BER of the single-antenna DT and the multiantenna methods for $N_t=2$ when the transmission rate is $R=1$ b/s/Hz (see Remark 4 for the definition of R). To achieve $R=1$ b/s/Hz, both DT and DSTF-NSE employ BPSK and DST utilizes the rate-1 2-Tx cyclic group code with quadrature phase shift keyng (QPSK) [9], whereas both DSTF-REP and DSTF-OPT use 8PSK. It is seen that DST and DSTF-NSE achieve a diversity order of 2 (viz., spatial diversity), as can be verified from the BER-SNR slope. This confirms that none of them yields spectral diversity. We see that spectral diversity is achieved by DSTF-REP and DSTF-OPT. An inspection of the BER-SNR slope indicates that both schemes achieve a diversity order of about 6 at high SNR. This is the maximum spatio-spectral diversity order offered by the system. Finally, we observe that DSTF-OPT yields a coding gain of about 2.5 dB relative to DSTF-REP, which agrees with the prediction made in Example 2 of Section V-C.

Fig. 4 depicts a scenario similar to that of the previous example, except that $R=2$ b/s/Hz. To achieve this transmission rate and both DT and DSTF-NSE employ QPSK, DST utilizes the rate-2 2-Tx cyclic group code with 16PSK [9], while both DSTF-REP and DSTF-OPT use 64PSK. Observations are

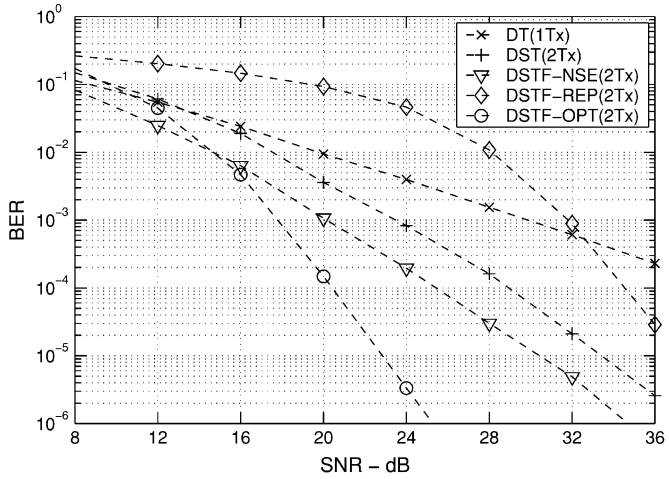

Fig. 4. BER versus E_s/N_0 with $N_t = 2$ Tx antennas, $N_r = 1$ Rx antenna, and transmission rate $R = 2$ b/s/Hz in independent Rayleigh block-fading channels of channel order $L = 2$.

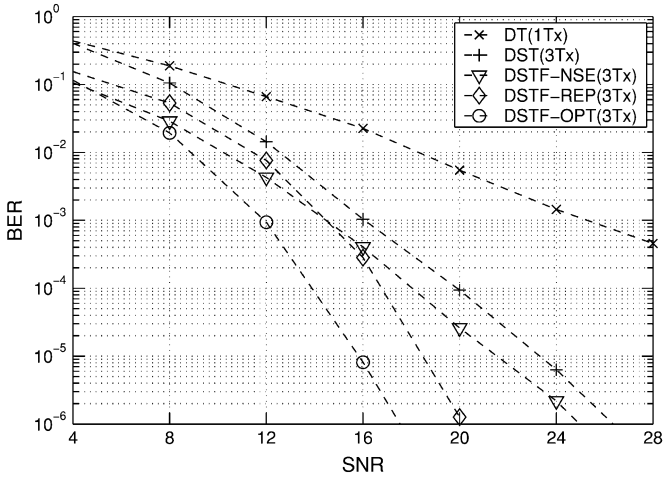

Fig. 5. BER versus E_s/N_0 with $N_t = 3$ Tx antennas, $N_r = 1$ Rx antenna, and transmission rate $R = .75$ b/s/Hz in independent Rayleigh block-fading channels of channel order $L = 2$.

similar to the last example with a few exceptions. First, we notice that even though both DSTF-NSE and DST achieve a similar diversity order, the former, which does not enforce a group structure in coding, yields an energy saving of about 3 dB over the group-based DST. This was also observed in [7] for high-rate transmission in flat-fading channels. Second, DSTF-REP becomes energy ineffective due to loss in coding gain. The loss relative to DSTF-OPT is about 15 dB at high SNR, which agrees with the prediction made in Section V-C. Note that the BER-SNR slope of DSTF-OPT at high SNR is similar to that of DSTF-OPT, indicating that the former is still a full-diversity scheme.

We now consider $N_t = 3$ Tx antennas for the multiantenna schemes. Since no rate-1 complex orthogonal design exists for $N_t = 3$, all DSTF methods employ the $R_{\rm d} = 3/4$ square orthogonal design in (8). To account for rate loss and make fair comparison, a rate-3/4 convolutional code is used for all non-DSTF schemes so that all methods are compared at the same spectral efficiency. The rate-3/4 code is obtained by puncturing a rate-1/2 convolutional code [24 , p. 499]. It is used along with subcarrier interleaving to provide coding across subcarriers (but not across OFDM symbols), which may provide spectral diversity.

Fig. 5 depicts the results when $N_t = 3$ Tx antennas are used by the multiantenna schemes, and the effective transmission rate is $R = .75$ b/s/Hz. In this case, both DT and DSTF-NSE employ BPSK, DST utilizes the rate-1 3-Tx cyclic group code with 8PSK [9], and both DSTF-REP and DSTF-OPT use 8PSK. Meanwhile, Fig. 6 depicts a similar scenario, except that the transmission rate is $R = 1.5$ b/s/Hz. To achieve this rate, both DT and DSTF-NSE employ QPSK, DST utilizes the rate-2 3-Tx cyclic group code with 64PSK [9], and both DSTF-REP and DSTF-OPT use 64PSK. From both figures, it is seen that DSTF-OPT overall compares favorably with the other schemes. In addition, the following remark is in order.

Remark 5: It is seen from Figs. 5 and 6 that the rate-3/4 convolution code cannot provide full spectral diversity. Interestingly, Wang and Giannakis [13] compared their hand-crafted full-diversity spectral codes with conventional error-control

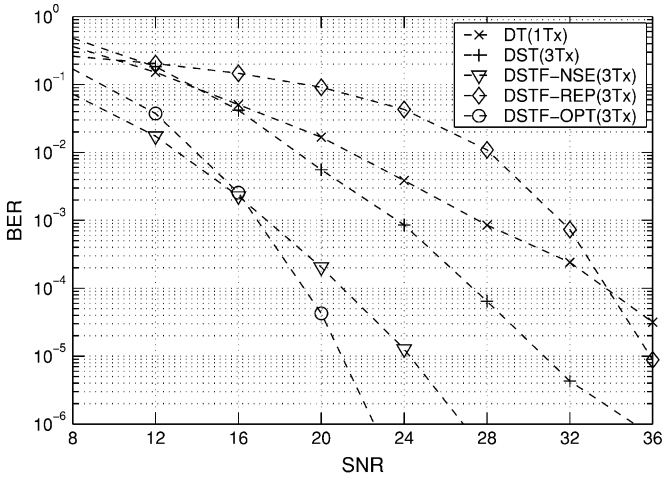

Fig. 6. BER versus E_s/N_0 with $N_t = 3$ Tx antennas, $N_r = 1$ Rx antenna, and transmission rate $R = 1.5$ b/s/Hz in independent Rayleigh block-fading channels of channel order $L = 2$.

codes (ECCs) in OFDM systems using coherent detection and had a similar observation. In principle, it is possible to use more powerful ECCs along with differential space-time modulation to provide full spatio-spectral diversity. However, it remains unclear how to select such codes, how they should be used with constellation mapping to provide the necessary minimum Hamming distance (cf. Theorem 4), and how subcarrier interleaving should be performed, etc. Indeed, all these issues are critical to the diversity and coding gain that is achieved in the end. Due to the popularity of standard ECCs in many wireless systems, an investigation into these issues is well motivated and should be pursued in the future.

B. Unknown Channel Order

In some cases, knowledge of the exact channel order L might be unavailable, and it would be of interest to examine the impact on DSTF if an incorrect channel order L' is assumed. To this end, we consider a scenario of Rayleigh fading involving $L+1 = 3$ independent paths with equal variance $1/3$, and the hypothesized channel order is $L' = 1, 2$ and 3, respectively.

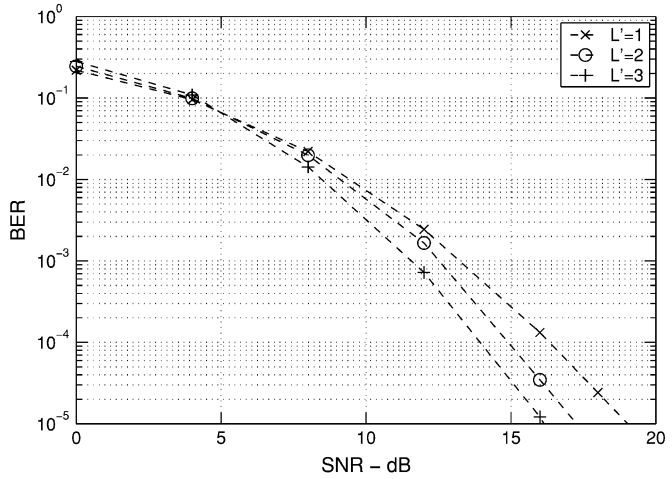

Fig. 7. BER versus E_s/N_0 for DSTF-OPT with $N_t = 2$ Tx antennas, $N_r = 1$ Rx antenna, and transmission rate $R = 1$ b/s/Hz in independent Rayleigh block-fading channels of channel order $L = 2$ when the channel order is assumed to be $L' = 1, 2$, and 3.

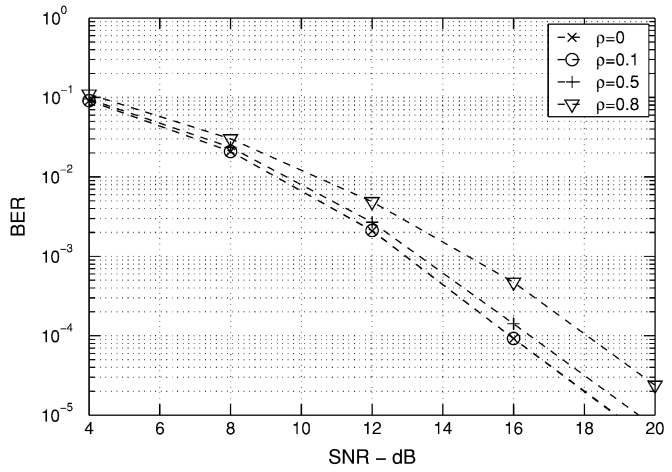

Fig. 8. BER versus E_s/N_0 for DSTF-OPT with $N_t = 2$ Tx antennas, $N_r = 1$ Rx antenna, and transmission rate $R = 1$ b/s/Hz in correlated Rayleigh block-fading channels of channel order $L = 1$ when the channel correlation coefficient is $\rho = 0, 0.1, 0.5$, and 0.8.

The results are shown in Fig. 7 for $R = 1$ b/s/Hz, where only DSTF-OPT is considered. The corresponding constellation is QPSK for $L' = 1$, 8PSK for $L' = 2$, and 16PSK for $L' = 3$. Fig. 7 suggests that overestimating L (i.e., $L' = 3$) incurs no loss of diversity since the BER-SNR slope for both $L' = 2$ and 3 are almost identical. In fact, it appears to slightly improve the coding gain compared to $L' = 2$. On the other hand, underestimating $L(L' = 1)$ leads to some loss of spectral diversity.

C. Correlated Fading Channels

We now examine the effect of correlated channel fading on DSTF by considering a scenario where the channel taps of each Tx-Rx pair are correlated, but channels among different Tx-Rx pairs are independent. In particular, the channel correlation matrix (see Section IV) is assumed to be $\mathbf{R}_h = (L+1)^{-1}(\mathbf{I}_{N_t} \otimes \mathbf{R}_0)$, where $\mathbf{R}_0$ is a symmetric Toeplitz matrix with the first row given by $[1, \rho, \ldots, \rho^L]$. Hence, ρ captures the significance of correlation among the channel taps. Fig. 8 depicts the BER for

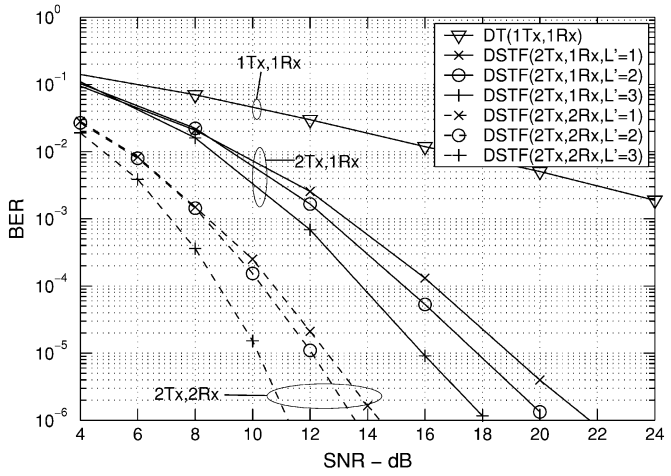

Fig. 9. BER versus E_s/N_0 for DT and DSTF-OPT with transmission rate $R = 1$ b/s/Hz in Rayleigh continuous-fading channels simulated according to Channel Model A ($L = 8$) specified in HIPERLAN/2.

DSTF-OPT with $R = 1$ b/s/Hz when the channel coefficients are independent $(\rho = 0)$, weakly correlated $(\rho = 0.1)$, moderately correlated $(\rho = 0.5)$, and strongly correlated $(\rho = 0.8)$. It is seen from Fig. 8 that for all cases, full diversity is achieved, which supports the previous analysis in Sections IV and V. However, as the correlated increases, a loss of coding gain becomes more pronounced.

D. Continuous-Fading Channels

In the final example, we consider continuous-fading channels. The channel coefficients are generated according to Channel Model A with $L = 8$, as specified in HIPERLAN/2 [33], and are time-varying according to the Jakes model with terminal speed of 3 m/s and carrier frequency of 5.2 GHz. As noted in Remark 2, attempting to achieve full spectral diversity in this case is computationally expensive. Instead, we seek full spatial but partial spectral diversity using a hypothesized channel order L'. Fig. 9 depicts the BER for DSTF-OPT with transmission rate $R = 1$ b/s/Hz, hypothesized channel $L' = 1, 2$, and 3, $N_t = 2$ Tx antennas, and $N_r = 1$ or 2 Rx antennas, respectively. As a reference, the BER of the single-Tx single-Rx DT scheme is also shown. It is seen that only incremental performance gain is achieved as L' increases. On the other hand, when $N_r = 1$ is increased to $N_r = 2$ Rx antennas, the gain becomes considerably larger. This suggests that spatial diversity should be maximized first before seeking additional spectral diversity.

VII. Conclusion

We have presented a differential modulation scheme (viz., DSTF) for systems equipped with multitransmit antennas in frequency-selective channels. Through both analysis and simulation, we have shown that the DSTF system guarantees full spatio-spectral diversity and significant coding gain. DSTF is computationally quite appealing due to several unique designs. In particular, a unitary structure based on block orthogonal designs has been imposed for differential encoding, which renders the ML decoding decoupled in space and time. For spectral encoding, we have introduced LCD codes that are nonbinary block

codes obtained by decimating a phase-shift-keying constellation with a group of decimation factors that are co-prime with the constellation size. Since LCD codes have the shortest code length for full diversity, they incur modest decoding complexity among all full-diversity codes.

It should be noted that although it is possible to use standard error-control codes for spectral encoding along with differential space-time modulation to provide joint spatio-spectral diversity, there are a number of issues that are yet to be resolved, including how to select such full-diversity codes, how they should be used with constellation mapping to provide the necessary minimum Hamming distance, and how subcarrier interleaving should be performed, etc., in order to ensure full diversity. On the other hand, it should also be noted that the proposed LCD codes can be combined with standard error-control codes to achieve additional coding gain without losing full diversity. These issues will be examined in the future.

APPENDIX A
PROOF OF THEOREM 1

For notational brevity, let

$$\begin{aligned}
\boldsymbol{s}(n) &\triangleq [\mathbf{s}^T(nN_s), \quad \ldots, \quad \mathbf{s}^T(nN_s+N_s-1)]^T \\
\tilde{\mathbf{A}}_k &\triangleq \sum_{l=0}^{N_d-1} [\mathbf{a}_{k,l} \otimes \mathbf{D}_y((n-1)N_d+l)] \\
\tilde{\mathbf{A}} &\triangleq [\tilde{\mathbf{A}}_0, \quad \ldots, \quad \tilde{\mathbf{A}}_{N_s-1}]_{N_dP\times N_sP} \\
\tilde{\mathbf{B}}_k &\triangleq \sum_{l=0}^{N_d-1} [\mathbf{b}_{k,l} \otimes \mathbf{D}_y((n-1)N_d+l)] \\
\tilde{\mathbf{B}} &\triangleq [\tilde{\mathbf{B}}_0, \quad \ldots, \quad \tilde{\mathbf{B}}_{N_s-1}]_{N_dP\times N_sP}.
\end{aligned}$$

Then, (21) can be written as

$$\boldsymbol{y}(n) = N_s^{-1/2}[\tilde{\mathbf{A}}\boldsymbol{s}(n) + \tilde{\mathbf{B}}\boldsymbol{s}^*(n)] + \boldsymbol{v}(n). \tag{47}$$

Let $\tilde{\boldsymbol{y}}(n) \triangleq [\boldsymbol{y}^T(n), \boldsymbol{y}^H(n)]^T, \tilde{\boldsymbol{s}}(n) \triangleq [\boldsymbol{s}^T(n), \boldsymbol{s}^H(n)]^T$, and $\tilde{\boldsymbol{v}}(n) \triangleq [\boldsymbol{v}^T(n), \boldsymbol{v}^H(n)]^T$. We have

$$\tilde{\boldsymbol{y}}(n) = N_s^{-1/2}\boldsymbol{\Gamma}\tilde{\boldsymbol{s}}(n) + \tilde{\boldsymbol{v}}(n) \tag{48}$$

where

$$\boldsymbol{\Gamma} \triangleq \begin{bmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{B}} \\ \tilde{\mathbf{B}}^* & \tilde{\mathbf{A}}^* \end{bmatrix}.$$

Recall that the real and imaginary components of the complex Gaussian vectors $\boldsymbol{v}(n)$ are i.i.d. with zero-mean and variance N_0 per dimension (see Section III). It follows that $\tilde{\boldsymbol{v}}(n)$ is also complex Gaussian with zero-mean and covariance $E\{\tilde{\boldsymbol{v}}(n)\tilde{\boldsymbol{v}}^H(n)\} = 2N_0\mathbf{I}_{2N_dP}$. Hence, the ML detection of $\boldsymbol{s}(n)$ based on (48) amounts to the minimization of the Euclidean distance $||\tilde{\boldsymbol{y}}(n) - N_s^{-1/2}\boldsymbol{\Gamma}\tilde{\boldsymbol{s}}(n)||^2$. A brute-force minimization would incur exponential complexity. This can be avoided by exploiting the following result.

Proposition 4: The $2N_dP \times 2N_sP$ matrix $\boldsymbol{\Gamma}_1 \triangleq \boldsymbol{\Gamma}[\mathbf{I}_{2N_s} \otimes \boldsymbol{\Omega}_y^{-1/2}(n-1)]$, where $\boldsymbol{\Omega}_y(n-1)$ is defined in (24), is semi-unitary:

$$\boldsymbol{\Gamma}_1^H\boldsymbol{\Gamma}_1 = \mathbf{I}_{2N_sP}. \tag{49}$$

Proof: We first note from (4)–(6) that

$$\mathbf{a}_{k_1,l_1}^T\mathbf{a}_{k_2,l_2} + \mathbf{b}_{k_2,l_1}^T\mathbf{b}_{k_1,l_2} = \delta(k_1-k_2)\delta(l_1-l_2) \tag{50}$$

$$\mathbf{a}_{k_1,l_1}^T\mathbf{b}_{k_2,l_2} + \mathbf{a}_{k_2,l_1}^T\mathbf{b}_{k_1,l_2} = 0, \quad \forall k_1,k_2,l_1,l_2. \tag{51}$$

Exploiting (50), we have [for notational brevity, we denote $\mathbf{D}_y((n-1)N+l)$ by $\mathbf{D}_y(l)$]

$$\begin{aligned}
&\tilde{\mathbf{A}}_{k_1}^H\tilde{\mathbf{A}}_{k_2} + \tilde{\mathbf{B}}_{k_1}^T\tilde{\mathbf{B}}_{k_2}^* \\
&\quad= \sum_{l_1}\sum_{l_2}\{[\mathbf{a}_{k_1,l_1}^T \otimes \mathbf{D}_y^H(l_1)]\,[\mathbf{a}_{k_2,l_2} \otimes \mathbf{D}_y(l_2)] \\
&\qquad+ [\mathbf{b}_{k_1,l_1}^T \otimes \mathbf{D}_y^T(l_1)]\,[\mathbf{b}_{k_2,l_2} \otimes \mathbf{D}_y^*(l_2)]\} \\
&\quad= \sum_{l_1}\sum_{l_2}[\mathbf{a}_{k_1,l_1}^T\mathbf{a}_{k_2,l_2} + \mathbf{b}_{k_1,l_2}^T\mathbf{b}_{k_2,l_1}] \\
&\qquad\otimes [\mathbf{D}_y^H(l_1)\mathbf{D}_y(l_2)] \\
&\quad= \boldsymbol{\Omega}_y(n-1)\delta(k_1-k_2)
\end{aligned}$$

where the second equality is due to the mixed product rule of the Kronecker product [23 , p. 244]. Likewise, using (51) yields

$$\begin{aligned}
&\tilde{\mathbf{A}}_{k_1}^H\tilde{\mathbf{B}}_{k_2} + \tilde{\mathbf{B}}_{k_1}^T\tilde{\mathbf{A}}_{k_2}^* \\
&\quad= \sum_{l_1}\sum_{l_2}[\mathbf{a}_{k_1,l_1}^T\mathbf{b}_{k_2,l_2} + \mathbf{b}_{k_1,l_2}^T\mathbf{a}_{k_2,l_1}] \\
&\qquad\otimes [\mathbf{D}_y^H(l_1)\mathbf{D}_y(l_2)] = \mathbf{0}.
\end{aligned}$$

Hence

$$\tilde{\mathbf{A}}^H\tilde{\mathbf{A}} + \tilde{\mathbf{B}}^T\tilde{\mathbf{B}}^* = \mathbf{I}_{N_s} \otimes \boldsymbol{\Omega}_y(n-1) = \tilde{\mathbf{A}}^T\tilde{\mathbf{A}}^* + \tilde{\mathbf{B}}^H\tilde{\mathbf{B}} \tag{52}$$

$$\tilde{\mathbf{A}}^H\tilde{\mathbf{B}} + \tilde{\mathbf{B}}^T\tilde{\mathbf{A}}^* = \tilde{\mathbf{A}}^T\tilde{\mathbf{B}}^* + \tilde{\mathbf{B}}^H\tilde{\mathbf{A}} = \mathbf{0} \tag{53}$$

where the second equality of (52) is because $\boldsymbol{\Omega}_y(n-1)$ is diagonal with non-negative entries [see (24)]. It follows that

$$\begin{aligned}
\boldsymbol{\Gamma}^H\boldsymbol{\Gamma} &= \begin{bmatrix} \tilde{\mathbf{A}}^H\tilde{\mathbf{A}} + \tilde{\mathbf{B}}^T\tilde{\mathbf{B}}^* & \tilde{\mathbf{A}}^H\tilde{\mathbf{B}} + \tilde{\mathbf{B}}^T\tilde{\mathbf{A}}^* \\ \tilde{\mathbf{A}}^T\tilde{\mathbf{B}}^* + \tilde{\mathbf{B}}^H\tilde{\mathbf{A}} & \tilde{\mathbf{A}}^T\tilde{\mathbf{A}}^* + \tilde{\mathbf{B}}^H\tilde{\mathbf{B}} \end{bmatrix} \\
&= \mathbf{I}_{2N_s} \otimes \boldsymbol{\Omega}_y(n-1)
\end{aligned}$$

which implies (49). ■

Based on Proposition 4, we can construct a unitary matrix $\tilde{\boldsymbol{\Gamma}} \triangleq [\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2]$, where $\boldsymbol{\Gamma}_2$ is a $2N_dP \times 2(N_d-N_s)P$ matrix (obtained by, e.g., a Gram–Schmidt procedure [30]) that is orthogonal to $\boldsymbol{\Gamma}_1$: $\boldsymbol{\Gamma}_2^H\boldsymbol{\Gamma}_1 = \mathbf{0}$. Without altering the ML solution, we multiply both sides of (48) by $\tilde{\boldsymbol{\Gamma}}$, which yields

$$\begin{aligned}
\tilde{\boldsymbol{z}}_1(n) &\triangleq \boldsymbol{\Gamma}_1^H\tilde{\mathbf{y}}(n) \\
&= N_s^{-1/2}[\mathbf{I}_{2N_s} \otimes \boldsymbol{\Omega}_y^{1/2}(n-1)]\tilde{\boldsymbol{s}}(n) + \tilde{\boldsymbol{\nu}}_1(n)
\end{aligned} \tag{54}$$

$$\tilde{\boldsymbol{z}}_2(n) \triangleq \boldsymbol{\Gamma}_2^H\tilde{\mathbf{y}}(n) = \tilde{\boldsymbol{\nu}}_2(n) \tag{55}$$

where $\tilde{\boldsymbol{\nu}}_1(n) \triangleq \boldsymbol{\Gamma}_1\tilde{\boldsymbol{v}}(n)$, and $\tilde{\boldsymbol{\nu}}_2(n) \triangleq \boldsymbol{\Gamma}_2\tilde{\boldsymbol{v}}(n)$. The reason that (55) contains only noise is because $\boldsymbol{\Gamma}_2^H\boldsymbol{\Gamma} = \mathbf{0}$, which cancels the signal component. Thus, we can discard $\tilde{\boldsymbol{z}}_2(n)$. Moreover, it is noted that the two halves of $\tilde{\boldsymbol{z}}_1(n)$ are conjugate of each other, i.e., $\tilde{\boldsymbol{z}}_1(n) = [\boldsymbol{z}^T(n), \boldsymbol{z}^H(n)]^T$, where the $N_sP \times 1$ vector $\boldsymbol{z}(n)$ is formed by the first N_sP elements of $\tilde{\boldsymbol{z}}_1(n)$. Hence, it is

sufficient to use $\boldsymbol{z}(n)$ for detection, and $\boldsymbol{z}(n)$ can be expressed as

$$\boldsymbol{z}(n) = N_s^{-1/2}\left[\mathbf{I}_{N_s} \otimes \boldsymbol{\Omega}_y^{1/2}(n-1)\right]\boldsymbol{s}(n) + \boldsymbol{\nu}(n) \quad (56)$$

where $\boldsymbol{\nu}(n)$ is similarly formed by the first N_sP elements of $\tilde{\boldsymbol{\nu}}_1(n)$. Finally, we observe that, due to the block diagonal structure of the matrix inside the square brackets, the above equation reduces to N_s separate equations:

$$\mathbf{z}(nN_s+k) = N_s^{-1/2}\boldsymbol{\Omega}_y^{1/2}(n-1)\mathbf{s}(nN_s+k) + \boldsymbol{\nu}_k(nN_s+k), \quad k=0,\ldots,N_s-1 \quad (57)$$

where $\mathbf{z}(nN_s+k)$ and $\boldsymbol{\nu}_k(nN_s+k)$ denote the kth $P\times 1$ subvector of $\boldsymbol{z}(n)$ and $\boldsymbol{\nu}(n)$, respectively. It is noted that $\boldsymbol{\nu}_k(nN_s+k)$ are independent Gaussian vectors with zero-mean and covariance matrix $2N_0\mathbf{I}_P$. One can easily verify that $\mathbf{z}(nN_s+k)$ can be collectively expressed as in (23). The ML detector is the minimizer of the Euclidean distance $D \triangleq ||\mathbf{z}(nN_s+k) - N_s^{-1/2}\boldsymbol{\Omega}_y^{1/2}(n-1)\mathbf{s}(nN_s+k)||^2$. A simple expansion of D yields (dropping indices nN_s+k and $n-1$)

$$D = \mathbf{z}^H\mathbf{z} + N_s^{-1}\mathbf{s}^H\boldsymbol{\Omega}_y\mathbf{s} - 2N_s^{-1/2}\Re\left\{\mathbf{z}^H\boldsymbol{\Omega}_y^{1/2}\mathbf{s}\right\}. \quad (58)$$

The first term can be discarded since it is independent of vector $\mathbf{s}$. Using the assumption that $\mathbf{s}$ is drawn from a constant-modulus unit-energy constellation, one can see that the second term can also be discarded. In particular, recalling that $\boldsymbol{\Omega}_y$ is diagonal [see (24)], we have

$$\begin{aligned}\mathbf{s}^H\boldsymbol{\Omega}_y\mathbf{s} &= \sum_{l=0}^{N_d-1}\sum_{p=0}^{P-1}|y((n-1)N_d+l;p)|^2|s(nN_s+k;p)|^2 \\ &= \sum_{l=0}^{N_d-1}\sum_{p=0}^{P-1}|y((n-1)N_d+l;p)|^2\end{aligned}$$

where $y((n-1)N_d+l;p)$ and $s(nN_s+k;p)$ denote the pth sample of vectors $\mathbf{y}((n-1)N_d+l)$ and $\mathbf{s}(nN_s+k)$, respectively. Hence, the ML detection reduces to the maximization of the decoupled correlation metric (22), and the proof is complete.

REFERENCES

[1] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[2] 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, Jul. 1999.

[3] S. N. Diggavi, N. Al-Dhahir, A. Stamoulis, and A. R. Calderbank, "Differential space-time coding for frequency-selective channels," *IEEE Commun. Lett.*, vol. 6, no. 6, pp. 253–255, Jun. 2002.

[4] H. Li, X. Lu, and G. B. Giannakis, "Capon multiuser receiver for CDMA systems with space-time coding," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1193–1204, May 2002.

[5] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 7, pp. 1169–1174, Jul. 2000.

[6] H. Jafarkhani and V. Tarokh, "Multiple transmit antenna differential detection from generalized orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2626–2631, Sep. 2001.

[7] G. Ganesan and P. Stoica, "Differential detection based on space-time block codes," *IEEE Signal Process. Lett.*, vol. 9, no. 2, pp. 57–60, Feb. 2002.

[8] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2567–2578, Nov. 2000.

[9] B. M. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2041–2052, Dec. 2000.

[10] A. Shokrollahi, B. Hassibi, B. M. Hochwald, and W. Sweldens, "Representation theory for high-rate multiple-antenna code design," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2335–2367, Sep. 2001.

[11] J. Liu, J. Li, H. Li, and E. G. Larsson, "Differential space-code modulation for interference suppression," *IEEE Trans. Signal Process.*, vol. 49, no. 8, pp. 1786–1795, Aug. 2001.

[12] H. Li and J. Li, "Differential and coherent decorrelating multiuser receivers for space-time-coded CDMA systems," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2529–2537, Oct. 2002.

[13] Z. Wang and G. B. Giannakis, "Complex-field coding for OFDM over fading wireless channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp. 707–720, Mar. 2003.

[14] S. Zhou and G. B. Giannakis, "Space-time coding with maximum diversity gains over frequency-selective fading channels," *IEEE Signal Process. Lett.*, vol. 8, no. 10, pp. 269–272, Oct. 2001.

[15] Z. Liu, Y. Xin, and G. B. Giannakis, "Linear constellation-precoding for OFDM with maximum multipath diversity and coding gain," in *Proc. 35th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 2001, pp. 1445–1449.

[16] J. Boutros and E. Viterbo, "Signal space diversity: A power and bandwidth efficient diversity techniques for the Rayleigh fading channels," *IEEE Trans. Inf. Theory*, vol. 44, pp. 1453–1467, Jul. 1998.

[17] Y. Larsen, G. Leus, and G. B. Giannakis, "Constant modulus block differential encoding for frequency-selective channels," in *Proc. 36th Annu. Conf. Inf. Sci. Syst.*. Princeton, NJ, Mar. 2002.

[18] Z. Liu and G. B. Giannakis, "Block differentially encoded OFDM with maximum multipath diversity," *IEEE Trans. Wireless Commun.*, vol. 2, no. 3, pp. 420–423, May 2003.

[19] H. Li, "Differential space-time-frequency modulation with maximum space-multipath diversity," in *Proc. 40th Annual Allerton Conf. Commun., Contr. Comput.*, Monticello, IL, Oct. 2002.

[20] ——, "Differential space-time-frequency modulation over frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 7, no. 8, pp. 349–351, Aug. 2003.

[21] A. V. Geramita and J. Seberry, *Orthogonal Designs, Quadratic Forms and Hadamard Matrices Vol. 43 of Lecture Notes in Pure and Applied Mathematics*. New York: Marcel Dekker, 1979.

[22] Q. Ma, C. Tepedelenlioğlu, and Z. Liu, "Differential space-time-frequency coded OFDM with maximum diversity," in *Proc. 37th Annual Conf. Inf. Sci. Syst.*. Baltimore, MD, Mar. 2003.

[23] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1991.

[24] J. G. Proakis, *Digital Communications*, Fourth ed. New York: McGraw-Hill, 2000.

[25] Z. Liu, Y. Xin, and G. B. Giannakis, "Space-time-frequency coded OFDM over frequency-selective fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2465–2476, Oct. 2002.

[26] G. Ganesan and P. Stoica, "Space-time block codes: A maximum SNR approach," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1650–1656, May 2001.

[27] Z. Liu and G. B. Giannakis, "Space-time block coded multiple access through frequency-selective fading channels," *IEEE IEEE Trans. Commun.*, vol. 49, no. 6, pp. 1033–1044, Jun. 2001.

[28] O. Tirkkonen and A. Hottinen, "Square-matrix embeddable space-time block codes for complex signal constellations," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 384–395, Feb. 2002.

[29] S. M. Alamouti, "A simple transmit diversity techniques for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[30] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

[31] P. Stoica, H. Li, and J. Li, "Amplitude estimation of sinusoidal signals: Survey, new results, and an application," *IEEE Trans. Signal Process.*, vol. 48, pp. 338–352, Feb. 2000.

[32] D. V. Sarwate and M. B. Pursley, "Crosscorrelation properties of pseudorandom and related sequences," *Proc. IEEE*, vol. 68, no. 5, pp. 593–619, May 1980.

[33] *Channel Models for HiperLAN/2 in Different Indoor Scenarios*, 1998. ETSI Normalization Committee, document 3ERI085B.

Hongbin Li (M'99) received the B.S. and M.S. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, in 1991 and 1994, respectively, and the Ph.D. degree from the University of Florida, Gainesville, in 1999, all in electrical engineering.

From July 1996 to May 1999, he was a Research Assistant with the Department of Electrical and Computer Engineering, University of Florida. He was a Summer Visiting Faculty Member at the Air Force Research Laboratory, Rome, NY, in the summers of 2003 and 2004. Since July 1999, he has been an Assistant Professor with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ. His current research interests include wireless communications, statistical signal processing, and radars.

Dr. Li is a member of Tau Beta Pi and Phi Kappa Phi. He received the Harvey N. Davis Teaching Award in 2003 and the Jess H. Davis Memorial Award for excellence in research in 2001 from Stevens Institute of Technology, and the Sigma Xi Graduate Research Award from the University of Florida in 1999. He is an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS.