WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput. 2002; **2**:693–717 (DOI: 10.1002/wcm.92)

Complex field coded MIMO systems: performance, rate, and trade-offs

Xiaoli Ma and Georgios B. Giannakis*†

Department of Electrical and Computer Engineering,
University of Minnesota,
200 Union Street,
Minneapolis,
MN 55455, USA

Summary

The quest for reliable high-rate wireless links motivates fading-resilient and bandwidth-efficient communication systems capable of capitalizing on available forms of diversity. This tutorial focuses on linear complex field (LCF) coding—a powerful tool that complements the traditional Galois field (GF) coding, in enabling diversity transmissions over single- and multi-antenna fading channels. Performance and capacity analyses are provided, along with systematic guidelines for constructing LCF encoders, and options available to the designer for selecting LCF decoders. Emphasis is placed on high-performance and high-rate designs, and the emerging performance-rate-complexity trade-offs.

KEY WORDS
complex field coding
space-time coding
diversity
capacity
complexity

1. Introduction

Rapid increases in cellular service subscribers and wireless applications have stimulated tremendous research efforts in developing systems that support reliable high-rate transmissions over wireless channels. A major challenge in designing such high-performance high-rate systems is the mitigation of fading propagation effects, within the prescribed bandwidth and power limitations. To cope with the deleterious effects fading has on system performance, diversity-enriched transmitters and receivers have well-appreciated merits.

Diversity comes in different flavors: time, frequency, or space [1]. Modulation (a.k.a. signal space) diversity is a relatively recent flavor that was first

*Correspondence to: Georgios B. Giannakis, Department of Electrical and Computer Engineering, University of Minnesota, 200 Union Street, Minneapolis, MN 55455, USA.
†E-mail: georgios@ece.umn.edu
Contract/grant sponsor: NSF Wireless Initiative; contract/grant number: 9979443.
Contract/grant sponsor: NSF; contract/grant number: 0122431.
Contract/grant sponsor: ARL/CTA; contract/grant number: DAAD19-01-2-011.

combined with a component interleaver to enable *bandwidth-efficient* transmissions, with signal dimensionalities 2, 3, or 5, that gain resilience over flat Rayleigh fading channels [2]. Later on, constructing modulation-induced diversity transmissions was recast as a lattice code design problem over the *real* algebraic number field, and some good lattice constellations were designed for Rayleigh fading channels up to signal dimensionality 8 [3]. As the resulting real constellations in Reference [3] were not as good for additive white Gaussian noise (AWGN) channels, better complex constellations were developed in Reference [4]. Independently, lattice generator matrices with entries drawn from the complex field were advocated for their fading-resilient features in the signal processing for communications literature, in which the focus was placed on frequency-selective fading channels, and these matrices appeared under terms such as linear precoding [5–9], linear constellation precoding [10,11], or complex field coding [12]. As these matrices are linear operators, to unite the information and coding-theoretic approach with the signal-processing viewpoint, we will collectively call them *LCF encoders*.

A comprehensive treatment of LCF encoding combined with ideal (de-) interleaving for single-antenna transmissions over flat-fading channels can be found in Reference [4]. LCF coding for multi-antenna transmissions over multi-input multi-output (MIMO) channels was introduced in Reference [13], in which real space-time (ST) codes were designed *via* exhaustive computer search. Based on Reference [14], complex LCF ST codes for MIMO flat-fading channels were first designed in Reference [9] to ensure space diversity (see also Reference [15]), in which the sphere-decoding (SD) algorithm of Reference [16] was applied to the same problem). LCF-based space-time coding (STC) was studied thoroughly in References [10,11], in which LCF encoders were generalized to account not only for diversity but also for coding gains.

The present paper aspires to motivate, provide a concise exposition devoid of undue rigor, delineate pertinent trade-offs, and put forth novel directions for single- and multi-antenna wireless systems equipped with LCF coding. Starting from a simple but general matrix-vector model, we will first illuminate the effects of LCF coding on system performance in the presence of AWGN and fading effects. Both (non) Gaussian and correlated channels will be considered along with LCF decoding options. Our unifying performance analysis will guide our systematic code design methodology, which will enable full-diversity and large coding gains in fading channels, without sacrificing bandwidth or compromising performance in AWGN channels. Our results will encompass (even correlated) single-antenna and MIMO channels exhibiting flat, time-selective, or frequency-selective multipath fading. On top of average and outage performance, ergodic and outage capacity metrics will be used for comparing LCF-coded systems with existing alternatives. To considerably boost performance with only a moderate increase in receiver complexity, concatenated Galois-LCF encoding along with iterative (a.k.a. turbo) decoding schemes will be discussed briefly along the lines of Reference [17]. Emphasis will be placed on the role of LCF coding in facilitating reliable high-rate MIMO transmissions, leading to the recent design of Reference [18], which remarkably enjoys maximum rate and maximum diversity for any number of antennas. Finally, throughout the paper we will stress the pertinent trade-offs among rate, performance, and complexity.

Notation. Upper (lower) bold face letters will be used for matrices (column vectors). Superscript $^{\mathcal{H}}$ will denote Hermitian, * conjugate, ^{T} transpose, and † pseudoinverse. We will reserve $\otimes$ for the Kronecker product, and $E[\cdot]$ for expectation. We will use $tr(\boldsymbol{A})$ for the trace of the matrix $\boldsymbol{A}$; $\boldsymbol{I}_N$ will denote the $N \times N$ identity matrix, and $\boldsymbol{F}_N$ the $N \times N$ normalized (unitary) Fast Fourier Transform (FFT) matrix; diag[$\boldsymbol{x}$] will stand for a diagonal matrix with $\boldsymbol{x}$ on its main diagonal; $\mathbb{Z}(j)$ will denote the ring of Gaussian integers, whose elements are in the form of $p + jq$ with $p, q \in \mathbb{Z}$.

2. LCF Code Construction

Consider the simple but general matrix-vector system model shown in Figure 1. Each symbol block $\boldsymbol{s}$ of length N comprises independent and identically distributed Gray mapped information bits, with each entry of $\boldsymbol{s}$ adhering to a finite alphabet $\mathcal{A}_s$. The $N \times 1$ information block $\boldsymbol{s}$ is precoded by an $N \times N$ matrix $\boldsymbol{\Theta}$, which we term LCF encoder, as its entries are drawn from the complex field $\mathbb{C}$. Each LCF encoded block $\boldsymbol{u} = \boldsymbol{\Theta s}$ is transmitted through an equivalent diagonal matrix channel represented by $\boldsymbol{D}_h := \text{diag}[h_1, \ldots, h_N]$. Later on, we will see that this equivalent channel matrix appears in various system

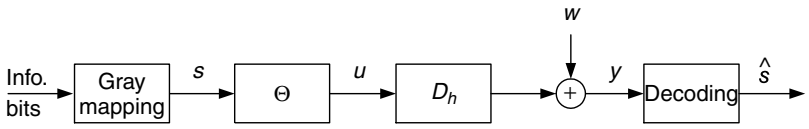

Fig. 1. An equivalent system model.

models such as time- or frequency-selective systems (Section 3) and multiple antenna systems (Section 4). The block $\boldsymbol{w}$ denotes a complex AWGN vector with mean zero and variance $\sigma^2 = N_0$.

At the receiver, we suppose that timing and carrier synchronization have been acquired. We sample the receive-filter output at the symbol rate and collect N samples into a vector $\boldsymbol{y}$, which obeys the following input–output relationship:

$$\boldsymbol{y} = \boldsymbol{D}_h\boldsymbol{u} + \boldsymbol{w} = \boldsymbol{D}_h\boldsymbol{\Theta}\boldsymbol{s} + \boldsymbol{w} \tag{1}$$

2.1. Encoder Design Criteria

Before discussing decoding options to obtain the detected block of symbols $\hat{\boldsymbol{s}}$, we will be concerned with the effect $\boldsymbol{\Theta}$ has on the system performance. Depending on the channel, $\boldsymbol{\Theta}$ will have different effects. In the ensuing sections, we will analyze these effects on AWGN channels and fading channels, respectively.

2.1.1. AWGN channels

For AWGN channels, the entries of $\boldsymbol{D}_h$ remain invariant from symbol to symbol. Without loss of generality, we set $\boldsymbol{D}_h = \boldsymbol{I}_N$. Since $\boldsymbol{w}$ is AWGN, the block pairwise error probability (PEP) is

$$P(\boldsymbol{s} \to \boldsymbol{s'}) = \frac{1}{2}\,\mathrm{erfc}\left(\frac{||\boldsymbol{u} - \boldsymbol{u'}||}{2\sqrt{N_0}}\right) \tag{2}$$

where $\boldsymbol{s} \neq \boldsymbol{s}' \in \mathcal{A}_s^{(N)}$, $\boldsymbol{u} = \boldsymbol{\Theta}\boldsymbol{s}$, $\boldsymbol{u}' = \boldsymbol{\Theta}\boldsymbol{s}'$, $\mathrm{erfc}(\cdot)$ is the complementary error function, $\mathrm{erfc}(x) = (2/\sqrt{\pi})\int_x^\infty e^{-t^2}\mathrm{d}t$, and $\mathcal{A}_s^{(N)}$ denotes the set of $N \times 1$ vectors $\boldsymbol{s}$ having their entries in $\mathcal{A}_s$. The union bound of the block error probability P_e is

$$\begin{aligned} P_\mathrm{e} &\leq \sum_{\forall \boldsymbol{s} \in \mathcal{A}_s^{(N)}} P(\boldsymbol{s}) \sum_{\boldsymbol{s}' \neq \boldsymbol{s}} P(\boldsymbol{s} \to \boldsymbol{s}') \\ &\leq \frac{M^N - 1}{2}\,\mathrm{erfc}\left(\frac{\delta_u}{2\sqrt{N_0}}\right) \end{aligned} \tag{3}$$

where $\delta_u^2 = \min_{\boldsymbol{u} \neq \boldsymbol{u}'} ||\boldsymbol{u} - \boldsymbol{u}'||^2$ is the minimum Euclidean distance between all $(\boldsymbol{u}, \boldsymbol{u}')$ pairs, and M is the cardinality of $\mathcal{A}_s$ (a.k.a. the constellation size). For fixed constellation size M and block length N, it is well known that P_e depends on the minimum Euclidean distance between any two possible blocks $\boldsymbol{u}$. To fix the power, we set $E[||\boldsymbol{u}||^2] = E[||\boldsymbol{s}||^2] = N\mathcal{E}_s$, where $\mathcal{E}_s := E[|s|^2]$. This means that $E[||\boldsymbol{u}||^2] = E[\boldsymbol{s}^{\mathcal{H}}\boldsymbol{\Theta}^{\mathcal{H}}\boldsymbol{\Theta}\boldsymbol{s}] = \mathrm{tr}(\boldsymbol{\Theta}^{\mathcal{H}}\boldsymbol{\Theta})\mathcal{E}_s = N\mathcal{E}_s$, with

$$\mathrm{tr}(\boldsymbol{\Theta}^{\mathcal{H}}\boldsymbol{\Theta}) = \mathrm{tr}(\boldsymbol{\Theta}\boldsymbol{\Theta}^{\mathcal{H}}) = N$$

Defining the minimum Euclidean distance between all (s, s') pairs in $\mathcal{A}_s$ as $\delta_s^2 := \min_{\forall s \neq s'} |s - s'|^2$, where $s \neq s' \in \mathcal{A}_s$, we establish the following property:

Property 1 *With minimum distance δ_s^2 between symbols, the Euclidean distance between symbol blocks obeys, under Equation (4), the following:*

(i) $\min_{\boldsymbol{s} \neq \boldsymbol{s}'} ||\boldsymbol{s} - \boldsymbol{s}'||^2 = \delta_s^2$; *and*
(ii) $\delta_u^2 \leq \delta_s^2$.

Proof: (i) Arguing by contradiction, suppose there exists one $(\boldsymbol{s}, \boldsymbol{s}')$ pair such that $0 < ||\boldsymbol{s} - \boldsymbol{s}'||^2 < \delta_s^2$, with $\boldsymbol{s} \neq \boldsymbol{s}'$. Then, there exists at least one element in $\boldsymbol{s} - \boldsymbol{s}'$, say the mth, which has norm $|s_m - s'_m| < \delta_s$. Since s_m and s'_m belong to $\mathcal{A}_s$, this contradicts the definition of δ_s. Hence, we have $\min_{\boldsymbol{s} \neq \boldsymbol{s}'} ||\boldsymbol{s} - \boldsymbol{s}'||^2 \geq \delta_s^2$. On the other hand, for a single error event, we have that $\boldsymbol{s} - \boldsymbol{s}'$ has only one nonzero entry. Supposing that the mth entry is nonzero, and $|s_m - s'_m|^2 = \delta_s^2$, we have that $\min_{\boldsymbol{s} \neq \boldsymbol{s}'} ||\boldsymbol{s} - \boldsymbol{s}'||^2 \leq |s_m - s'_m|^2 = \delta_s^2$. This completes the proof of (i).

(ii) From Equation (4), we obtain $\sum_{n=1}^N ||\boldsymbol{\theta}_n||^2 = N$, where $\boldsymbol{\theta}_n$ is the nth column of $\boldsymbol{\theta}$. Hence, there exists at least one column of $\boldsymbol{\Theta}$ such that $||\boldsymbol{\theta}_m||^2 \leq 1$. Selecting a pair $(\boldsymbol{s}_1, \boldsymbol{s}_2)$ with a single entry in error implies that $\boldsymbol{s}_1 - \boldsymbol{s}_2$ has only one nonzero entry at the mth location, and $||\boldsymbol{s}_1 - \boldsymbol{s}_2||^2 = \delta_s^2$. We can then verify that

$$\begin{aligned} \delta_u^2 &= \min_{\boldsymbol{u} \neq \boldsymbol{u}'} ||\boldsymbol{u} - \boldsymbol{u}'||^2 = \min_{\boldsymbol{s} \neq \boldsymbol{s}'} ||\boldsymbol{\Theta}(\boldsymbol{s} - \boldsymbol{s}')||^2 \\ &\leq ||\boldsymbol{\theta}_m||^2 \delta_s^2 \leq \delta_s^2 \end{aligned}$$

The inequality in Property 1 shows that the LCF encoder $\boldsymbol{\Theta}$ changes the Euclidean geometry of the original constellation. More importantly, the minimum Euclidean distance of the LCF-coded blocks $\boldsymbol{u}$ is no greater than that between the original information blocks $\boldsymbol{s}$. This implies that the performance in Equation (3) can only decrease if we do not choose $\boldsymbol{\Theta}$ carefully. Note that when $\boldsymbol{\Theta}^{\mathcal{H}}\boldsymbol{\Theta} = \boldsymbol{I}_N$; that is, $\boldsymbol{\Theta}$ is a unitary matrix, for each pair of $\boldsymbol{s}$ and $\boldsymbol{s}'$, we have $||\boldsymbol{u} - \boldsymbol{u}'|| = ||\boldsymbol{s} - \boldsymbol{s}'||$. Hence, the LCF encoder $\boldsymbol{\Theta}$ has no positive effect on performance when the channel is AWGN. In contrast, we will see that $\boldsymbol{\Theta}$ can have a positive impact when the wireless channel is fading.

2.1.2. Fading channels

To model fading effects, we view channels as random processes. For every time snapshot, each channel tap can be viewed as a random variable, which causes the error probability, P_{e}, to be also a random variable. To fully characterize a random variable, we need its probability density function (pdf). In general, it is difficult to find the pdf of P_{e} in closed form. This motivates studying the average, P_{av}, and the outage error probability, P_{out}.

Average error probability. We will first use a special pdf to introduce some definitions and later generalize to any pdf. Assume that the channel matrix $\boldsymbol{D}_h$ has independent diagonal entries and each entry h_n is a complex Gaussian distributed with zero mean and unit variance; that is, $h_n \sim \mathcal{CN}(0, 1)$, $\forall n \in [1, N]$. For each realization of $\boldsymbol{D}_h$, the PEP is denoted as $P(\boldsymbol{s} \to \boldsymbol{s}'|\boldsymbol{D}_h)$, with $\boldsymbol{s} \neq \boldsymbol{s}'$. The average PEP is defined as $P_{av}(\boldsymbol{s} \to \boldsymbol{s}') = E_h[P(\boldsymbol{s} \to \boldsymbol{s}'|\boldsymbol{D}_h)]$. Applying the Chernoff bound (see e.g. Reference [19], p. 717), we can upper-bound the average PEP as

$$P_{\mathrm{av}}(\boldsymbol{s} \to \boldsymbol{s}') \leq \prod_{u_n \neq u'_n} \frac{1}{1 + |u_n - u'_n|^2/4N_0} \tag{5}$$

where $u_n(u'_n)$ denotes the nth element of $\boldsymbol{u}(\boldsymbol{u}')$. Without loss of generality, we assume that $u_n \neq u'_n$, for $n = 1, \ldots, g_{\mathrm{d}}$, for some $g_{\mathrm{d}} \in [1, N]$. From Equation (5), we observe that unlike AWGN channels the performance now depends on g_{d}, the number of distinct entries between two blocks $\boldsymbol{u}$ and $\boldsymbol{u}'$. This corroborates the fact that it is the Hamming distance, not the Euclidean distance, that plays the more important role in performance when communicating over fading channels.

For compactness, let us define $g_{\mathrm{c}} = (\prod_{n=1}^{g_{\mathrm{d}}} |u_n - u'_n|^2)^{1/g_{\mathrm{d}}}$. At high signal-to-noise ratio (SNR) values, the union bound on the error probability becomes

$$P_{\mathrm{e}} \leq \sum_{g_{\mathrm{d}}} \sum_{g_{\mathrm{c}}} g_\kappa \left(g_{\mathrm{c}} \frac{1}{4N_0} \right)^{-g_{\mathrm{d}}} \tag{6}$$

where g_κ is an integer that equals the number of pairs with the same g_{d} and g_{c}.

Let us introduce the following definitions:

$$\begin{aligned} G_{\mathrm{d}} &= \min_{\boldsymbol{u} \neq \boldsymbol{u}'} g_{\mathrm{d}}, \\ G_{\mathrm{c}} &= \min_{\boldsymbol{u} \neq \boldsymbol{u}'} g_{\mathrm{c}}, \quad \text{when } g_{\mathrm{d}} = G_{\mathrm{d}}, \\ G_\kappa &= g_\kappa, \quad \text{when } g_{\mathrm{d}} = G_{\mathrm{d}}, \text{ and } g_{\mathrm{c}} = G_{\mathrm{c}} \end{aligned} \tag{7}$$

We call G_{d} the system diversity gain, G_{c} the system coding gain, and G_κ the kissing number. These three constants play important roles in parameterizing the system performance. At high SNR, we can rewrite Equation (6) as

$$P_{\mathrm{e}} \leq G_\kappa \left(G_{\mathrm{c}} \frac{1}{4N_0} \right)^{-G_{\mathrm{d}}} + o(*), \tag{8}$$

where $o(*)$ represents other terms from Equation (6), that can be ignored relative to the first term in the right-hand side of Equation (8). At high SNR, the $\log(P_{\mathrm{e}})$ *versus* SNR(dB) curve shows up as a straight line (c.f. Equation 8). Graphically, G_{d} decides the slope of the error probability curve in log-log P_{av} *versus* SNR figures, while G_{c} and G_κ determine the amount of shift along the SNR axis. Better performance means larger G_{d}, G_{c}, and smaller G_κ. To combat fading, enhancing the diversity gain is most important. Note that we defined G_{c} and G_κ in Equation (7) based on G_{d}. Therefore, when we want to compare different systems based on G_{c} and G_κ, we need to first make sure that they have the same diversity order. Since the kissing number G_κ depends on the constellation $\mathcal{A}_s$, we will henceforth focus on diversity and coding gains.

For each pair $(\boldsymbol{u}, \boldsymbol{u}')$, it follows by definition that the Hamming distance g_{d} equals the rank of $\mathrm{diag}[\boldsymbol{u} - \boldsymbol{u}'] = \mathrm{diag}[\boldsymbol{\Theta}(\boldsymbol{s} - \boldsymbol{s}')]$. Hence, G_{d} in Equation (7) depends on the LCF encoder $\boldsymbol{\Theta}$. Since the rank of a square matrix can never exceed its dimensionality ($G_{\mathrm{d}} \leq N$), the maximum-diversity gain $G_{\mathrm{d}}^{\max} = N$. To guarantee full diversity, we need to design $\boldsymbol{\Theta}$ such that $\forall \boldsymbol{s} \neq \boldsymbol{s}'$, the matrix $\mathrm{diag}[\boldsymbol{\Theta}(\boldsymbol{s} - \boldsymbol{s}')]$ has full rank. If the channel correlation matrix is $\boldsymbol{R}_h = E[\boldsymbol{h}\boldsymbol{h}^{\mathcal{H}}]$, where $\boldsymbol{h} := [h_1, \ldots, h_N]^T$, and the eigenvalue decomposition yields $\boldsymbol{R}_h = \boldsymbol{U}_h \boldsymbol{\Lambda}_h \boldsymbol{U}_h^{\mathcal{H}}$, where $\boldsymbol{U}_h$ is a unitary matrix and $\boldsymbol{\Lambda}_h$ is a diagonal matrix including all

nonzero eigenvalues of $\boldsymbol{R}_h$, then the diversity order for correlated channels is given as [11,12,20,21]

$$G_{\mathrm{d}} = \min_{\forall \boldsymbol{s} \neq \boldsymbol{s}'} \mathrm{rank}(\mathrm{diag}[\boldsymbol{\Theta}(\boldsymbol{s} - \boldsymbol{s}')]\boldsymbol{U}_h \boldsymbol{\Lambda}_h^{1/2}) \tag{9}$$

Therefore, the diversity order G_{d} also depends on the rank of the channel correlation matrix. In general, to achieve the maximum possible diversity N, we need to design both transmitter and receiver properly, and of course the channel should be able to provide at least diversity of order N.

For correlated channels, the coding gain is given by [11,22]

$$G_{\mathrm{c}} = \min_{\forall \boldsymbol{s} \neq \boldsymbol{s}'} \left\{ \det \left[\left(\mathrm{diag}[\boldsymbol{\Theta}(\boldsymbol{s} - \boldsymbol{s}')]\boldsymbol{U}_h \boldsymbol{\Lambda}_h^{1/2} \right)^{\mathcal{H}} \cdot \mathrm{diag}[\boldsymbol{\Theta}(\boldsymbol{s} - \boldsymbol{s}')]\boldsymbol{U}_h \boldsymbol{\Lambda}_h^{1/2} \right] \right\}^{1/G_{\mathrm{d}}} \tag{10}$$

which shows that the coding gain G_{c} also depends on the design of $\boldsymbol{\Theta}$. In the next section, we will use some examples to illustrate these dependencies.

So far, we have only considered fading channels with Rayleigh distributed taps. If we define $\sigma_n^2 := E[|h_n|^2]$ and $\beta_{\mathrm{n}} := |h_n|^2/\sigma_n^2$, we have that β_{n} is a nonnegative channel dependent random variable with unit mean. For a large class of channel pdfs, we can express the pdf of β_{n} near the origin as

$$p(\beta_{\mathrm{n}}) = a_n \beta_{\mathrm{n}}^{t_n} + o(\beta_{\mathrm{n}}^{t_n+\varepsilon}), \quad \forall n \in [1, N], \tag{11}$$

where $\varepsilon > 0$ and $t_n > -1$. We underscore that Equation (11) is assumed only near the origin. Considering the convergence of $p(\beta_{\mathrm{n}})$, t_n can be any real number that is greater than -1. On the basis of Equation (11), we can assess G_{d} and G_{c} for non-Gaussian fading channels as well. To support this claim, we borrow the following result:

Lemma 1 *[23, Proposition 1] Given the model*

$$y = hs + w \tag{12}$$

where $|h|^2 = \beta E[|h|^2]$, $p(\beta) = a\beta^t + o(\beta^{t+\varepsilon})$ *as in Equation (11), w is AWGN,* $s = \pm\sqrt{\mathcal{E}_s}$, *and* $\overline{\gamma} = \mathcal{E}_s E[|h|^2]/N_0$, *the average symbol error probability at high SNR depends only on the behavior of* $p(\beta)$ *at* $\beta \to 0^+$. *Specifically, at high SNR, the average symbol error probability is given by*

$$P_E = \frac{2^t a \Gamma(t + 3/2)}{\sqrt{\pi}(t+1)} \overline{\gamma}^{-(t+1)} + o\left(\overline{\gamma}^{-(t+1+\varepsilon)}\right) \tag{13}$$

where Gamma function is defined as $\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx$.

For our model in Equation (1), the pairwise probability is

$$P(\boldsymbol{u} \to \boldsymbol{u}'|\boldsymbol{D}_h) = \frac{1}{2}\,\mathrm{erfc}\left(\frac{||\boldsymbol{D}_h(\boldsymbol{u} - \boldsymbol{u}')||}{2\sqrt{N_0}} \right) = \frac{1}{2}\,\mathrm{erfc}\left(\frac{1}{2}\sqrt{\frac{\sum_{n=1}^{N} |h_n(u_n - u_n')|^2}{N_0}} \right)$$

Based on Equation (11), the independence of h_n, and the definition $g_{\mathrm{d}} = \sum_{u_n \neq u_n'} (t_n + 1)$, the pdf of the equivalent SNR $\gamma = \sum_{u_n \neq u_n'} |h_n(u_n - u_n')|^2/(4N_0)$ can be expressed as [23–25]

$$p(\gamma) = \left(\prod_{u_n \neq u_n'} \frac{a_n \Gamma(t_n + 1)}{(|u_n - u_n'|^2 \sigma_n^2)^{t_n+1}} \right) \frac{\gamma^{g_{\mathrm{d}}-1}}{\Gamma(g_{\mathrm{d}})} \cdot \left(\frac{1}{4N_0} \right)^{-g_{\mathrm{d}}} + o(\gamma^{g_{\mathrm{d}}-1+\varepsilon}) \tag{14}$$

From Equation (14) and Lemma 1, we can obtain the average pairwise probability for $\boldsymbol{u}$ and $\boldsymbol{u}'$ as

$$P(\boldsymbol{u} \to \boldsymbol{u}') = \left(\prod_{u_n \neq u_n'} \frac{a_n \Gamma(t_n + 1)}{(|u_n - u_n'|^2 \sigma_n^2)^{t_n+1}} \right) \cdot \frac{2^{g_{\mathrm{d}}-1} \Gamma(g_{\mathrm{d}} + 1/2)}{\sqrt{\pi}\Gamma(g_{\mathrm{d}} + 1)} \left(\frac{1}{4N_0} \right)^{-g_{\mathrm{d}}} + o\left(\left(\frac{1}{4N_0} \right)^{-(g_{\mathrm{d}}+\varepsilon)} \right) \tag{15}$$

Relying on Equation (15), we can establish the following proposition:

Proposition 1 *If all* h_n*'s are independent, and* h_n *has a pdf that is expressible near the origin as in Equation (11), then for any constellation, the diversity and coding gains for the average pairwise probability are*

$$G_{\mathrm{d}} = \min_{\boldsymbol{u} \neq \boldsymbol{u}'} g_{\mathrm{d}} = \min_{\boldsymbol{u} \neq \boldsymbol{u}'} \sum_{u_n \neq u_n'} (t_n + 1),$$

$$G_{\mathrm{c}} = \left(\frac{2^{G_{\mathrm{d}}-1} \Gamma(G_{\mathrm{d}} + 1/2)}{\sqrt{\pi}\Gamma(G_{\mathrm{d}} + 1)} \right)^{-1/G_{\mathrm{d}}} \cdot \min_{\boldsymbol{u} \neq \boldsymbol{u}'} \left(\prod_{u_n \neq u_n'} \frac{a_n \Gamma(t_n + 1)}{(|u_n - u_n'|^2 \sigma_n^2)^{t_n+1}} \right)^{-1/G_{\mathrm{d}}} \tag{16}$$

Notice that different channel taps h_n can have not only different SNRs but also different pdfs.

Outage probability. In addition to the average error probability, the outage probability P_{out} is another often used performance indicator. It is defined as [25]

$$P_{\text{out}} = P(0 \le \gamma \le \gamma_{\text{th}}) = \int_0^{\gamma_{\text{th}}} p(\gamma)\mathrm{d}\gamma \tag{17}$$

On the basis of Equation (14), we can verify that the outage error probability for pair $(\boldsymbol{u}, \boldsymbol{u}')$ is

$$P_{\text{out}} = \left(\prod_{u_n \neq u_n'} \frac{a_n \Gamma(t_n+1)}{(|u_n - u_n'|^2 \sigma_n^2)^{t_n+1}} \frac{\gamma_{\text{th}}^{g_{\text{d}}}}{\Gamma(g_{\text{d}}+1)} \right) \cdot \left(\frac{1}{4N_0} \right)^{-g_{\text{d}}} + o(*) \tag{18}$$

Similar to the average error probability, we can also parameterize the outage probability using the diversity and coding gains (see also Reference [23]) as

$$G_{\text{d}}' = \min_{\boldsymbol{u} \neq \boldsymbol{u}'} \sum_{u_n \neq u_n'} (t_n + 1)$$

$$G_{\text{c}}' = \min_{\boldsymbol{u} \neq \boldsymbol{u}'} \left(\prod_{u_n \neq u_n'} \frac{a_n \Gamma(t_n+1)}{(|u_n - u_n'|^2 \sigma_n^2)^{t_n+1}} \cdot \frac{\gamma_{\text{th}}^{G_{\text{d}}}}{\Gamma(G_{\text{d}}+1)} \right)^{-1/G_{\text{d}}} \tag{19}$$

Comparing Equation (19) with Equation (16), we deduce that the average error probability and the outage probability have the same diversity (slope in log–log curves). More important, both of them depend on the design of $\boldsymbol{\Theta}$ and the maximum diversity order that the underlying channel can provide. The difference between this general case and the Rayleigh case is that certain channel pdfs may provide diversity order more than one [23]. In the following, we will give an example to illustrate the theoretical analysis.

Example 1 (Nakagami-m) Suppose the block size is N and the channel taps have identical Nakagami-m pdfs

$$p(\beta_n) = \frac{m^m \beta^{m-1}}{\Gamma(m)} \exp(-m\beta), \quad n \in [1, N]$$

with variance $\sigma_n^2 = 1$ and $m \in \mathbb{R} > 0$. With respect to Equation (11), this pdf has $t_n = m - 1$ and $a_n = m^m/\Gamma(m)$ [23]. In this case, we can simplify the diversity and coding gains for the average pairwise probability in Equation (16) as

$$G_{\text{d}} = m \min_{\boldsymbol{u} \neq \boldsymbol{u}'} \operatorname{rank}(\operatorname{diag}(\boldsymbol{u} - \boldsymbol{u}')) \le mN$$

$$G_{\text{c}} = \frac{1}{m} \left(\frac{\Gamma(G_{\text{d}} + 1/2)}{2^{G_{\text{d}}+1}\sqrt{\pi}\Gamma(G_{\text{d}}+1)} \right)^{-1/G_{\text{d}}} \cdot \min_{\boldsymbol{u} \neq \boldsymbol{u}'} \left(\prod_{u_n \neq u_n'} \frac{1}{|u_n - u_n'|^2} \right)^{-m/G_{\text{d}}} \tag{20}$$

If $N = 1$, it follows that $G_{\text{d}} = m$. Since the parameter m can assume any positive real number, we obtain that the diversity order can be a noninteger real number and greater than 1, if $m > 1$.

If $m = 1$, the channel taps h_n are Rayleigh distributed. In this case, G_{d} reduces to what we have for the Gaussian case in Equation (9). The coding gain G_{c} in Equation (20) is more precise than the one given in Equation (10). Because the term before 'min' in the right-hand side of Equation (20) does not depend on $\boldsymbol{\Theta}$, using Equation (10) as a criterion for designing $\boldsymbol{\Theta}$ is equivalent to using Equation (20).

2.2. LCF Encoder Examples

In Section 2.1, we have seen that different channels (AWGN or fading) call for different $\boldsymbol{\Theta}$ designs to improve their performance. In this section, we will provide some examples to verify the claims in Section 2.1.

Example 2 With the system model of Figure 1, we select block size $N = 2$ and binary phase shift keying (BPSK) modulation. At the receiver, we perform maximum-likelihood (ML) decoding using exhaustive search. At the transmitter, we consider the following three choices for $\boldsymbol{\Theta}$:

$$\boldsymbol{\Theta}_1 = \boldsymbol{I}_2,$$

$$\boldsymbol{\Theta}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & e^{j\frac{\pi}{4}} \\ 1 & -e^{j\frac{\pi}{4}} \end{bmatrix},$$

$$\boldsymbol{\Theta}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & e^{j\frac{\pi}{4}} \\ 1 & e^{j\frac{\pi}{4}} \end{bmatrix} \tag{21}$$

When channels are AWGN, that is, $\boldsymbol{D}_h = \boldsymbol{I}_2$, the bit error rate (BER) curves are plotted in Figure 2 for the various choices of $\boldsymbol{\Theta}$. We observe from Figure 2

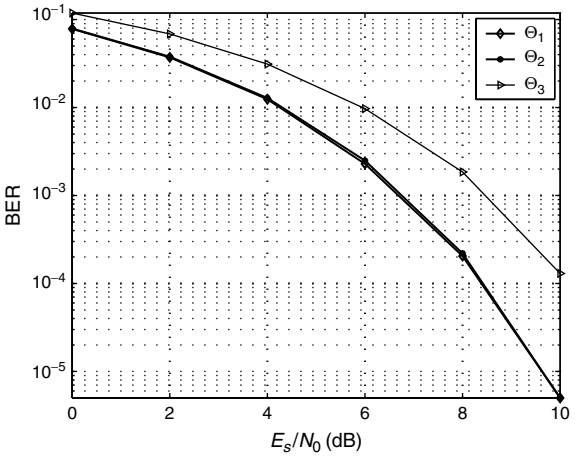

Fig. 2. LCF coding effects for AWGN channels.

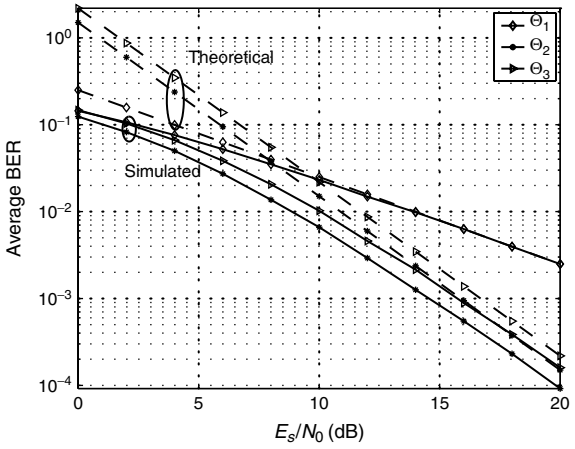

Fig. 3. LCF-coding effects for fading channels [solid lines are Monte Carlo simulation results, and dashed lines are theoretical ones based on Equation (16)].

that the systems using Θ_1 and Θ_2 exhibit the same performance, since Θ_1 and Θ_2 are both unitary matrices. The system with Θ_3 leads to the worst performance, since for Θ_3 the minimum Euclidean distance $\delta_u^2 = 4(2-\sqrt{2})$, which is less than the corresponding $\delta_u^2 = 4$ for Θ_1 and Θ_2. This verifies our theoretical analysis in Section 2.1.

We also apply the three LCF encoders in Equation (21) to fading channels. The results are depicted in Figure 3. We assume independent $h_0, h_1 \sim \mathcal{CN}(0,1)$. From the slopes of the BER curves, we infer that Θ_2 and Θ_3 enable higher diversity orders than Θ_1. Although the curves based on Θ_2 and Θ_3 exhibit the same slope at high SNR, they have different coding gains. In the same figure, the theoretical results from Equation (16) are plotted. We observe that the theoretical results are quite close to simulated ones.

This example demonstrates that an LCF encoder (e.g. Θ_3) which is good for fading channels may not be good for AWGN channels. Interestingly, we also notice that Θ_2 works well for both AWGN and fading channels.

Example 3 In Example 2, we have seen that good performance calls for a judiciously designed Θ. In this example, we will show that good Θ designs are also modulation-dependent. We select $N = 2$, Rayleigh independently fading channels, and test the following two unitary LCF encoders:

$$\Theta_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & e^{j\frac{\pi}{4}} \\ 1 & -e^{j\frac{\pi}{4}} \end{bmatrix}, \quad \Theta_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & j \\ 1 & -j \end{bmatrix} \tag{22}$$

Figure 4 depicts the BER curves for BPSK and quadrature phase shift keying (QPSK) modulations. We observe that with BPSK, Θ_2 outperforms Θ_1 by less than 0.2 dB. However, when QPSK modulation is used, Θ_1 enjoys higher diversity and outperforms Θ_2. In fact, Θ_2 does not unleash the maximum diversity order 2, when QPSK modulation is used. For example, if $\boldsymbol{s} = [1+j, 1+j]^T$ and $\boldsymbol{s}' = [-1+j, 1-j]^T$, then $\Theta_2(\boldsymbol{s}-\boldsymbol{s}') = [0, 2\sqrt{2}]^T$, which confirms that the diversity order G_{d} is only one.

2.3. General LCF Encoder Design

The following class of LCF encoders represented by Vandermonde matrices in $\mathbb{C}^{(N)}$ has been considered in [4]:

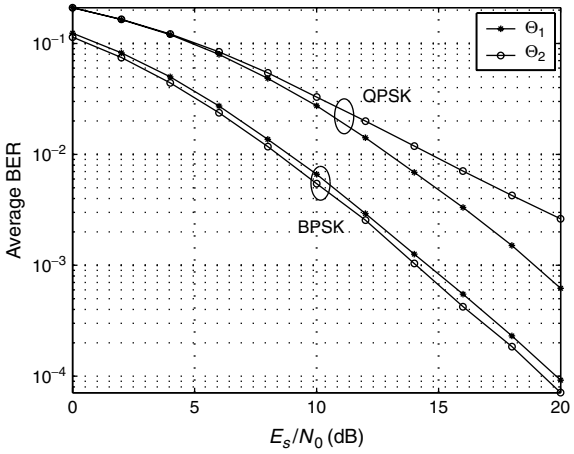

Fig. 4. LCF encoder effects on different modulations.

$$\Theta = \frac{1}{\lambda}\begin{bmatrix} 1 & \alpha_1 & \dots & \alpha_1^{N-1} \\ 1 & \alpha_2 & \dots & \alpha_2^{N-1} \\ \vdots & \vdots & & \vdots \\ 1 & \alpha_N & \dots & \alpha_N^{N-1} \end{bmatrix} \tag{23}$$

where the so-called generators $\{\alpha_n\}_{n=1}^N$ have unit modulus in $\mathbb{C}$ and $1/\lambda$ is a normalizing factor to enforce the condition in Equation (4). Recall from our discussion before Equation (9) that designing Θ judiciously amounts to selecting $\{\alpha_n\}_{n=1}^N$ so that for any $\boldsymbol{s} \neq \boldsymbol{s}'$, the vector $\Theta(\boldsymbol{s} - \boldsymbol{s}')$ has no zero entry.

A related criterion was introduced in Reference [2], in which Θ was designed to rotate the constellation such that a so-called modulation gain can be obtained. Later, References [3,4,26] generalized this design to obtain LCF-coded modulations adhering to higher dimensional lattices over both the fading and the AWGN channels. Recently, members of LCF class of encoders in Equation (23) have also been advocated for multi-antenna transmissions [10,11,13,27,28]. In most of the existing literature, the design of LCF encoders relies heavily on algebraic-theoretic tools. In an effort to bring these designs closer to practitioners, we will summarize the design of LCF encoders in relatively simpler terms.

Before we pursue the design of Θ though, we need to introduce the concept of Euler numbers and a couple of their properties (see also References [4,11]).

Euler numbers. Given an integer P, the Euler number $\phi(P)$ of P is the cardinality of the set $\{p : \gcd(p, P) = 1,\ p \in [1, P)\}$, where gcd stands for the greatest common divisor.

Two properties of Euler numbers are useful (see also References [4,11]):

E1) if P is prime, then $\phi(P^n) = P^{n-1}(P-1)$;
E2) if P, Q are relatively prime, that is, $\gcd(P, Q) = 1$, then $\phi(P \cdot Q) = \phi(P)\phi(Q)$.

Euler numbers are also important for LCF coding, as they are for GF coding. We will show that to design an LCF encoder the bottom line is to find an integer whose Euler number is a multiple of the block size N.

We can now summarize our LCF encoder design as follows:

General LCF encoder design. For a given block size N, select an integer multiple P of N such that $\phi(P) = 2mN$, where m is a positive integer. Then the generators $\{\alpha_n\}_{n=1}^N$ (and thus Θ in Equation (23)) can be designed as $\alpha_n = e^{j2\pi(P/N(n-1)+p)/P}$, where p is selected from $[1, P/N)$ such that $\gcd(P/N(n-1) + p, P) = 1, \forall n$.

There are some interesting ramifications from this general LCF encoder design

Proposition 2 *The following hold for the general LCF encoder:*

i. The LCF encoder Θ in Equation (23) exists for any size N;
ii. The normalizing factor of Θ is $\lambda = \sqrt{N}$;
iii. Θ is a unitary matrix and can be written as $\Theta = \boldsymbol{F}_N^{\mathcal{H}}\boldsymbol{D}_\alpha$, where $\boldsymbol{D}_\alpha = \mathrm{diag}[1, \alpha_1, \dots, \alpha_1^{N-1}]$;
iv. Θ enables full diversity N over quadrature amplitude modulated (QAM) or pulse amplitude modulated (PAM) constellations;
v. To maximize the coding gain, we need to select P so that m is as small as possible;
vi. The selection of p may not be unique.

The proofs of (i) to (vi) are similar to those in References [4,11], and we will skip them here. In the following, we will use some special cases to illustrate the LCF encoder design:

Special case 1. If $N = 2^k$, $k \in \mathbb{N}$, then we can select $P = 2^{k+2}$ such that $\phi(P) = 2N$ (c.f. property E1). There are two choices for p, namely, $p = 1$ or $p = 3$, with corresponding generators $\alpha_n = e^{j\pi(4n-3)/(2N)}$ or $\alpha_n = e^{j\pi(4n-1)/(2N)}$ for $n = 1, \dots, N$. This last choice coincides with that in References [4,11]. Since $m = 1$ in this case, the coding gain is also maximized (see Reference [11] for a proof).

Special case 2. If $N = 3{\cdot}2^k$, $k \in \mathbb{N} \cup \{0\}$, then we can select $P = 3^2{\cdot}2^{k+1}$, and the Euler number $\phi(P) = 2N$ (cf. property E2). We can choose $p = 1$ or 5 such that $\gcd(6(n-1) + p, P) = 1, \forall n$. Hence, $\alpha_n = e^{j2\pi(6n-5)/P}$ or $\alpha_n = e^{j2\pi(6n-1)/P}$. The latter selection of α_n coincides with that in Reference [4].

Special case 3. If $N = 5$, one choice of P is 50. In this case, $\phi(P) = 20 = 2 \cdot 2 \cdot N$. Selecting $p = 1$, the resulting generators are $\alpha_n = e^{j2\pi(10n-9)/P}$. Interestingly, this case had not been identified before.

The idea behind our general LCF encoder design can be explained as follows. The block $\boldsymbol{s}$ from

QAM(PAM) can be viewed as one symbol on an N-dimensional lattice, with $\boldsymbol{\Theta}$ mapping $\boldsymbol{s}$ to another N-dimensional lattice. If one then considers that $\overline{\boldsymbol{\theta}}_n^T\boldsymbol{s}$ is the nth coordinate of $\boldsymbol{\Theta s}$ projecting $\boldsymbol{s}$ on the nth row of $\boldsymbol{\Theta}$, then the properties of $\boldsymbol{\Theta}$ guarantee that $\boldsymbol{s}$ is mapped to a new constellation that has N unique coordinates. In other words, for all $\boldsymbol{s} \neq \boldsymbol{s}'$, we have $\overline{\boldsymbol{\theta}}_n^T(\boldsymbol{s} - \boldsymbol{s}') \neq 0$, $\forall n \in [1, N]$. Therefore, we obtain that diag$[\boldsymbol{\Theta}(\boldsymbol{s} - \boldsymbol{s}')]$ has full rank, $\forall \boldsymbol{s} \neq \boldsymbol{s}'$, which ensures full diversity.

2.4. LCF Decoding

To collect the full diversity enabled by judiciously designing $\boldsymbol{\Theta}$, proper LCF decoding is also necessary. The Maximum-Likelihood (ML) decoder for Equation (1) is given as

$$\hat{\boldsymbol{s}} = \min_{\boldsymbol{s} \in \mathcal{A}_s^{(N)}} ||\boldsymbol{y} - \boldsymbol{D}_h \boldsymbol{\Theta s}||^2. \tag{24}$$

When the ML decoder relies on exhaustive search, it has complexity $\mathcal{O}(M^N)$, which is prohibitively high for the large block size N and/or the constellation size M. In the following, we will consider three alternative decoders with complexity lower than ML.

2.4.1. Near-optimum decoding schemes

Sphere decoding. LCF encoder designs are based on lattices that are obtained by carving a space in $\mathbb{C}^{(N)}$ (complex field with dimension N). ML decoding of a lattice code is equivalent to finding the closest lattice point to the received point [16,29]. The sphere-decoding algorithm of References [16,29,30] was introduced to reduce decoding complexity, when the transmitted constellation is a subset of a lattice. The sphere-decoding algorithm searches through the lattice points inside a sphere of a certain radius and is centered at the received point, as shown in Figure 5.

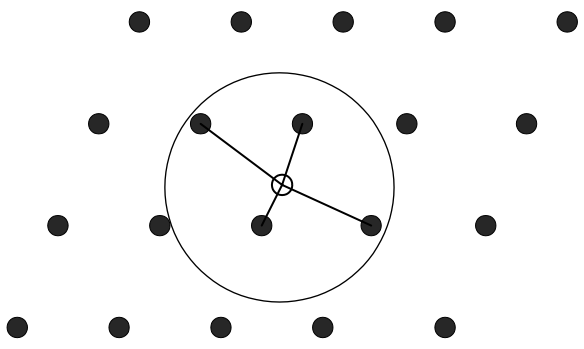

Fig. 5. Sphere decoding.

It has been shown that for a fixed search radius, and a given lattice structure, the decoding complexity for a real block of length N is approximately $\mathcal{O}(N^6)$ [15]. For complex-valued blocks, the complexity is approximately $\mathcal{O}((2N)^6)$. Recently, it has been purported that the *expected* decoding complexity is often $\mathcal{O}(N^3)$ [31]. Selecting the radius properly is not easy. Complex field decoding with further reduced complexity has been reported in References [32,33].

Semidefinite programming. Quasi-ML detection based on semidefinite programming is another promising near-optimal decoding option [34]. From Equation (24), we have that

$$\begin{aligned}\hat{\boldsymbol{s}} = \max_{\boldsymbol{s} \in \mathcal{A}_s^{(N)}} (&\boldsymbol{s}^{\mathcal{H}} \boldsymbol{\Theta}^{\mathcal{H}} \boldsymbol{D}_h^{\mathcal{H}} \boldsymbol{y} + \boldsymbol{y}^{\mathcal{H}} \boldsymbol{D}_h \boldsymbol{\Theta s} \\ &- \boldsymbol{s}^{\mathcal{H}} \boldsymbol{\Theta}^{\mathcal{H}} \boldsymbol{D}_h^{\mathcal{H}} \boldsymbol{D}_h \boldsymbol{\Theta s})\end{aligned} \tag{25}$$

Define $\boldsymbol{A} = \boldsymbol{D}_h\boldsymbol{\Theta}$, and

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{s} \\ c \end{bmatrix}, \quad \boldsymbol{Q} = \begin{bmatrix} -\boldsymbol{A}^{\mathcal{H}}\boldsymbol{A} & \boldsymbol{A}^{\mathcal{H}}\boldsymbol{y} \\ \boldsymbol{y}^{\mathcal{H}}\boldsymbol{A} & 0 \end{bmatrix}$$

where c is a known constant with $|c|^2 = \mathcal{E}_s$. Equation (25) can be rewritten as

$$\hat{\boldsymbol{s}} = \max_{\boldsymbol{x} \in \mathcal{A}_s^{(N+1)}} \boldsymbol{x}^{\mathcal{H}} \boldsymbol{Q} \boldsymbol{x}.$$

If s is a constant modulus, this integer maximization problem can be solved with the semidefinite relaxation (SDR) algorithm outlined in Reference [34, Table I]. The complexity of SDR decoding is $\mathcal{O}(N^{3.5})$.

2.4.2. Suboptimal decoding schemes

When the channel is rich in diversity gains, we can further reduce the complexity by giving up some performance. Nulling-cancelling (NC) [35] and the block minimum mean-square error (MMSE) decision feedback equalization (DFE) algorithm [36,37], both have lower complexity (roughly $\mathcal{O}(N^3)$) but worse performance than ML or near-ML decoders. Both NC and MMSE-DFE are decoding schemes that are based on decision feedback. The decision feedback here is only used for interference cancellation, but one could also use it for channel estimation in a decision-directed mode.

2.4.3. Linear decoding schemes

Two common linear decoding schemes are block zero-forcing (ZF) and MMSE equalizers:

$$\hat{\boldsymbol{s}}_{\text{zf}} = (\boldsymbol{D}_h\boldsymbol{\Theta})^{\dagger}\boldsymbol{y}$$
$$\hat{\boldsymbol{s}}_{\text{mmse}} = \boldsymbol{R}_s(\boldsymbol{D}_h\boldsymbol{\Theta})^{\mathcal{H}}(\sigma_w^2\boldsymbol{I}_N + \boldsymbol{D}_h\boldsymbol{\Theta}\boldsymbol{R}_s(\boldsymbol{D}_h\boldsymbol{\Theta})^{\mathcal{H}})^{-1}\boldsymbol{y} \tag{26}$$

where † denotes pseudoinverse and $\boldsymbol{R}_s := E[\boldsymbol{s}\boldsymbol{s}^{\mathcal{H}}]$ is the autocorrelation matrix of $\boldsymbol{s}$.

As a closing comment for all decoding options, we point out the well-known trade-off between performance and complexity. When the maximum diversity order is large, we may use linear decoders to reduce decoding complexity. However, sometimes this kind of trade-off is not easy to delineate analytically, since it is difficult to evaluate the performance of suboptimal and linear schemes in closed form. There is a considerable performance gap between ML decoding and linear decoding. Later on, we will show how a 'continuously controllable' trade-off is possible via a flexible LCF encoding that relies on variable-size block partitioning.

3. Single-Antenna Systems

In Section 2, we have established that LCF encoding is useful for enabling diversity gains over fading channels. In the limit (as $N \to \infty$), the $N \times N$ encoder $\boldsymbol{\Theta}$ converts a fading channel to an AWGN channel if the channel provides enough diversity [17]. On this limiting AWGN channel, GF coding has well-documented merits for improving its BER performance. But when used alone, GF coding may not offer equally efficient improvement over fading channels. In the following, we will advocate combining GF and LCF coding, in an effort to benefit from their complementary strengths in AWGN and fading channels, respectively.

3.1. Joint LCF and GF Coding for Fast Fading Channels

With reference to Figure 6, we first consider flat-fading channels. At the transmitter, GF coding is used as an outer code. In the literature, most GF codes are designed for AWGN channels. It is well known that a code that is good for AWGN may not be good for fading channels. The LCF encoder serves as an inner coder to turn the equivalent fading channel 'closer' to an AWGN channel. Another motivation for concatenating GF with LCF coding is provided through the required Hamming distance: relying on GF coding alone to improve system performance over fading channels implies a commensurate increase in the Hamming distance of the GF code with the possible decrease in transmission rate. However, we have shown that LCF encoding enables high diversity without sacrificing bandwidth efficiency. On the other hand, if we use LCF encoding alone, the price paid is twofold. First, to collect full-diversity gains, ML (or near ML) decoding is needed, which comes with high computational complexity that increases as the block size N increases. Second, the performance when using LCF coding only for fading channels is bounded by that of an uncoded transmission over an AWGN channel. For all these reasons, we are motivated to consider the combination of GF and LCF coding.

3.1.1. Encoder design

When GF coding is used together with LCF coding, two interleavers must be employed. One interleaver Π_2 is needed to decorrelate the channel symbols entering the LCF decoder. The other interleaver Π_1 is used to guarantee that code words enter different LCF encoded blocks [17,38]. Thanks to the ideal interleaver Π_1, $d_{\min}$ coded bits are mapped to $d_{\min}$ different symbols, and then these $d_{\min}$ symbols are assigned to $d_{\min}$ LCF encoded blocks. The interleaver Π_2 renders the channel taps (at least approximately) independent from symbol to symbol.

It has been shown that with a small size ($N \leq 4$) $\boldsymbol{\Theta}$, we can collect most of the gains provided by LCF coding [17]. If the interleavers (Π_1 and Π_2) are ideal, the maximum achievable diversity by joint GF-LCF coding is

$$G_{\text{d}}^{\max} = d_{\min}N, \tag{27}$$

where $d_{\min}$ denotes the minimum Hamming distance of the GF code. This multiplicative diversity

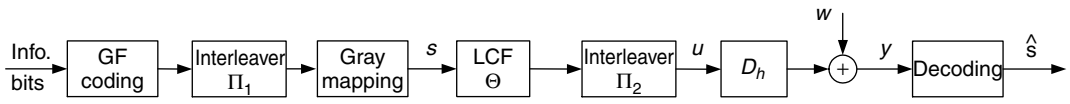

Fig. 6. System model for joint GF-LCF coding.

in Equation (27) shows that joint GF-LCF coding achieves better performance than either GF or LCF coding used alone [17].

For illustration purposes, we plot in Figure 7 the Chernoff bounds for the performance of joint GF-LCF coding with different GF codes and variable-size LCF encoders. As GF codes, we select rate $1/2$ and $3/4$ convolutional codes (CC). The free distances for these two codes are $d_{\min} = 10$ and 5, respectively. There are four curves in Figure 7. The dashed line depicts the performance with a rate-$1/2$ CC only. Rate-$1/2$ CC outperforms rate-$3/4$ CC (solid line), as it achieves a much higher diversity order. Combining the LCF encoder of Size 2 with a rate-$3/4$ CC brings the performance of joint GF-LCF coding quite close to that of rate-$1/2$ CC. They have identical diversity orders. Furthermore, if we select $\boldsymbol{\Theta}$ to have size $N = 4$, GF-LCF coding outperforms the rate-$1/2$ CC with rate $3/4$. Compared with the scheme without LCF coding, joint GF-LCF decoding incurs moderately higher complexity. In the following sections, we will outline decoding options for this joint GF-LCF coding [17].

3.1.2. Turbo decoding

Exhaustive search over all possible input sequences certainly offers the exact ML decoding for our joint GF-LCF-coded transmissions. However, presence of the interleaver Π_1 prevents the application of near-optimal decoding. It has been widely demonstrated that iterative (turbo) decoding is effective in dealing with such joint ML detection problems. Recently, turbo decoding of joint GF-LCF-coded transmissions has been derived in Reference [17], to which the reader is referred for detailed performance and complexity analyses.

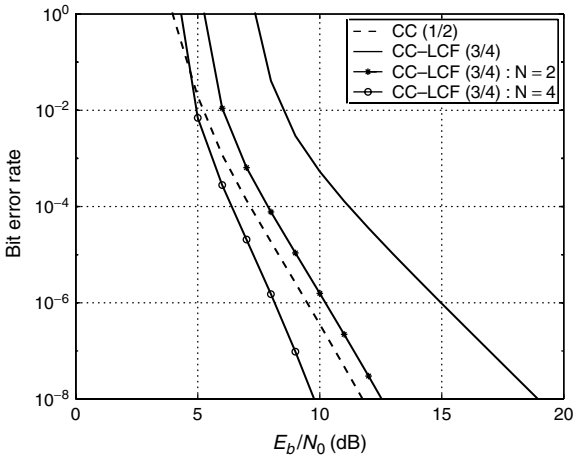

Fig. 7. Comparisons between convolutional coded systems.

3.2. LCF Coding for Frequency-selective Channels

It has been well documented that frequency-selective channels provide multipath diversity [12]. Orthogonal frequency division multiplexing (OFDM) converts a frequency-selective fading channel into a set of parallel flat-fading subchannels. Therefore, OFDM significantly reduces complexity both at the equalization and the decoding stages. However, uncoded OFDM's performance suffers from loss of diversity [12]. In the following, we will show how LCF-coded OFDM restores the full multipath diversity and offers large coding gains.

Instead of LCF coding, GF-Coded OFDM (C-OFDM) [39,40] has so far been adopted for the same purpose by many standards [41]. We will not elaborate on the analysis C-OFDM because:

1. it is difficult to optimize trellis or convolutional codes over fading channels for any channel order L;
2. improving performance requires a commensurate increase in the minimum Hamming distance, which often comes at the price of bandwidth loss and/or decoding complexity;
3. as we mentioned earlier, the combination of GF and LCF coding turns out to provide an attractive low-complexity alternative with better performance.

3.2.1. LCFC-OFDM

The underlying frequency-selective channel is modeled as an Lth-order finite impulse response (FIR) filter, denoted by $\boldsymbol{h} := [h(0), \ldots, h(L)]^T$, with $h(l)$ standing for the lth tap. After inserting a cyclic prefix of size L in every transmitted block and removing it from each received block, the interblock interference is eliminated and the equivalent channel matrix becomes a circulant matrix $\tilde{\boldsymbol{H}}$ with first row $[h(0), 0, \ldots, 0, h(L), \ldots, h(1)]$ (see Reference [8] for details). Performing N-point Inverse Fast Fourier Transform (IFFT) at the transmitter (postmultiplication of $\tilde{\boldsymbol{H}}$ by the IFFT matrix $\boldsymbol{F}^{\mathcal{H}}$) and FFT at the receiver (premultiplication of $\tilde{\boldsymbol{H}}$ by $\boldsymbol{F}$), OFDM yields a diagonal equivalent channel matrix $\boldsymbol{D}_h := \text{diag}[\tilde{h}_1, \ldots, \tilde{h}_N]$, where $\tilde{h}_n = \sum_{l=0}^{L} h(l) e^{-j2\pi l(n-1)/N}$

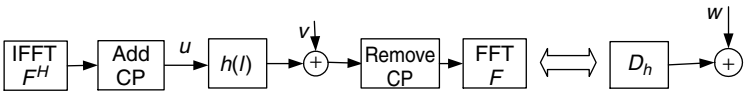

Fig. 8. Equivalent channel model.

is the frequency response of the underlying FIR channel evaluated at the FFT grid (see Figure 8). Now, the equivalent model is the same as that in Figure 1:

$$\boldsymbol{y} = \boldsymbol{D}_h \boldsymbol{\Theta} \boldsymbol{s} + \boldsymbol{w}$$

which shows that with the use of OFDM, our LCF coding design (including LCF-GF combinations) developed for flat-fading channels carries over to frequency-selective fading channels too. However, we still need to keep block sizes (equal to the number of subcarriers) small to ensure low decoding complexity. Alternatively, we can have large size blocks, but follow the 'divide-and-conquer' approach that we describe next.

3.2.2. Grouping with LCFC

Let us suppose that the channel taps $\{h(l)\}_{l=0}^{L}$ are complex Gaussian random variables with zero mean and correlation matrix $\boldsymbol{R}_h = E[\boldsymbol{h}\boldsymbol{h}^{\mathcal{H}}]$. When $\boldsymbol{\Theta} = \boldsymbol{I}_N$, it is easy to verify that (LCFC) OFDM has diversity order $G_{\mathrm{d}} = 1$. Note that the maximum-diversity order is $G_{\mathrm{d}}^{\max} = \mathrm{rank}(\boldsymbol{R}_h) \leq L + 1$ for any block of size $N > L$. To enable the full diversity, one direct approach here is to design $\boldsymbol{\Theta}$ with size $N \times N$ according to Section 2.3. However, this approach has two drawbacks: i) although it guarantees full diversity, it does not necessarily guarantee a decent coding gain; ii) the SD algorithm is computationally prohibitive when N is large. Viterbi's algorithm (VA) can be used in this case with extra L redundant symbols per block as discussed in Reference [42]. Later, we will describe designs that provide more flexibility than the VA in achieving desirable performance-complexity trade-offs.

Suppose we split the N subcarriers into N_g groups and select each group to have size N_{sub} [8,43]. Then, the input–output relationship for the gth group is

$$\boldsymbol{y}_g = \boldsymbol{D}_g \boldsymbol{\Theta}_{\mathrm{sub}} \boldsymbol{s}_g + \boldsymbol{w}_g \tag{28}$$

where $\boldsymbol{D}_g := \mathrm{diag}[\tilde{h}_{g_1}, \ldots, \tilde{h}_{g_{N_{\mathrm{sub}}}}]$. On the basis of the definition of $\tilde{h}_n$, we have

$$\tilde{\boldsymbol{h}}_g := \begin{bmatrix} \tilde{h}_{g_1} \\ \vdots \\ \tilde{h}_{g_{N_{\mathrm{sub}}}} \end{bmatrix} = \sqrt{N} \boldsymbol{F}_g \begin{bmatrix} h(0) \\ \vdots \\ h(L) \end{bmatrix} \tag{29}$$

where $\boldsymbol{F}_g$ denotes the $\{n_l\}_{l=1}^{N_{\mathrm{sub}}}$ rows of the first $L+1$ columns drawn from the N-point FFT matrix $\boldsymbol{F}$. If $\boldsymbol{R}_h = E[\tilde{\boldsymbol{h}}\tilde{\boldsymbol{h}}^{\mathcal{H}}]$ has full rank and $N_{\mathrm{sub}} \leq L + 1$, then $E[\tilde{\boldsymbol{h}}_g \tilde{\boldsymbol{h}}_g^{\mathcal{H}}]$ has full rank. For each group, we can design an LCF encoder $\boldsymbol{\Theta}_{\mathrm{sub}}$ with size $N_{\mathrm{sub}} \times N_{\mathrm{sub}}$ to enable diversity gains. If the SD algorithm is applied per group, the diversity order for each group is

$$G_{\mathrm{d}} = \min(N_{\mathrm{sub}}, L + 1) \tag{30}$$

and the decoding complexity depends on N_{sub}, which can be chosen to be much smaller than N. Now the trade-off between diversity (performance) and complexity has been delineated (cf. Equation (30)). If $N_{\mathrm{sub}} \leq L + 1$, as the group size increases from N_{sub} to $N_{\mathrm{sub}} + 1$, the decoding complexity increases from $\mathcal{O}((N_{\mathrm{sub}})^3)$ to $\mathcal{O}((N_{\mathrm{sub}} + 1)^3)$. Properly designed $\boldsymbol{\Theta}$ can span the entire diversity range from 1 to $L + 1$. The role of N_{sub} as a tuning parameter is important when the channel order L is large. If $N_{\mathrm{sub}} > L + 1$, no extra diversity order is provided, but the decoding complexity still increases with N_{sub}.

When $N_{\mathrm{sub}} = L + 1$, full diversity is enabled. The coding gain when $N_{\mathrm{sub}} = L + 1$ is (cf. Equation (10))

$$G_{\mathrm{c}} = \min_{\boldsymbol{s}_g \neq \boldsymbol{s}_g'} \left[\det(\boldsymbol{D}_{\mathrm{e}} \boldsymbol{D}_{\mathrm{e}}^{\mathcal{H}}) \det(\boldsymbol{R}_h) \det(\boldsymbol{F}_g^{\mathcal{H}} \boldsymbol{F}_g) \right]^{1/G_{\mathrm{d}}} \tag{31}$$

where $\boldsymbol{D}_{\mathrm{e}} := \mathrm{diag}[\boldsymbol{\Theta}_{\mathrm{sub}}(\boldsymbol{s}_g - \boldsymbol{s}_g')]$. The coding gain in Equation (31) depends not only on $\boldsymbol{\Theta}_{\mathrm{sub}}$ but also on the subcarrier placement (via the $\det(\boldsymbol{F}_g^{\mathcal{H}} \boldsymbol{F}_g)$), which is different from what we have seen for flat-fading channels. To maximize G_{c}, we need to maximize $\det(\boldsymbol{F}_g^{\mathcal{H}} \boldsymbol{F}_g)$. The latter is achieved if the groups are selected as (see also Reference [43])

$$\boldsymbol{s}_g = [s_g, s_{g+N_g}, \ldots, s_{g+(N_{\mathrm{sub}}-1)N_g}]^T \tag{32}$$

This equi-spaced grouping method can be implemented via a block interleaver with depth N_{sub}, as shown in Figure 9.

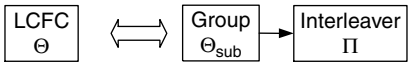

Fig. 9. Equivalent LCF encoder.

The equivalent LCF encoder with grouping can also be expressed as

$$\Theta = \begin{bmatrix} \boldsymbol{I}_{N_g} \otimes \overline{\boldsymbol{\theta}}_1^T \\ \vdots \\ \boldsymbol{I}_{N_g} \otimes \overline{\boldsymbol{\theta}}_{N_{\text{sub}}}^T \end{bmatrix} \tag{33}$$

where $\overline{\boldsymbol{\theta}}_l^T$ is the lth row of Θ_{sub}. The bottom line here is related to the influence interleaving has on coding schemes: the interleaver distributes the LCF (GF)-coded symbols to 'less correlated' channels.

This grouped LCFC-OFDM scheme guarantees full-diversity and large coding gains with low decoding complexity that depends on the subblock length N_{sub}. More importantly, it offers (by tuning the parameter N_{sub}) flexibility to trade-off performance for reduced decoding complexity.

Example 4 (performance-complexity trade-off) Here, we select $L = 7$ and $N = 48$. The channel taps $\{h(l)\}_{l=0}^{L}$ are independent zero-mean Gaussian distributed with variance satisfying an exponential power profile. Figure 10 depicts the average BER curves with variable-size N_{sub}. We observe that as N_{sub} increases, the performance increases, along with the decoding complexity.

This grouped LCF encoding approach can also be used with time-selective channels and/or MIMO systems [22,44]. It can certainly be combined with GF coding to further improve the performance, while keeping the receiver complexity manageable by employing turbo decoding as described in Reference [17]. Furthermore, training-based channel estimation blends well with the grouped LCF encoder as discussed in Reference [22].

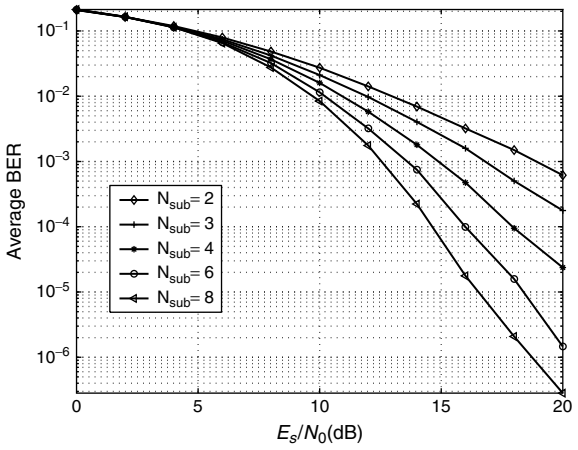

Fig. 10. BER performance with different size Θ_{sub}.

Having outlined LCF coding for single-antenna systems, we now turn our attention to space-time multi-antenna systems.

4. LCF Space-time Coding

Space-time coding (STC) is an effective means of gaining space diversity and thereby improving performance. At the same time, multi-antenna systems can provide multiplexing gain to boost capacity and increase transmission rate. Typical examples of STC include ST trellis codes [20] and ST block codes from orthogonal designs [45]. We call the STC schemes in, for example, [20,45] 'performance-oriented' ST codes, because these codes aim at collecting full space diversity with rate at most 1 symbol per channel use (pcu). Another category of ST schemes is designed to enable multiplexing gain that results in high transmission rate [35,46]. We call this category of codes 'rate-oriented'. In this section, we will design both 'performance-oriented' as well as 'rate-oriented LCF codes suitable for multi-antenna systems.

Consider a wireless system with N_{t} transmit- and N_{r} receive-antennas, as shown in Figure 11. At the transmitter, the $N \times 1$ information-bearing block $\boldsymbol{s}$ is mapped to the N_{t} substreams $\{\boldsymbol{c}_\mu\}_{\mu=1}^{N_{\text{t}}}$. Each block $\boldsymbol{c}_\mu$ per substream has length P. For a constellation $\mathcal{A}_s$ with size M, the transmission rate is therefore

$$R = \frac{N}{P} \log_2 M \text{ bits pcu} \tag{34}$$

Because from each antenna we are not going to transmit more than 1 symbol pcu, we have that

$$0 < R \leq N_{\text{t}} \log_2 M \text{ bits pcu} \tag{35}$$

We will consider ST codes with different rates R, but first, we will introduce 'performance-oriented' LCF-based ST codes with maximum $R = 1$ symbol pcu.

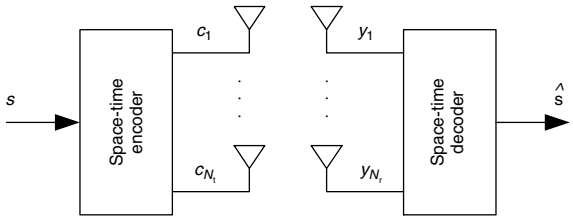

Fig. 11. MIMO model.

4.1. Performance-oriented LCF-STC

Although LCF codes were originally designed for single-antenna systems (see e.g. References [3,4,26]), recently, they have been applied also to multi-antenna systems [10,11,13,27]. For such MIMO channels, the LCF-ST encoder is designed as

$$\boldsymbol{C} = \left[\boldsymbol{c}_1 \cdots \boldsymbol{c}_{N_t}\right] = \mathrm{diag}[\boldsymbol{\Theta s}] \tag{36}$$

that is, the information block $\boldsymbol{s}$ is first LCF-coded by an $N_t \times N_t$ matrix $\boldsymbol{\Theta}$ and then the coded block $\boldsymbol{\Theta s}$ is transmitted via the N_t antennas in an *antenna-switching* fashion. Since $\boldsymbol{\Theta}$ is a square matrix, and the total number of transmitting time slots is $P = N_t$, the rate of LCF-STC is $R = 1$ symbol pcu. We referred to this LCF-STC scheme as antenna-switching, because in any time slot only one antenna is active; certainly, N_t antennas are used for N_t time slots. One may argue that the multiple antennas are not used efficiently here: one could envision having all multiple antennas transmit simultaneously, and thus increase the transmission rate. We will discuss this issue in detail in the next section.

But for now suppose that the $N_t N_r$ channels are Rayleigh flat-fading; hence, the $h_{\mu,\nu}$ channel tap from the μth transmitter to the νth receiver is a zero-mean complex Gaussian random variable. We assume that the $N_t N_r \times N_t N_r$ channel correlation matrix $\boldsymbol{R}_h = E[\boldsymbol{h}\boldsymbol{h}^{\mathcal{H}}]$ has full rank, where $\boldsymbol{h} := [h_{1,1}, \ldots, h_{N_t,N_r}]^T$. Channel coefficients $h_{\mu,\nu}$ are invariant during N_t time slots and are known to the receiver, but unknown to the transmitter.

Let $\boldsymbol{y}_\nu$ denote the received block at the νth receive-antenna; let $\boldsymbol{H}$ stand for the $N_t \times N_r$ channel matrix with (μ, ν)th entry $h_{\mu,\nu}$; and $\boldsymbol{W}$ be the AWGN matrix. The input–output relationship for this multi-antenna system is

$$[\boldsymbol{y}_1 \cdots \boldsymbol{y}_{N_r}] = \boldsymbol{CH} + \boldsymbol{W} \tag{37}$$

On the basis of Equation (36), and the property of interchanging the diagonal matrix with a vector, we can rewrite Equation (37) as

$$[\boldsymbol{y}_1 \cdots \boldsymbol{y}_{N_r}] = [\boldsymbol{D}_{h_1}\boldsymbol{\Theta s} \cdots \boldsymbol{D}_{h_{N_r}}\boldsymbol{\Theta s}] + \boldsymbol{W} \tag{38}$$

where $\boldsymbol{D}_{h_\nu} = \mathrm{diag}[h_{1,\nu}, \ldots, h_{N_t,\nu}]$. Note that for each receive-antenna, we have

$$\boldsymbol{y}_\nu = \boldsymbol{D}_{h_\nu}\boldsymbol{\Theta s} + \boldsymbol{w}_\nu, \quad \forall \nu \in [1, N_r] \tag{39}$$

Recalling Figure 1, the model in Equation (1) is the same as our simple equivalent model in Equation (1). Hence, we know that by judiciously designing $\boldsymbol{\Theta}$, full transmit diversity and large coding gains can be guaranteed. Since all the blocks $\boldsymbol{y}_\nu$ contain the same information block $\boldsymbol{s}$, we can use maximum ratio combining (MRC) at the receiver to collect receive diversity. As each transmit-antenna provides transmit diversity of order N_t, and MRC collects diversity from all N_r receive-antennas, it is intuitively reasonable to expect that LCF-STC enables diversity of order $N_t N_r$. The detailed proof establishing the diversity ($N_t N_r$) and coding gains of LCF-STC can be found in Reference [11].

Compared to ST trellis codes [20], LCF-STC can afford low design complexity for any N_t. Specifically, instead of exhaustive computer search required for designing good ST trellis codes for some N_t, LCF-STC offers a systematic design that ensures full-diversity and large coding gains, $\forall N_t$. Compared to ST orthogonal codes (ST-OD) [45], LCF-STC has higher rate. Thorough performance comparisons between LCF-STC and ST-OD can be found in Reference [11]. Here, we will highlight their rate differences by analyzing the corresponding mutual information.

When there is no channel knowledge at the transmitter, to maximize the mutual information through our MIMO channel, the multi-antenna transmitter should transmit white Gaussian distributed symbol blocks with correlation matrix, $\boldsymbol{R}_s = E[\boldsymbol{s}\boldsymbol{s}^{\mathcal{H}}] = \mathcal{E}_s \boldsymbol{I}_N$ [47]. In this case, the capacity for LCF-STC and ST-OD is given, respectively, by [11,48,49]

$$\mathcal{C}_{\mathrm{CF}} = \frac{1}{N_t}\sum_{\mu=1}^{N_t} \log_2\left(1 + \overline{\gamma}\sum_{\nu=1}^{N_r}|h_{\mu,\nu}|^2\right) \quad \text{bits pcu}$$

$$\mathcal{C}_{\mathrm{OD}} = R_{\mathrm{OD}} \log_2\left(1 + \frac{\overline{\gamma}}{N_t}\sum_{\mu=1}^{N_t}\sum_{\nu=1}^{N_r}|h_{\mu,\nu}|^2\right) \quad \text{bits pcu} \tag{40}$$

where R_{OD} denotes the rate for ST-OD and $\overline{\gamma} = \mathcal{E}_s/N_0$ denotes the average SNR. In comparing the capacity of these two schemes, we distinguish between the following two cases:

Case 1: $N_t = 2$ and $N_r \geq 1$, with $R_{\mathrm{OD}} = 1$

For the $\mathcal{C}_{\mathrm{OD}}$ and $\mathcal{C}_{\mathrm{CF}}$ capacities in Equation (40), we will show that

$$\mathcal{C}_{\mathrm{CF}} \leq \mathcal{C}_{\mathrm{OD}} \tag{41}$$

To prove Equation (41), let us first define $|h_\mu|^2 = \sum_{\nu=1}^{N_r}|h_{\mu,\nu}|^2$. Since $\log_2(\cdot)$ is a monotonic function,

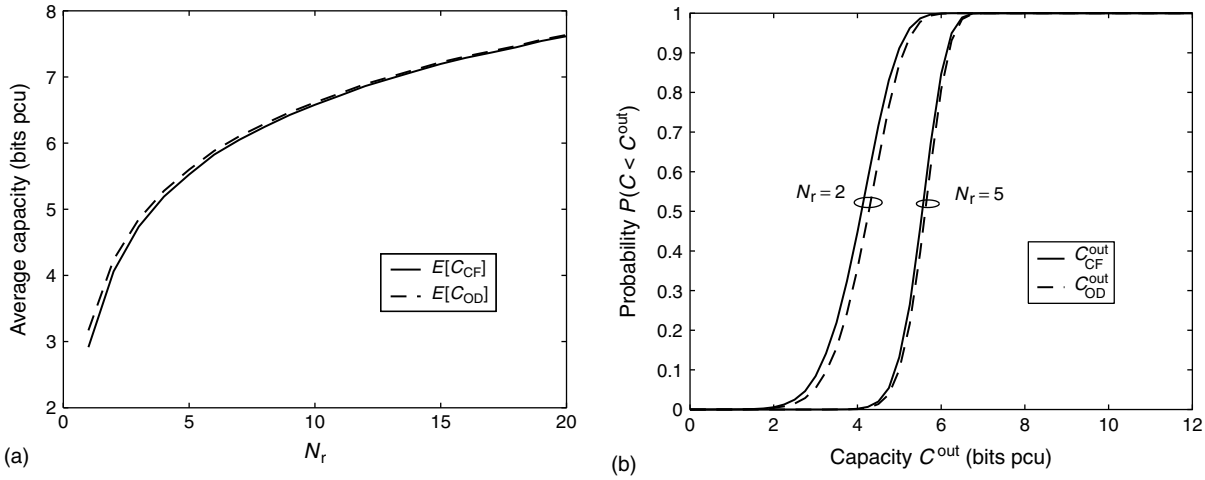

Fig. 12. Comparisons of $\mathcal{C}_{CF}$ and $\mathcal{C}_{OD}$ when $N_t = 2$.

proving Equation (41) is equivalent to showing that

$$\left(\prod_{\mu=1}^{N_t}(1+\overline{\gamma}|h_\mu|^2)\right)^{\frac{1}{N_t}} \leq 1+\frac{\overline{\gamma}}{N_t}\sum_{\mu=1}^{N_t}|h_\mu|^2 \quad (42)$$

As the geometric mean is less than or equal to the arithmetic mean, the result follows readily.

Based on Equation (41), we deduce that the average mutual information (ergodic capacity) of ST-OD and LCF-STC designs obey the following relationship:

$$E[\mathcal{C}_{CF}] < E[\mathcal{C}_{OD}] \quad (43)$$

For a given probability P_o, the outage capacity $\mathcal{C}^{out}$ can be found from $P(\mathcal{C} < \mathcal{C}^{out}) = P_o$. From Equation (41), we obtain that

$$\mathcal{C}^{out}_{CF} < \mathcal{C}^{out}_{OD}, \quad \forall P_o \in (0,1) \quad (44)$$

Furthermore, we find that as N_r increases, the difference between $\mathcal{C}_{OD}$ and $\mathcal{C}_{CF}$ decreases. Thus, we have

$$\lim_{N_r\to\infty} E[\mathcal{C}_{CF}] = E[\mathcal{C}_{OD}], \text{ and } \lim_{N_r\to\infty} \mathcal{C}^{out}_{CF} = \mathcal{C}^{out}_{OD} \quad (45)$$

Figure 12(a) plots average capacity *versus* N_r. In accordance with Equation (45), we observe that as N_r increases $E[\mathcal{C}_{CF}]$ converges fast to $E[\mathcal{C}_{OD}]$. Figure 12(b) depicts the cumulative distribution function (cdf) of these capacities, when $N_r = 2, 5$ and $\overline{\gamma} = 10$ dB. It shows that $\mathcal{C}_{OD} > \mathcal{C}_{CF}$ for any outage probability $P(\mathcal{C} < \mathcal{C}^{out})$. As N_r increases, however, $\mathcal{C}^{out}_{CF}$ converges to $\mathcal{C}^{out}_{OD}$.

Case 2:$N_t > 2, N_r \geq 1$, with $R_{OD} < 1$

In this case, ST-OD incurs rate loss. Using Equation (40) at high SNR, we obtain the following approximate expressions for $\mathcal{C}_{OD}$ and $\mathcal{C}_{CF}$, respectively

$$\mathcal{C}_{CF} \approx \frac{1}{N_t}\sum_{\mu=1}^{N_t}\log_2\left(\sum_{\nu=1}^{N_r}|h_{\mu,\nu}|^2\right) + \log_2\overline{\gamma} \text{ bits pcu} \quad (46)$$

$$\mathcal{C}_{OD} \approx R_{OD}\log_2\left(\frac{1}{N_t}\sum_{\mu=1}^{N_t}\sum_{\nu=1}^{N_r}|h_{\mu,\nu}|^2\right) + R_{OD}\log_2\overline{\gamma} \text{ bits pcu} \quad (47)$$

The resulting capacity curves are straight lines when we plot capacity *versus* $10\log_{10}\overline{\gamma}$. On the basis of this observation, we find that when $\overline{\gamma}$ is large enough, $\mathcal{C}_{CF} > \mathcal{C}_{OD}$. To prove this, consider the slope of each capacity curve *versus* SNR $10\log_{10}\overline{\gamma}$

$$\lim_{\overline{\gamma}\to\infty}\frac{\mathcal{C}_{CF}}{\log_{10}\overline{\gamma}} = \log_2 10 > 0$$

$$\lim_{\overline{\gamma}\to\infty}\frac{\mathcal{C}_{OD}}{\log_{10}\overline{\gamma}} = R_{OD}\log_2 10 < \log_2 10 \quad (48)$$

Since $\mathcal{C}_{CF}$ and $\mathcal{C}_{OD}$ are monotonically increasing functions and at the high SNR they become approximately straight lines, we infer that when $\overline{\gamma}$ is large enough, $\mathcal{C}_{CF} > \mathcal{C}_{OD}$.

Similar to Case 1, if $\overline{\gamma}$ is large, then

$$E[\mathcal{C}_{CF}] > E[\mathcal{C}_{OD}], \text{ and } \mathcal{C}^{out}_{CF} > \mathcal{C}^{out}_{OD} \quad (49)$$

Furthermore, as N_r increases, we have that

$$\mathcal{C}_{OD} \approx R_{OD}\mathcal{C}_{CF} \quad (50)$$

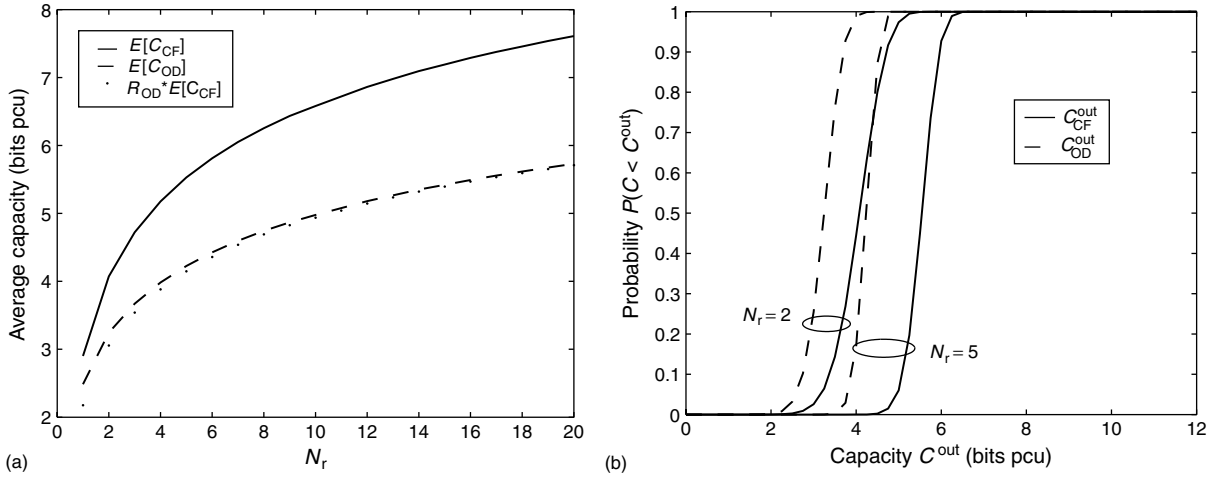

Fig. 13. Comparisons of $\mathcal{C}_{\mathrm{CF}}$ and $\mathcal{C}_{\mathrm{OD}}$ when $N_{\mathrm{t}} = 4$.

Figure 13(a) shows that when $N_{\mathrm{t}} = 4$, the gap between $E[\mathcal{C}_{\mathrm{CF}}]$ and $E[\mathcal{C}_{\mathrm{OD}}]$ is large, and $E[\mathcal{C}_{\mathrm{OD}}]$ converges to $R_{\mathrm{OD}}E[\mathcal{C}_{\mathrm{CF}}]$. In this case, $R_{\mathrm{OD}} = 3/4$. Figure 13(b) illustrates the outage performance when $N_{\mathrm{r}} = 2, 5$, and $\overline{\gamma} = 10$ dB. These simulations confirm the validity of Equation (49).

4.2. High-rate ST Coding

In the previous section, we dealt with performance-oriented STC schemes capable of rates up to 1 symbol pcu. However, multiple transmit- and/or receive-antennas allow for a significant increase in the capacity of coherent wireless systems communicating over flat Rayleigh-fading environments [48]. Various multi-antenna systems have been developed in recent years targeting such high-rate designs; VBLAST [35], and DBLAST [46] provide two early examples. Recently, designs enabling desirable performance-rate trade-offs have received increasing attention; these include combining array processing with STC [50], threaded space-time (TST) coding [51], and precoded layered STC [52,53] to name a few. However, none of these existing schemes achieve full diversity and full rate simultaneously, when the number of transmit-antennas is greater than 2 and/or the number of receive-antennas is greater than 1.

4.2.1. Full-diversity full-rate designs

For simplicity, we assume here that $N_{\mathrm{t}}N_{\mathrm{r}}$ flat-fading channels are independent and are identically distributed. Using average PEP analysis, it follows that the maximum diversity order is $G_{\mathrm{d}}^{\max} = N_{\mathrm{t}}N_{\mathrm{r}}$ (see e.g. References [20,11]). For our transmission depicted in Figure 11, the transmission rate is $R = N/P$ symbols pcu, while the maximum transmission rate is $R^{\max} = N_{\mathrm{t}}$ symbols pcu. We call $G_{\mathrm{d}}^{\max}$ and $R^{\max}$ as full diversity and full rate (FDFR), respectively. VBLAST in Reference [35] achieves full rate $R^{\max}$ but loses diversity, while LCF-STC achieves full diversity $G_{\mathrm{d}}^{\max}$ at rate $R = 1$ symbol pcu for any N_{t}. In this section, we will outline a design that guarantees full rate and full diversity, simultaneously [18].

Suppose we select the information block $\boldsymbol{s}$ to have length $N = N_g N_{\mathrm{sub}}$, that is, each block $\boldsymbol{s}$ is grouped into N_g subblocks of length N_{sub} each. Let the gth LCF-coded subblock be $\boldsymbol{u}_g := \boldsymbol{\Theta}_g \boldsymbol{s}_g$. We will henceforth view $\boldsymbol{u}_g$ as the gth *layer* (subblock). Each layer will be transmitted through all the N_{t} transmit-antennas. Each LCF-coded symbol in $\boldsymbol{u}_g$ will be transmitted once through one antenna during one time slot. As shown in Figure 11, the transmitted $P \times N_{\mathrm{t}}$ ST matrix is given as

$$\boldsymbol{C} = [\boldsymbol{c}_1 \cdots \boldsymbol{c}_{N_{\mathrm{t}}}] = \begin{bmatrix} c_1(1) & \cdots & c_{N_{\mathrm{t}}}(1) \\ \vdots & & \vdots \\ c_1(P) & \cdots & c_{N_{\mathrm{t}}}(P) \end{bmatrix} \tag{51}$$

We will specify the relationship between $\boldsymbol{C}$ and $\boldsymbol{u}$ later. Define the *spatial span* of each layer as the number of columns that contain one or more of this layer's symbols and the *temporal span* as the number of rows in which one layer's symbols are present. For example, VBLAST in Reference [35] has spatial span one, while LCF-STC has only one layer with spatial span N_{t} and with temporal span N_{t}. On the basis of these two definitions, we will introduce

two lemmas that provide necessary conditions for achieving FDFR.

Lemma 2 (Necessary condition for full diversity) *For one LCF-coded layer in $\boldsymbol{C}$ to achieve full diversity $N_t N_r$, each layer must have full spatial span ($= N_t$) and enough temporal span ($\geq N_t$).*

LCF-STC introduced in Section 4.1 has full spatial span ($=N_t$) and has enough temporal span ($=N_t$). This explains why LCF-STC can achieve full diversity. VBLAST [35] on the other hand does not have enough spatial span. The high-rate LCF-coded layered STC in Reference [52] does not have full spatial span for each layer (only some of the transmit-antennas are conveying each layer's symbols). Therefore, full diversity cannot be achieved with the scheme in Reference [52] except for the single-layer case ($N_g = 1$).

Lemma 3 (Necessary condition for full rate) *Assume that one LCF-coded symbol is transmitted through one antenna in one time slot. For an LCF-coded $P \times N_t$ ST matrix $\boldsymbol{C}$ to achieve full rate of $R^{\max} = N_t$ symbols pcu, one needs to guarantee that $N_t = N/P$, where N is the number of information symbols and P is number of time slots required to transmit the N information symbols.*

Although LCF-STC enables full diversity, it cannot achieve full rate, because only one antenna is used per time slot. Similarly, DBLAST in Reference [46] has the potential to collect full diversity, while the rate is reduced by the 'all-zero wedges' that are present in its ST matrix and by the redundancy introduced by GF coding. To achieve full rate and full diversity simultaneously, the designed ST code has to satisfy these two necessary conditions.

Toward this objective, let us now select the block length $N = N_t^2$, and let each subblock (and thus each layer) have length N_t, for a total of $N_g = N_t$ layers. The $N_t \times N_t$ ST matrix $\boldsymbol{C}$ is designed as

$$\boldsymbol{C} = \begin{bmatrix} u_1(1) & u_2(1) & \cdots & u_{N_t}(1) \\ u_{N_t}(2) & u_1(2) & \cdots & u_{N_{t-1}}(2) \\ \vdots & \vdots & \cdots & \vdots \\ u_2(N_t) & u_3(N_t) & \cdots & u_1(N_t) \end{bmatrix} \tag{52}$$

where $u_\mu(n)$ denotes the nth element of the μth layer. Checking the subscripts in Equation (52), we observe that the layers are arranged in a row circular fashion, similar to the design reported in Reference [51]. As per Lemma 3, the ST matrix in Equation (52) enjoys full transmission rate $R^{\max} = N_t$ symbols pcu. From the structure of $\boldsymbol{C}$, it is evident that each layer has full spatial span (N_t), equal to the temporal span (N_t); hence, Lemma 2 is satisfied, and the ST matrix in Equation (52) could potentially enable full diversity $G_d^{\max} = N_t N_r$.

Recall that $\boldsymbol{C}$ depends on $\{\boldsymbol{u}_g\}_{g=1}^{N_g}$ and $\boldsymbol{u}_g = \boldsymbol{\Theta}_g \boldsymbol{s}_g$. To guarantee full diversity, we need to design $\{\boldsymbol{\Theta}_g\}_{g=1}^{N_t}$ such that for any two symbol blocks $\boldsymbol{s}$ and $\boldsymbol{s}'$ [20]

$$\det[\boldsymbol{C} - \boldsymbol{C}'] \neq 0, \quad \forall \boldsymbol{s} \neq \boldsymbol{s}' \tag{53}$$

Defining $\tilde{\boldsymbol{C}} := \boldsymbol{C} - \boldsymbol{C}'$, $\tilde{\boldsymbol{u}}_g := \boldsymbol{u}_g - \boldsymbol{u}_g'$, and $\tilde{\boldsymbol{s}}_g := \boldsymbol{s}_g - \boldsymbol{s}_g'$, we can express the determinant of $\tilde{\boldsymbol{C}}$ as (cf. Equations 52 and 53)

$$\det(\tilde{\boldsymbol{C}}) = \sum_{(i_1,\ldots,i_{N_t})} (-1)^{\tau(i_1,\ldots,i_{N_t})} \prod_{n=1}^{N_t} \tilde{c}_{i_n}(n) \tag{54}$$

where $(i_1, \ldots, i_{N_t})$ is a permutation of the sequence $(1, \ldots, N_t)$; $\tau(i_1, \ldots, i_{N_t})$ is the number of inversions [54, p. 11] of the sequence $(i_1, \ldots, i_{N_t})$; and $\tilde{c}_{i_n}(n)$ is defined as in Equation (51), corresponding to $\tilde{u}_{g_n}(n)$ in Equation (52). At this point, we will need the following result proved in Reference [11].

Lemma 4 *If the constellation of $\boldsymbol{s}_g$ is carved from the ring of Gaussian integers $\mathbb{Z}(j)$, then there exists a matrix $\boldsymbol{\Theta}$, which guarantees that $\tilde{\boldsymbol{u}}_g = \boldsymbol{\Theta}\tilde{\boldsymbol{s}}_g$ has no zero entry when $\tilde{\boldsymbol{s}}_g \neq \boldsymbol{0}$. Furthermore, the entries of $\tilde{\boldsymbol{u}}_g$ belong to $\mathbb{Z}(j)$.*

Lemma 4 shows that if the Vandermonde matrix $\boldsymbol{\Theta}$ is selected properly, when $\tilde{\boldsymbol{s}}_g$ belongs to $\mathbb{Z}(j)$, then $\tilde{\boldsymbol{u}}_g = \boldsymbol{\Theta}\tilde{\boldsymbol{s}}_g$ also belongs to $\mathbb{Z}(j)$. This implies that if $\tilde{\boldsymbol{u}}_g \neq \boldsymbol{0} \in \mathbb{Z}(j)$, then we can design $\boldsymbol{\Theta}'$ such that $\boldsymbol{\Theta}'\tilde{\boldsymbol{u}}_g$ has all its entries as nonzero. On the basis of Lemma 4, we establish the following proposition:

Proposition 3 *For any constellation of $\boldsymbol{s}_g$ carved from $\mathbb{Z}(j)$ (QAM or PAM), if the ST matrix is designed as in Equation (52), then there exists at least one pair of $(\boldsymbol{\Theta}, \beta)$ for which*

$$\boldsymbol{\Theta}_g = \beta^{g-1}\boldsymbol{\Theta}, \quad g \in [1, N_t] \tag{55}$$

guarantee Equation (53), and thus enable full-diversity full-rate ST transmissions. The design of $\boldsymbol{\Theta}$ is the same as that in References [11,4] for any N_t, and β can be selected as the N_tth root of any of the generators in the Vandermonde matrix $\boldsymbol{\Theta}$; that is, $\beta = \alpha_i^{1/N_t}$, $\forall \alpha_i$, $i = 1, \ldots, N_t$.

Proof: For a given N_t, we can always design an $N_t \times N_t$ matrix Θ such that $\boldsymbol{u}_g = \beta^{g-1}\Theta \boldsymbol{s}_g$ has no zero entry (cf. Lemma 4). For an $N_t \times N_t$ matrix Θ chosen according to Lemma 4, $\det(\tilde{\boldsymbol{C}})$ becomes a polynomial in β with degree $(N_t - 1)N_t$. Note that the coefficients of this polynomial are $\prod_{n=1}^{N_t} \tilde{u}_{g_n}(n) = \prod_{n=1}^{N_t} \overline{\boldsymbol{\theta}}_n^T \tilde{\boldsymbol{s}}_{g_n}$. Since $\tilde{u}_{g_n}(n) = \overline{\boldsymbol{\theta}}_n^T \tilde{\boldsymbol{s}}_{g_n} \in \mathbb{Z}(j), \forall n$ (cf. Lemma 4), we have that the $\prod_{n=1}^{N_t} \overline{\boldsymbol{\theta}}_n^T \tilde{\boldsymbol{s}}_{g_n} \in \mathbb{Z}(j)$.

Comparing Equation (51) with Equation (52), we have that when $\tilde{c}_{i_n}(n) = \tilde{u}_{g_n}(n)$, then

$$g_n = \text{mod}(N_t + i_n - n + 1, N_t), \quad \forall n \in [1, N_t]$$

Thus, $g_n = i_{n-n+1}$, or $N_t + i_{n-n+1}$, from which it follows that

$$\sum_{n=1}^{N_t} (g_{n-1}) = mN_t, \quad m \in [0, N_t - 1] \tag{56}$$

Therefore, we deduce that for each permutation $(i_1, \ldots, i_{N_t})$,

$$\prod_{n=1}^{N_t} \tilde{c}_{i_n}(n) = \beta^{mN_t} \prod_{n=1}^{N_t} \overline{\theta}_n^T \tilde{s}_{g_n}, \quad m \in [0, N_t - 1] \tag{57}$$

Plugging Equation (57) into Equation (54), we obtain that $\det(\tilde{\boldsymbol{C}})$ is a polynomial of β^{N_t} with the coefficients in $\mathbb{Z}(j)$. Therefore, if these N_t coefficients are not equal to zero at the same time, then according to Lemma 4, there always exists β^{N_t} such that $\det(\tilde{\boldsymbol{C}}) \neq 0$ for $\boldsymbol{s} \neq \boldsymbol{s}'$. Now we need to prove that $\det(\tilde{\boldsymbol{C}})$ is not a zero polynomial; that is, that not all the coefficients of $\det(\tilde{\boldsymbol{C}})$ equal to zero simultaneously, when $\boldsymbol{s} \neq \boldsymbol{s}'$.

Suppose g is the smallest index of the layers that satisfy $\tilde{\boldsymbol{s}}_g \neq \boldsymbol{0}$, that is, $\forall g' < g$, $\tilde{\boldsymbol{s}}_{g'} = \boldsymbol{0}$. On the basis of Equation (54), we have

$$\det(\tilde{\boldsymbol{C}}) = \beta^{N_t(g-1)} \prod_{n=1}^{N_t} \boldsymbol{\theta}_n^T \tilde{\boldsymbol{s}}_g + \beta^{N_t g}(*) \tag{58}$$

where $(*)$ denotes the other terms from Equation (54). Here $\prod_{n=1}^{N_t} \boldsymbol{\theta}_n^T \tilde{\boldsymbol{s}}_g$ is the coefficient for the lowest power of $\beta^{N_t(g-1)}$, because all other terms have at least one LCF-coded symbol from the layer with index greater than g, that is, the power of β^{N_t} is greater than $\beta^{N_t(g-1)}$. Since $\tilde{\boldsymbol{s}}_g \neq \boldsymbol{0}$, from the design of Θ, we have $\prod_{n=1}^{N_t} \boldsymbol{\theta}_n^T \tilde{\boldsymbol{s}}_g \neq 0$. Therefore, $\det(\tilde{\boldsymbol{C}})$ cannot be a zero polynomial for any error pattern $\boldsymbol{s} \neq \boldsymbol{s}'$. ■

Thanks to the structure of Equation (52), we have shown that the power of β in $\det(\tilde{\boldsymbol{C}})$ is a multiple of N_t, that is, $\det(\tilde{\boldsymbol{C}})$ is a polynomial in β^{N_t} with degree $N_t - 1$. Therefore, instead of designing β according to Lemma 4 for the dimension $N_t(N_t - 1)$, we can just design β^{N_t} directly, that is, in order to guarantee full diversity, we do not need to design a new β but instead select β^{N_t} as any of the generators $\{\alpha_n\}_{n=1}^{N_t}$ for Θ. This design subsumes Reference [28], where $\beta = \alpha_1^{1/2}$ was used for $N_t = 2$. It also confirms the nonuniqueness of the full-diversity full-rate design and offers flexibility to maximize also the coding gain.

Example 5 In the following, we will use $N_t = 3$ as an application example. According to Reference [4], when $N_t = 3$, we select Θ according to the special case 2 in Section 2.3 with $\alpha = e^{j5\pi/9}$ and $\beta = e^{j5\pi/18}$. As before, we define $\tilde{\boldsymbol{C}} := \boldsymbol{C} - \boldsymbol{C}'$ and $\tilde{\boldsymbol{s}} := \boldsymbol{s} - \boldsymbol{s}'$. Using Equation (55), we find that

$$\begin{aligned}\det[\tilde{\boldsymbol{C}}] = \sum_{g=1}^{N_t} \beta^{(g-1)N_t} \Bigg(& \prod_{n=1}^{N_t} \overline{\boldsymbol{\theta}}_n^T \tilde{\boldsymbol{s}}_g(n) - \beta^{N_t} \prod_{n=1}^{N_t} \overline{\boldsymbol{\theta}}_n^T \tilde{\boldsymbol{s}}_n(n) \\ & + \prod_{n=1}^{N_t} \overline{\boldsymbol{\theta}}_n^T \tilde{\boldsymbol{s}}_n(\text{mod}(n - 1 + N_t, N_t)) \\ & + \prod_{n=1}^{N_t} \overline{\boldsymbol{\theta}}_n^T \tilde{\boldsymbol{s}}_n(\text{mod}(n - 2 + N_t, N_t)) \Bigg)\end{aligned} \tag{59}$$

where $\overline{\boldsymbol{\theta}}_n^T$ is the nth row of Θ. Given Θ in Section 2.3 and on the basis of Lemma 4, $\det[\tilde{\boldsymbol{C}}]$ is a polynomial in β^{N_t} with coefficients in $\mathbb{Z}(j)$. If we select $\beta^{N_t} = e^{j\frac{5\pi}{9}}$, $e^{j\frac{11\pi}{9}}$, or, $e^{j\frac{17\pi}{9}}$, then Equation (53) is guaranteed.

4.2.2. Decoding

From our definition of $\boldsymbol{C}$ in Equation (52), we deduce that $\boldsymbol{C}$ contains N_t^2 unknowns. This structure causes difficulty in the decoding stage. In the following, we will introduce near-optimal and suboptimal decoders for our full-rate full-diversity design.

Based on the model in Equation (37), we define the nth column of $\boldsymbol{Y}^T$ as $\overline{\boldsymbol{y}}_n$ and stack N_t columns of $\boldsymbol{Y}^T$ to form a super vector $\boldsymbol{y}$. We can then rewrite Equation (37) as

$$\boldsymbol{y} = (\boldsymbol{I}_{N_t} \otimes \boldsymbol{H}^T) \begin{bmatrix} \overline{\boldsymbol{c}}_1 \\ \vdots \\ \overline{\boldsymbol{c}}_N \end{bmatrix} + \boldsymbol{w} \tag{60}$$

where $\overline{\boldsymbol{c}}_n$ is the nth column of $\boldsymbol{C}^T$. On the basis of Equations (52) and (55), we obtain

$$\overline{\boldsymbol{c}}_n = [(\boldsymbol{P}_n \boldsymbol{D}_\beta) \otimes \overline{\boldsymbol{\theta}}_n^T]\boldsymbol{s}$$

where the permutation matrix $\boldsymbol{P}_n$, and the diagonal matrix $\boldsymbol{D}_\beta$ are defined, respectively, as

$$\boldsymbol{P}_n = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{I}_{n-1} \\ \boldsymbol{I}_{N_\mathrm{t}-n+1} & \boldsymbol{0} \end{bmatrix}, \text{ and}$$

$$\boldsymbol{D}_\beta = \mathrm{diag}[1, \beta, \ldots, \beta^{N_\mathrm{t}-1}]$$

and $\overline{\boldsymbol{\theta}}_n^T$ is the nth row of $\boldsymbol{\Theta}$. By defining $\boldsymbol{H} := \boldsymbol{I}_{N_\mathrm{t}} \otimes \boldsymbol{H}^T$ and

$$\boldsymbol{\Phi} := \begin{bmatrix} \boldsymbol{P}_1 \boldsymbol{D}_\beta \otimes \overline{\boldsymbol{\theta}}_1^T \\ \vdots \\ \boldsymbol{P}_{N_\mathrm{t}} \boldsymbol{D}_\beta \otimes \overline{\boldsymbol{\theta}}_{N_\mathrm{t}}^T \end{bmatrix}$$

we obtain

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{\Phi}\boldsymbol{s} + \boldsymbol{w} \tag{61}$$

Note that $\boldsymbol{\Phi}$ is a unitary matrix. Now we can use sphere decoding or semidefinite programming algorithms to estimate $\boldsymbol{s}$ based on Equation (61). Recall that the decoding complexity depends on the length of $\boldsymbol{s}$, which here is $N = N_\mathrm{t}^2$. When N_t is large, the decoding complexity is high even for near-optimal decoders.

Based on Equations (37) and (52), for the nth time slot, we have

$$\overline{\boldsymbol{y}}_n = \boldsymbol{H}_n \begin{bmatrix} u_1(n) \\ \vdots \\ u_{N_\mathrm{t}}(n) \end{bmatrix} + \boldsymbol{w}_n \tag{62}$$

where $\boldsymbol{H}_n = \boldsymbol{H}^T \boldsymbol{P}_n^T$. For a suboptimal decoding alternative inspired by the NC algorithm [35], we follow these major steps (see also Reference [52]):

step 1) (nulling:) On the basis of Equation (62), perform QR decomposition of $\boldsymbol{H}_n = \boldsymbol{Q}_n \boldsymbol{R}_n$ for each n and then multiply $\boldsymbol{Q}_n^{\mathcal{H}}$ on both sides of Equation (62);

step 2) (sphere decoding:) Starting from the N_tth layer, and supposing that the current layer is the gth layer, collect the received data from step 1) that depend on $\boldsymbol{u}_g$ alone and then perform SD to decode $\boldsymbol{s}_g$;

step 3) (cancelling:) On the basis of step 2), cancel the gth layer from all other layers and proceed to step 2) for the $(g-1)$st layer.

Example 6 Selecting $N_\mathrm{t} = 2$ and varying N_r, we compare our high-rate design with the LCF-STC outlined in Section 4.1 and with the ST-OD in References [45,55]. To maintain the same transmission rate, we use 16QAM for LCF-STC and ST-OD, and 4QAM for our high-rate scheme.

In Figure 14(a), we set $N_\mathrm{r} = 2$. We use the SD algorithm for the LCF-STC [11] designs, while for ST-OD, we use the ML exhaustive search for each symbol. For our FDFR-LCF-STC, we use the SD and the NC schemes. It can be seen that the performance of our FDFR design with SD is quite close to that of ST-OD. When N_r increases to 5 as shown in Figure 14(b), the gap between FDFR and ST-OD widens. This is because ST-OD loses information rate when $N_\mathrm{r} > 1$. To reduce the decoding complexity of

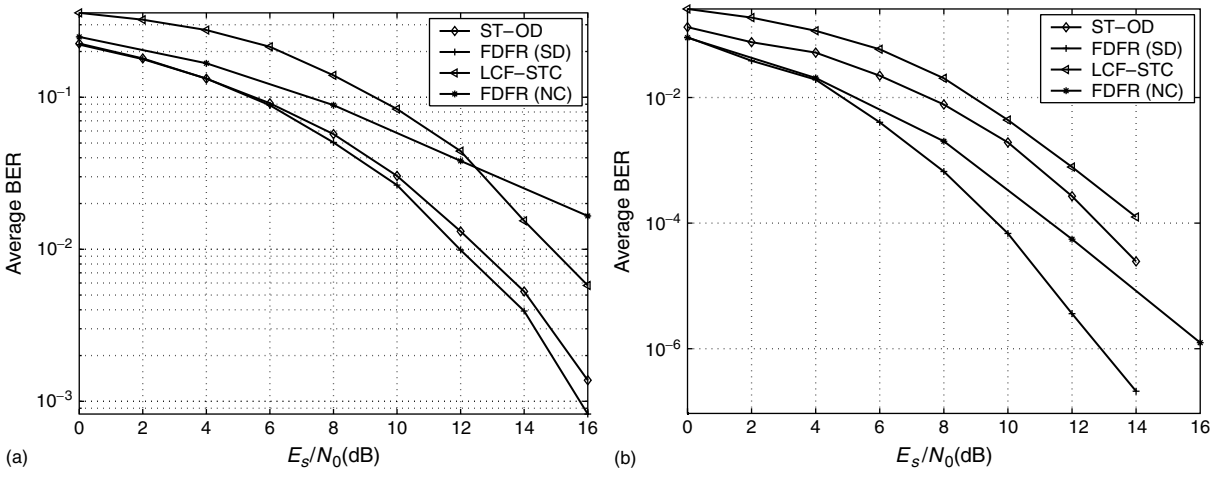

Fig. 14. Performance comparisons.

our FDFR design, we use suboptimal decoding (NC) that sacrifices performance (diversity) to lower the decoding complexity. When $N_r = 2$ [Figure 14(a)], the performance with NC is worse than that of ST-OD and the performance of FDFR with SD. However, when N_r increases to 5, FDFR with NC outperforms ST-OD and LCF-STC over a large SNR range.

4.2.3. Capacity of full-rate full-diversity designs

On the basis of Equation (60) and assuming that $\boldsymbol{s}$ is Gaussian with $\boldsymbol{R}_s = \mathcal{E}_s \boldsymbol{I}_N$, we can write the mutual information of our FDFR design as

$$\mathcal{C}_{\text{FDFR}} = \frac{1}{N_t} \log_2 \det\left(\boldsymbol{I}_N + \frac{\overline{\gamma}}{N_t} \boldsymbol{H}^{\mathcal{H}} \boldsymbol{H}\right) \text{ bits pcu}$$
$$= \log_2 \det\left(\boldsymbol{I}_{N_t} + \frac{\overline{\gamma}}{N_t} \boldsymbol{H}^{\mathcal{H}} \boldsymbol{H}\right) \text{ bits pcu} \quad (63)$$

Equation (63) shows that there is no information rate loss for our FDFR design.

Suppose that the first $\min(N_t, N_r)$ eigenvalues of $\boldsymbol{H}^{\mathcal{H}}\boldsymbol{H}$ are $\lambda_1, \ldots, \lambda_{\min(N_t,N_r)}$; then $\mathcal{C}_{\text{FDFR}}$ in Equation (63) can be rewritten as

$$\mathcal{C}_{\text{FDFR}} = \sum_{\mu=1}^{\min(N_t,N_r)} \log_2\left(1 + \frac{\overline{\gamma}}{N_t}\lambda_\mu\right) \text{ bits pcu} \quad (64)$$

Since $\sum_{\mu=1}^{\min(N_t,N_r)} \lambda_\mu = \text{tr}(\boldsymbol{H}^{\mathcal{H}}\boldsymbol{H}) = \sum_{\mu=1}^{N_t}\sum_{\nu=1}^{N_r} |h_{\mu,\nu}|^2$, we have that

$$\mathcal{C}_{\text{FDFR}} \geq \log_2\left(1 + \frac{\overline{\gamma}}{N_t} \sum_{\mu=1}^{\min(N_t,N_r)} \lambda_\mu\right) \text{ bits pcu} \quad (65)$$

At high SNR, the average capacity is

$$E[\mathcal{C}_{\text{FDFR}}] \approx \min(N_t, N_r)\log_2 \overline{\gamma}$$
$$+ E\left[\sum_{\mu=1}^{\min(N_t,N_r)} \log_2\left(\frac{\lambda_\mu}{N_t}\right)\right] \text{ bits pcu} \quad (66)$$

Equation (66) shows that $E[\mathcal{C}_{\text{FDFR}}]$ increases linearly at high SNR with slope $\min(N_t, N_r)$ (see also Reference [48]).

Closed-form expressions of the average capacity in Equation (66) can be computed in some cases (see e.g. Reference [56]). Furthermore, we observe that for fixed N_t and fixed $\overline{\gamma}$, the difference between $E[\mathcal{C}_{\text{FDFR}}]$ and $E\left[\log_2\left(1 + \frac{\overline{\gamma}}{N_t}\sum_{\mu=1}^{\min(N_t,N_r)}\lambda_\mu\right)\right]$ increases as N_r increases.

Figure 15 depicts the cdf of the capacity in Equation (63) when $N_t = 2$ and $\overline{\gamma} = 10$ dB. When $N_r = 1$, $\mathcal{C}_{\text{FDFR}}$ is the same as $\mathcal{C}_{\text{OD}}$. This result coincides with that of References [28,49]. As N_r increases, the gap between $\mathcal{C}_{\text{FDFR}}$ and $\mathcal{C}_{\text{OD}}$ (or $\mathcal{C}_{\text{CF}}$) increases.

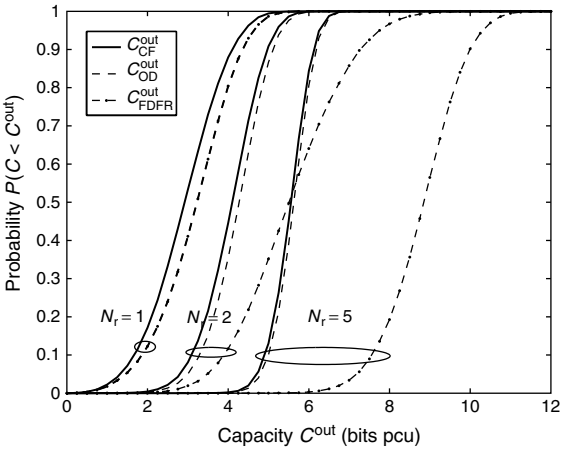

Fig. 15. Capacity cdf.

4.2.4. Rate-performance-complexity trade-offs

In Sections 4.2.1 to 4.2.3, we introduced an ST encoding and decoding scheme to achieve full diversity with full rate. Because the number of unknowns present in the ST code matrix $\boldsymbol{C}$ is high, the decoding complexity required is also high. To reduce complexity, we often have to compromise performance and/or rate. Complexity is affected by two choices: i) suboptimal *versus* (near) optimal ML decoding (this choice affects diversity); and ii) the number of layers in ST matrix (this choice affects rate).

Suppose that the channel coherence time is $T = mN_t$, for some positive integer m. Clearly, as T increases, our ST matrix $\boldsymbol{C}$ can have more rows, which increases its time span. If ML (or near-ML) decoding is used at the receiver, then we have (proof is skipped because of space limitation):

Proposition 4 *For any rate $R = N_t r/(mN_t)$, $r \in [1, mN_t]$, there always exists a layered LCF-coded ST design that guarantees full diversity $N_t N_r$, provided that ML (or near-ML) decoding is employed at the receiver over a MIMO channel with coherence time mN_t.*

This proposition shows that without sacrificing performance, it is possible to reduce decoding complexity by reducing transmission rate.

An alternative way to reduce complexity is to use suboptimal decoding (e.g. NC). In this case, the performance will be compromised.

Define the QR-decomposition of the channel matrix $\boldsymbol{H}$ that has zero-mean i.i.d. complex Gaussian elements as

$$\boldsymbol{H} = \boldsymbol{QR} = \boldsymbol{Q}\begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,N} \\ 0 & r_{2,2} & \cdots & r_{2,N} \\ 0 & 0 & \ddots & \vdots \\ 0 & \cdots & 0 & r_{N,N} \end{bmatrix} \quad (67)$$

It has been shown that [46,57,58]

$$|r_{j,k}|^2 \sim \chi^2_2, \quad \text{when } j < k;$$
$$|r_{j,j}|^2 \sim \chi^2_{2(N-j+1)}, \quad \text{when } j \in [1, N] \quad (68)$$

where χ^2_i denotes the chi-square distribution with i degrees of freedom. To quantify the price paid in performance with the NC decoder, we need the following lemma:

Lemma 5 *Suppose that matrix $\boldsymbol{H}$ is an $N \times N$ random matrix having i.i.d. complex Gaussian entries with zero mean and unit variance. The QR-decomposition of $\boldsymbol{H}$ provides a unitary matrix $\boldsymbol{Q}$ and an upper triangular matrix $\boldsymbol{R}$. Suppose $\boldsymbol{H}'$ is an $N \times N$ matrix after permuting the columns of $\boldsymbol{H}$ such that no column of $\boldsymbol{H}$ stays in the same place in $\boldsymbol{H}'$. The QR-decomposition of $\boldsymbol{H}'$ is $\boldsymbol{H}' = \boldsymbol{Q}'\boldsymbol{R}'$. Then, each entry of $\boldsymbol{R}$ is independent from its counterpart entry in $\boldsymbol{R}'$.*

An example of $\boldsymbol{H}'$ is obtained after circularly shifting the columns of $\boldsymbol{H}$. Actually this is the case for our design in Equation (52). On the basis of Lemma 5, we obtain

Proposition 5 *Using NC as proposed in Section 4.2.2, the design in Equation (52) achieves diversity order $N_t(N_r - N_t + g)$ for the gth decoded layer, when there is no error propagation.*

When error propagation is accounted for, the system performance is dominated by the worst layer. Therefore, the system diversity order can be defined as the worst layer's diversity, $N_t(N_r - N_t + 1)$. To further reduce error-propagation effects, we can use the loopback method suggested in Reference [57].

When the transmission rate is fixed, we can sacrifice diversity to lower the decoding complexity. The interesting case is if we can compromise the rate, to make up for some diversity order. In the following, we will elaborate further on these performance-rate trade-offs.

Suppose one of the layers in Equation (52) is zero. Using Proposition 5, we can then verify that the system diversity order becomes $N_t(N_r - N_t + 2)$, which is greater than that of Equation (52). However, the transmission rate now becomes $N_t - 1$ symbols pcu. How to optimize this performance-rate trade-off is an interesting open problem. Subsequently, we will use an example to illustrate this trade-off on the basis of two important parameters. One is the channel's coherence time T and the other is the layer size, which also controls the decoding complexity. A clarification is due at this point: when we talk about the performance-rate trade-off, we need to maintain comparable decoding complexity, by fixing the length of each layer. For a fixed channel coherence time T, we obtain a family of performance-rate trade-off curves by varying T.

Example 7 Select $N_t = N_r = 4$. The channel matrix $\boldsymbol{H}$ has i.i.d. complex Gaussian entries. The length of each layer is 4. The LCF-STC encoder design (see References [11,4] or Section 4.2.1) is used for each layer so that Proposition 5 holds true. We use layer as a unit to adjust the rate, that is, if we want to reduce the rate, we remove the layers instead of individual symbols.

When the channel coherence time is $T = N_t$, we can directly obtain diversity order $G_d = N_t(N_r + 1 - R)$, where $R = 1, \ldots, N_t$. When $T = 2N_t$, we have more flexibility in this trade-off since the rate is $R = r/2$, where $r = 2, \ldots, 2N_t$. When $T < N_t$, full diversity N_tN_r cannot be achieved per layer. We then select the layer length equal to $T = N_t - 1$.

From Figure 16, we observe that

i. as T increases, we have more flexibility to trade-off rate with performance;
ii. as T increases, diversity with NC suboptimal decoding comes closer to the optimal decoding; and
iii. to achieve full diversity, we need $T \geq N_t$.

Note that our diversity *versus* symbol rate figure here is different from the diversity *versus* rate figure in Reference [59]. The rate in Reference [59] is defined as the limit of transmission rate divided by $\log_2(\text{SNR})$ when the SNR goes to infinity; while in our approach, we define transmission rate as the

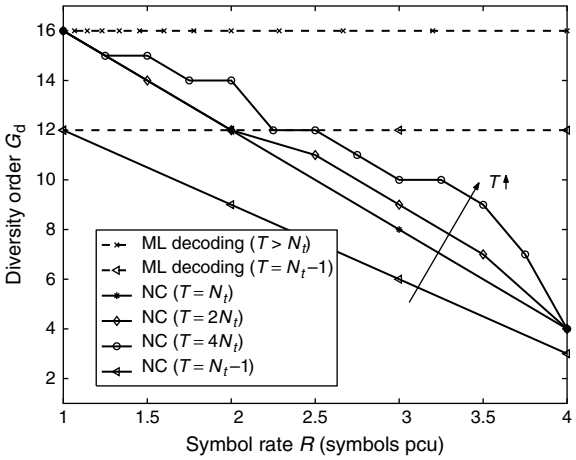

Fig. 16. Diversity *versus* rate trade-offs.

number of symbols pcu. Therefore, if we fix the constellation for any SNR, all curves in Figure 16 will show up as points at rate zero in the figure of Reference [59].

5. Concluding Summary

In this paper we have dealt with linear complex field (LCF) coding that enhances the performance of wireless fading links without rate loss, which is different from GF coding that boosts performance at the expense of rate. Albeit not interesting for AWGN channels, LCF coding holds great promise in the presence of fading as it enhances both diversity and coding gains of single- or multi-antenna transmissions over realistic fading channels that may be Gaussian, or non-Gaussian, and possibly correlated.

We relied on both average and outage performance metrics to develop a systematic design methodology of LCF block encoders. We presented explicit forms for practical block sizes and outlined the available options for batch LCF decoding. We argued that GF coding is more suitable than LCF coding in AWGN channels, while for a given rate and block size limitations LCF coding offers well-documented merits over GF coding with respect to diversity advantages in fading channels. Their complementary strengths can be combined through a concatenated GF-LCF encoder whose significant performance gains, which are unleashed with iterative (turbo) decoding, justify the resulting decoding complexity.

LCF coding offers benefits to a wide variety of fading channels beyond the flat ones. Combining it with OFDM, we saw its impact in enabling the maximum multipath diversity gains for wireless transmissions over frequency-selective fading channels. By performing LCF coding over groups of small size subblocks, it became possible not only to collect full-diversity gains without undue rate loss but also to trade off diversity (when in excess) for reduced decoding complexity. Recently, LCF coding has been utilized to collect multipath diversity [12,43], Doppler diversity [60], as well as *joint* Doppler-multipath diversity [61], space-multipath diversity [21,42,44], and space-Doppler [22] diversity. Extensions to multiuser communications have also appeared in References [8,62].

In addition to enabling various forms of diversity without sacrificing bandwidth, LCF coding also offers the *flexibility* to trade off diversity for transmission rate [52], and/or decoding complexity [22,43,44]. Such trade-offs are particularly desirable when the channel is rich in diversity, and one can afford to give up some performance in order to reduce decoding delay or reduce receiver complexity [11,12,30]. More room for delineating and optimizing these trade-offs *via* LCF coding emerges with multi-antenna systems, in which the associated multiplexing gain allows for an increased capacity of the underlying MIMO channel [48]. The trade-off between multiplexing gain and diversity has been studied in Reference [59]. The rate-performance trade-off, which is different from the multiplexing-diversity trade-off, has also received growing attention recently [18,50–52].

Our unifying framework encompassed application of our performance analysis and LCF code design methodology to MIMO fading channels. We documented the merits of LCF space-time (ST)-coded multi-antenna transmissions both in terms of achieving full-diversity gains and in terms of approaching the average and outage capacity of 'performance-oriented' space-time codes at one symbol per channel use, for any number of transmit-antennas. For 'rate-oriented' space-time codes, we identified necessary conditions for full-diversity and full-rate multi-antenna signaling and provided an LCF-ST-coded design capable of achieving them. The development of near-optimal and suboptimal decoders and the analysis of average and outage capacity revealed one of the most attractive features of LCF-ST-coded multi-antenna systems: they provide the designer with ample flexibility to delineate

and optimize the performance-rate-complexity tradeoffs that emerge depending on the wireless application at hand.

Acknowledgments

The authors would like to thank Dr. A. Swami from the Army Research Lab, Prof. Z. Wang from Iowa State University and Dr. S. Zhou from the University of Minnesota for their helpful suggestions that improved the presentation of the material. The work in this paper was supported by the NSF Wireless Initiative Grant No. 9979443, the NSF Grant No. 0122431, and by the ARL/CTA Grant No. DAAD19-01-2-011.

References

1. Biglieri E, Proakis J, Shamai S. Fading channels: information-theoretic and communications aspects. *IEEE Transactions on Information Theory* 1998; **44**: 1619–1692.
2. Boullè K, Belfiore JC. Modulation scheme designed for the Rayleigh fading channel. In *Proceedings of the CISS*, Princeton, NJ, March 1992; pp. 288–293.
3. Giraud X, Belfiore JC. Constellations matched to the Rayleigh fading channel. *IEEE Transactions on Information Theory* 1996; **42**(1): 106–115.
4. Giraud X, Boutillon E, Belfiore JC. Algebraic tools to build modulation schemes for fading channels. *IEEE Transactions on Information Theory* 1997; **43**(3): 938–952.
5. Giannakis GB. Filterbanks for blind channel identification and equalization. *IEEE Signal Processing Letters* 1997; **4**: 184–187.
6. Scaglione A, Giannakis GB, Barbarossa S. Redundant filterbank precoders and equalizers, Part I: unification and optimal designs. *IEEE Transactions on Signal Processing* 1999; **47**: 1988–2006.
7. Scaglione A, Giannakis GB, Barbarossa S. Redundant filterbank precoders and equalizers, Part II: blind channel estimation, synchronization, and direct equalization. *IEEE Transactions on Signal Processing* 1999; **47**: 2007–2022.
8. Wang Z, Giannakis GB. Wireless multicarrier communications: where Fourier meets Shannon. *IEEE Signal Processing Magazine* 2000; **17**(3): 29–48.
9. Xin Y, Wang Z, Giannakis GB. Linear unitary precoders for maximum diversity gains with multiple transmit- and receive-antennas. In *Proceedings of the 34th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, October 29–November 1 2000; pp. 1553–1557.
10. Xin Y, Wang Z, Giannakis GB. Space-time constellation-rotating codes maximizing diversity and coding gains. In Proceedings of the GLOBECOM, Vol. 1, San Antonio, TX, 25–27 November 2001; pp. 455–459.
11. Xin Y, Wang Z, Giannakis GB. Space-time diversity systems based on linear constellation precoding. *IEEE Transactions on Wireless Communications* 2002; to appear.
12. Wang Z, Giannakis GB. Complex-field coding for OFDM over fading wireless channels. *IEEE Transactions on Information Theory* 2002; to appear.
13. DaSilva VM, Sousa ES. Fading-resistant modulation using several transmitter antennas. *IEEE Transactions on Communications* 1997; **45**(10): 1236–1244.
14. Boutros J, Viterbo E. Signal space diversity: a power- and bandwidth-efficient diversity technique for the Rayleigh fading channel. *IEEE Transactions on Information Theory* 1998; **44**(4): 1453–1467.
15. Damen MO, Chkeif A, Belfiore J-C. Lattice codes decoder for space-time codes. *IEEE Communications Letters* 2000; **4**: 161–163.
16. Finkce U, Pohst M. Improved methods for calculating vectors of short length in lattice, including a complexity analysis. *Mathematics of Computation* 1985; **44**: 463–471.
17. Wang Z, Zhou S, Giannakis GB. Joint coded-precoded OFDM with low-complexity turbo-decoding. In *Proceedings of the European Wireless Conference*, Florence, Italy, 25–28 February 2002; pp. 648–654.
18. Ma X, Giannakis GB. Layered space-time complex field coding: full-diversity with full-rate, and tradeoffs. In *Proceedings of the 2nd Sensor Array and Multichannel SP Workshop*, Rosslyn, VA, 4–6 August 2002.
19. Benedetto S, Biglieri E. *Principles of Digital Transmission with Wireless Applications*. Kluwer Academic/Plenum Publishers: New York, 1999.
20. Tarokh V, Seshadri N, Calderbank AR. Space-time codes for high data rate wireless communication: performance criterion and code construction. *IEEE Transactions on Information Theory* 1998; **44**(2): 744–765.
21. Zhou S, Giannakis GB. Space-time coding with maximum diversity gains over frequency-selective fading channels. *IEEE Signal Processing Letters* 2001; **8**(10): 269–272.
22. Ma X, Leus G, Giannakis GB. Space-time-Doppler coding for correlated time-selective fading channels. *IEEE Transactions on Information Theory* 2002; submitted;. see also In *Proceedings of the International Conference on ASSP*, Orlando, FL, May 13–17, 2002; pp. 2217–2220.
23. Wang Z, Giannakis GB. A simple and general parameterization quantifying performance in fading channels. *IEEE Transactions on Communications* 2002; submitted.
24. Simon MK, Alouini M-S. *Digital Communications over Fading Channels: A Unified Approach to Performance Analysis*. John Wiley & Sons: New York, 2000.
25. Stüber GL. *Principles of Mobile Communications*. Kluwer Academic Publishers: Norwell, MA, 1996.
26. Boutros J, Viterbo E, Rastello C, Belfiore JC. Good lattice constellations for both Rayleigh and Gaussian channels. *IEEE Transactions on Information Theory* 1996; **42**(2): 502–518.
27. Damen MO, Abed-Meraim K, Belfiore J-C. Diagonal algebraic space-time block codes. *IEEE Transactions on Information Theory* 2002; **48**(3): 628–636.
28. Damen MO, Tewfik A, Belflore J-C. A construction of a space-time code based on number theory. *IEEE Transactions on Information Theory* 2002; **48**(3): 753–760.
29. Pohst M. On the computation of lattice vectors of minimal length, successive minima and reduced bases with applications. *ACM SIGSAM Bulletin* 1981; **15**: 37–44.
30. Viterbo E, Boutros J. A universal lattice code decoder for fading channels. *IEEE Transactions on Information Theory* 1999; **45**(5): 1639–1642.
31. Hassibi B, Vikalo H. On the expected complexity of sphere decoding. In *Proceedings of the 35th Asilomar Conference on Signals, Systems, and Computers*, Vol. 2, Pacific Grove, CA, 4–7 November 2001; pp. 1051–1055.
32. Hochwald B, Ten Brink S. Achieving Near-Capacity on a Multiple-Antenna Channel; downloadable at http://mars.bell-labs.com.
33. Agrell E, Eriksson T, Vardy A, Zeger K. Closest point search in lattices. *IEEE Transactions on Information Theory* 2002; **48**(8): 2201–2214.
34. Ma W-K, Davidson TN, Wong KM, Luo Z-Q, Ching P-C. Quasi-maximum-likelihood multiuser detection using semi-definite relaxation with application to synchronous CDMA.

IEEE Transactions on Signal Processing 2002; **50**(4): 912–922.
35. Wolniansky PW, Foschini GJ, Golden GD, Valenzuela RA. V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel. In *Proceedings of the URSI International Symposium Signals, Systems, and Electronics*, Italy, September 1998.
36. Al-Dhahir N, Sayed AH. The finite-length multi-input multi-output MMSE-DFE. *IEEE Transactions on Signal Processing* 2000; **48**(10): 2921–2936.
37. Stamoulis A, Giannakis GB, Scaglione A. Block FIR decision-feedback equalizers for filterbank precoded transmissions with blind channel estimation capabilities. *IEEE Transactions on Communications* 1999; **49**(7): 1456–1466.
38. Rainish D. Diversity transform for fading channels. *IEEE Transactions on Communications* 1996; **44**(12): 1653–1661.
39. Zou W, Wu Y. COFDM: an overview. *IEEE Transactions on Broadcasting* 1995; **41**: 1–8.
40. Le Floch B, Alard M, Berrou C. Coded orthogonal frequency division multiplex [TV broadcasting]. *Proceedings of the IEEE* 1995; **83**(6): 982–996.
41. ETSI Normalization Committee. Channel Models for Hiper-Lan/2 in Different Indoor Scenarios, Norme ETSI, Document 3ERI085B, European Telecommunications Standards Institute, Sophia-Antipolis, France, 1998.
42. Zhou S, Giannakis GB. Single-carrier space-time block coded transmissions over frequency-selective fading Channels. *IEEE Transactions on Information Theory*, submitted May 2001 revised April 2002.
43. Liu Z, Xin Y, Giannakis GB. Linear constellation precoding for OFDM with maximum multipath diversity and coding gains. In *Proceedings of the 35th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, 4–7 November 2001; pp. 1445–1449.
44. Liu Z, Xin Y, Giannakis GB. Space-time-frequency coded OFDM with sub-carrier grouping and constellation precoding. In *Proceedings of the International Conference on ASSP*, Orlando, FL, 13–17 May 2002.
45. Tarokh V, Jafarkhani H, Calderbank AR. Space-time block codes from orthogonal designs. *IEEE Transactions on Information Theory* 1999; **45**(5): 1456–1467.
46. Foschini GJ. Layered space-time architecture for wireless communication in fading environments when using multiple antennas. *Bell Labs. Tech. Journal* 1996; **2**: 41–49.
47. Emre Telatar I. Capacity of multi-antenna Gaussian channels. *European Telecommunications Transactions* 1999; **10**(6): 585–596.
48. Foschini GJ, Gans MJ. On limits of wireless communication in a fading environment when using multiple antennas. *Wireless Personal Communications* 1998; **6**(3): 311–335.
49. Hassibi B, Hochwald BM. High-rate codes that are linear in space and time. *IEEE Transactions on Information Theory* 2002; **48**(7): 1804–1824.
50. Tarokh V, Naguib A, Seshadri N, Calderbank AR. Combined array processing and space-time coding. *IEEE Transactions on Information Theory* 1999; **45**: 1121–1128.
51. El Gamal H, Hammons Jr AR. A new approach to layered space-time coding and signal processing. *IEEE Transactions on Information Theory* 2001; **47**(6): 2321–2334.
52. Xin Y, Liu Z, Giannakis GB. High-rate layered space-time coding based on linear constellation precoding. In *Proceedings of the Wireless Communications and Networking Conference*, Vol. 1, Orlando, FL, 17–21 March 2002; pp. 471–476.
53. Sidiropoulos ND, Budampati R. Khatri-Rao space-time codes. *IEEE Transactions on Signal Processing*, special issue on signal processing techniques for space-time coded transmissions, 2002; to appear;. see also In *Proceeding of the 39th Allerton Conference*, University of Illinois at U-C, Monticello, IL, October 3–5, 2001; pp. 1316–1325.
54. Knuth DE. *Art of Computer Programming Vol. 3: Sorting and Searching*. 2nd edition, Addison-Wesley: Reading, MA, 1998.
55. Alamouti SM. A simple transmit diversity technique for wireless communications. *IEEE Journal on Selected Areas in Communications* 1998; **16**(8): 1451–1458.
56. Hochwald BM, Marzetta TL, Hassibi B. Space-time autocoding. *IEEE Transactions on Information Theory* 2001; **47**(7): 2761–2781.
57. Liu K, Sayeed AM. Improved layered space-time processing in multiple antenna systems. In *Proceedings of the 39th Annual Allerton Conference on Communication, Control and Computing*, Monticello, IL, October 2001.
58. Shiu D, Kahn JM. Design of high-throughput codes for multiple-antenna wireless systems. *IEEE Transactions on Information Theory* 1999; downloadable from http://www.cs.berkeley.edu/ jmk//res.areas/mawc.html; submitted.
59. Zheng L, Tse D. Optimal diversity-multiplexing tradeoff in multi-antenna channels. In *Proceeding of the 39th Allerton Conference on Communication, Control and Computing*, Monticello, IL, October 2001; pp. 835–844.
60. Ma X, Giannakis GB. Maximum-diversity transmissions over time-selective wireless channels. In *Proceedings of the Wireless Communications and Networking Conference*, Vol. 1, Orlando, FL, 17–21 March 2002; pp. 497–501.
61. Ma X, Giannakis GB. Designing maximum multipath-Doppler diversity transmissions over time- and frequency-selective wireless channels. In *Proceedings of the 39th Allerton Conference*, University of Illinois at U-C, Monticello, IL, 3–5 October 2001; pp. 682–691.
62. Liu Z, Giannakis GB. Space-time block coded multiple access through frequency-selective fading channels. *IEEE Transactions on Communications* 2001; **49**(6): 1033–1044.
63. Hassibi B, Hochwald BM. Cayley differential unitary space-time codes. *IEEE Transactions on Information Theory* 2002; **48**(6): 1485–1503.
64. Helard JF, Le Floch B. Trellis coded orthogonal frequency division multiplexing for digital video transmission. In *Proceedings of the Global Telecommunications Conference*, Vol. 2, Phoenix, AZ, 2–5 December 1991; pp. 785–791.
65. Jeličić BD, Roy S. Design of trellis coded QAM for flat fading and AWGN channels. *IEEE Transactions on Vehicular Technology* 1994; **44**(1): 192–201.
66. Lamy C, Boutros J. On random rotations diversity and minimum MSE decoding of lattices. *IEEE Transactions on Information Theory* 2000; **46**(4): 1584–1589.

Authors' Biographies

Xiaoli Ma received her B.S. degree in automatic control from Tsinghua University, P.R. China, in 1998, and her M.S. degree in electrical engineering from University of Virginia in 1999. She is currently a Ph.D. student at the Department of Electrical and Computer Engineering, University of Minnesota.

Her research interests include transmitter and receiver diversity techniques for fading channels, communications over time- and frequency-selective channels, space-time coding, channel estimation and equalization algorithms, and carrier synchronization for OFDM systems.

Georgios B. Giannakis (F'97) G. B. Giannakis received his diploma in electrical engineering from the National Technical University of Athens, Greece, 1981. From September 1982 to July 1986 he was with the University of Southern California (USC), where he received his M.S. in electrical engineering, 1983, M.S. in mathematics, 1986, and Ph.D. in electrical engineering, 1986. After lecturing for one year at USC, he joined the University of Virginia in 1987, where he became a professor of electrical engineering in 1997. Since 1999, he has been a professor with the Department of Electrical and Computer Engineering at the University of Minnesota, where he now holds an ADC Chair in Wireless Telecommunications.

His general interests span the areas of communications and signal processing, estimation and detection theory, time-series analysis, and system identification—subjects on which he has published more than 150 journal papers, 270 conference papers, and two edited books. Current research topics focus on transmitter and receiver diversity techniques for single- and multiuser fading communication channels, complex field and space-time coding for block transmissions, multicarrier, and wide-band wireless communication systems.

G. B. Giannakis is the (co-) recipient of four best paper awards from the IEEE Signal Processing (SP) Society (1992, 1998, 2000, 2001). He also received the Society's Technical Achievement Award in 2000. He co-organized three IEEE-SP Workshops, and guest (co-) edited four special issues. He has served as editor-in-chief for the *IEEE SP Letters*, as associate editor for the *IEEE Trans. on Signal Proc.* and the *IEEE SP Letters*, as secretary of the SP Conference Board, as member of the SP Publications Board, as member and vice-chair of the Statistical Signal and Array Processing Technical Committee, and as chair of the SP for Communications Technical Committee. He is a member of the editorial board for the *Proceedings of the IEEE*, and the steering committee of the *IEEE Trans. on Wireless Communications*. He is a Fellow of the IEEE, a member of the IEEE Fellows Election Committee, the IEEE-SP Society's Board of Governors, and a frequent consultant for the telecommunications industry.