Single-Carrier Space–Time Block-Coded Transmissions Over Frequency-Selective Fading Channels

Shengli Zhou, *Member, IEEE,* and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*—We study space–time block coding for single-carrier block transmissions over frequency-selective multipath fading channels. We propose novel transmission schemes that achieve a maximum diversity of order $N_tN_r(L+1)$ in rich scattering environments, where N_t (N_r) is the number of transmit (receive) antennas, and L is the order of the finite impulse response (FIR) channels. We show that linear receiver processing collects full antenna diversity, while the overall complexity remains comparable to that of single-antenna transmissions over frequency-selective channels. We develop transmissions enabling maximum-likelihood optimal decoding based on Viterbi's algorithm, as well as turbo decoding. With single receive and two transmit antennas, the proposed transmission format is capacity achieving. Simulation results demonstrate that joint exploitation of space-multipath diversity leads to significantly improved performance in the presence of frequency-selective fading channels.**

***Index Terms*—Block transmissions, frequency-selective multipath channels, space–time block coding.**

I. Introduction

SPACE–TIME (ST) coding has by now been well documented as an attractive means of achieving high data rate transmissions with diversity and coding gains in wireless applications; see, e.g., [27], [30] for tutorial treatments. So far, ST codes are mainly designed for frequency-flat channels. However, future broad-band wireless systems will communicate symbols with duration smaller than the channel delay spread, which gives rise to frequency-selective propagation effects. Targeting broad-band wireless applications, it is thus important to design ST codes in the presence of frequency-selective multipath channels.

Unlike flat fading channels, optimal design of ST codes for dispersive multipath channels is complex because signals from different antennas are mixed not only in space but also in time. In order to maintain decoding simplicity and take advantage of existing ST coding designs for flat fading channels, most existing works have pursued (suboptimal) two-step approaches. First, they mitigate intersymbol interference (ISI) by converting frequency-selective fading channels to flat fading ones, and then design ST coders and decoders for the resulting flat fading channels. One approach to ISI mitigation is to employ a rather complex multiple-input multiple-output equalizer (MIMO-EQ) at the receiver to turn finite impulse response (FIR) channels into temporal ISI-free ones [10], [11]. Another approach, with lower receiver complexity, is to employ orthogonal frequency division multiplexing (OFDM), which converts frequency-selective multipath channels into a set of flat fading subchannels through inverse fast Fourier transform (IFFT) and cyclic prefix (CP) insertion at the transmitter, together with CP removal and fast Fourier transform (FFT) processing at the receiver [44]. On the flat fading OFDM subchannels, many authors have applied ST coding for transmissions over frequency-selective channels, including [31] that assumes channel knowledge, and [1], [20], [21], [26], [27] that require no channel knowledge at the transmitter. The ST trellis codes of [38] are employed in [1], [21] across OFDM subcarriers, while the orthogonal ST block codes (STBCs) of [3], [36] are adopted by [20], [26], [27] on each OFDM subcarrier.

Although using ST codes designed for flat fading channels can at least achieve full multiantenna diversity [37], the potential diversity gains embedded in multipath propagation have not been addressed thoroughly. Recently, in OFDM-based systems, it was first claimed in [6], and then [28], that it is possible to achieve both multiantenna and multipath diversity gains of order equal to the product of the number of transmit antennas, the number of receive antennas, and the number of FIR channel taps. However, code designs which guarantee full exploitation of the embedded diversity were not provided in [6], [28]. The simple design of [7] achieves full diversity, but it is essentially a repeated transmission, which decreases the transmission rate considerably (see Section II-F for details). On the other hand, for *single-antenna* transmissions, it is shown in [45] that a diversity order equal to the number of FIR taps is achievable when OFDM transmissions are linearly precoded across subcarriers.

An inherent limitation of all multicarrier (OFDM) based ST transmissions is their nonconstant modulus, which necessitates power amplifier back-off, and thus reduces power efficiency [34]. In addition, multicarrier schemes are more sensitive to carrier frequency offsets relative to their single-carrier counterparts [34]. These two facts motivate well ST codes for single-carrier transmissions over frequency-selective channels, that have been looked upon recently in [2], [22], [42], [47] with block coding, and in [23] using trellis coding.

In this paper, we design *ST block codes* for *single-carrier block transmissions* in the presence of frequency-selective

Manuscript received May 4, 2001; revised April 1, 2002 and August 30, 2002. This work was supported by the NSF Wireless Initiative under Grant 99-79443, the NSF under Grant 0105612, and by the ARL/CTA under Grant DAAD19-01-2-011. The material in this paper was presented in part at the IEEE Global Communications Conference, San Antonio, TX, November 2001.

The authors are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: szhou@ece.umn.edu; georgios@ece.umn.edu).

Communicated by G. Caire, Associate Editor for Communications.

Digital Object Identifier 10.1109/TIT.2002.806158

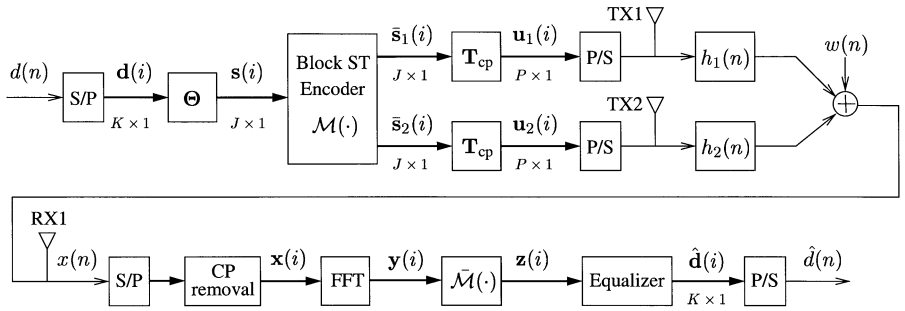

Fig. 1. Single-carrier ST transceiver model.

fading channels. We propose novel transmission formats, that subsume those in [2], [22], [42], as special cases. Furthermore, we show that a maximum diversity up to order $N_tN_r(L+1)$ is achieved in a rich scattering environment, where N_t is the number of transmit antennas, N_r is the number of receive antennas, and $(L+1)$ is the number of taps corresponding to each FIR channel. With single receive and two transmit antennas, our transmission offers a capacity-achieving scheme.

Being counterparts of orthogonal STBCs [3], [36], but for frequency- selective channels, our proposed schemes enable simple linear processing to collect full antenna diversity, and incur receiver complexity that is comparable to single-antenna transmissions. Interestingly, our transmissions enable exact application of Viterbi's algorithm for maximum-likelihood (ML) optimal decoding, in addition to various reduced-complexity suboptimal equalization alternatives. Equally important, when our ST transmissions are combined with channel coding, they facilitate application of iterative (turbo) equalizers. Simulation results demonstrate that joint exploitation of space-multipath diversity leads to significantly improved performance in the presence of frequency-selective multipath channels.

The rest of this paper is organized as follows. Section II deals with the important special case of single receive and two transmit antennas. Section III details the equalization and decoding designs. Section IV generalizes the proposed schemes to multiple transmit and receive antennas. Simulation results are presented in Section V, while conclusions are drawn in Section VI.

Notation: Bold upper case letters denote matrices, bold lower case letters stand for column vectors; $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^{\mathcal{H}}$ denote conjugate, transpose, and Hermitian transpose, respectively; $\mathrm{E}\{\cdot\}$ for expectation, $\mathrm{tr}\{\cdot\}$ for the trace of a matrix, $\|\cdot\|$ for the Euclidean norm of a vector; $\mathbf{I}_K$ denotes the identity matrix of size K, $\mathbf{0}_{M\times N}$ $(\mathbf{1}_{M\times N})$ denotes an all-zero (all-one) matrix with size $M\times N$, and $\mathbf{F}_N$ denotes an $N\times N$ FFT matrix with the $(p+1, q+1)$st entry $(1/\sqrt{N})\exp(-j2\pi pq/N)$, $\forall\, p, q \in [0, N-1]$; $\mathrm{diag}(\mathbf{x})$ stands for a diagonal matrix with $\mathbf{x}$ on its diagonal. $[\cdot]_p$ denotes the $(p+1)$st entry of a vector, and $[\cdot]_{p,q}$ denotes the $(p+1, q+1)$st entry of a matrix.

II. Single-Carrier Block Transmissions

Fig. 1 depicts the discrete-time equivalent baseband model of a communication system with $N_t = 2$ transmit antennas and $N_r = 1$ receive antenna. We detail this important special case first, and then generalize to more than two transmit antennas and multiple receive antennas in Section IV. The information-bearing data symbols $d(n)$ belonging to an alphabet $\mathcal{A}$ are first parsed to $K\times 1$ blocks $\mathbf{d}(i) := [d(iK), \ldots, d(iK+K-1)]^T$, where the serial index n is related to the block index i by $n = iK + k$, $k \in [0, K-1]$. The blocks $\mathbf{d}(i)$ are precoded by a $J\times K$ matrix $\mathbf{\Theta}$ (with entries in the complex field) to yield $J\times 1$ symbol blocks $\mathbf{s}(i) := \mathbf{\Theta}\mathbf{d}(i)$. The linear precoding by $\mathbf{\Theta}$ can be either nonredundant with $J = K$ or redundant when $J > K$. The ST encoder takes as input two consecutive blocks $\mathbf{s}(2i)$ and $\mathbf{s}(2i+1)$ to output the following $2J\times 2$ ST block-coded matrix:

$$\begin{bmatrix} \bar{\mathbf{s}}_1(2i) & \bar{\mathbf{s}}_1(2i+1) \\ \bar{\mathbf{s}}_2(2i) & \bar{\mathbf{s}}_2(2i+1) \end{bmatrix} := \begin{bmatrix} \mathbf{s}(2i) & -\mathbf{P}\mathbf{s}^*(2i+1) \\ \mathbf{s}(2i+1) & \mathbf{P}\mathbf{s}^*(2i) \end{bmatrix} \begin{matrix} \rightarrow \text{time} \\ \downarrow \text{space} \end{matrix} \quad (1)$$

where $\mathbf{P}$ is a permutation matrix that is drawn from a set of permutation matrices $\{\mathbf{P}_J^{(n)}\}_{n=0}^{J-1}$, with J denoting the dimensionality $J\times J$. Each $\mathbf{P}_J^{(n)}$ performs a reverse cyclic shift (that depends on n) when applied to a $J\times 1$ vector

$$\mathbf{a} := [a(0), a(1), \ldots, a(J-1)]^T.$$

Specifically, the $(p+1)$st entry of $\mathbf{P}_J^{(n)}\mathbf{a}$ is

$$[\mathbf{P}_J^{(n)}\mathbf{a}]_p = a((J-p+n-1)\mathrm{mod}J).$$

Two important special cases are $\mathbf{P}_J^{(0)}$ and $\mathbf{P}_J^{(1)}$. The output of

$$\mathbf{P}_J^{(0)}\mathbf{a} = [a(J-1), a(J-2), \ldots, a(0)]^T$$

performs *time reversal* of $\mathbf{a}$, while

$$\begin{aligned} \mathbf{P}_J^{(1)}\mathbf{a} &= [a(0), a(J-1), a(J-2), \ldots, a(1)]^T \\ &= \mathbf{F}_J^{-1}\mathbf{F}_J^{\mathcal{H}}\mathbf{a} \\ &= \mathbf{F}_J^{\mathcal{H}}\mathbf{F}_J^{\mathcal{H}}\mathbf{a} \end{aligned}$$

corresponds to taking the J-point IFFT twice on the vector $\mathbf{a}$. This double IFFT operation in the ST coded matrix is, in fact, a special case of the Z-transform approach originally proposed in [25], with the Z-domain points chosen to be equally spaced on the unit circle $\{e^{j\frac{2\pi}{J}n}\}_{n=0}^{J-1}$. Note that in our notation, [22] uses only $\mathbf{P}_J^{(0)}$, [2] uses only $\mathbf{P}_J^{(1)}$, and [42] uses both $\mathbf{P}_J^{(0)}$ and $\mathbf{P}_J^{(1)}$. Our unifying view here allows for any $\mathbf{P}$ from the set $\{\mathbf{P}_J^{(n)}\}_{n=0}^{J-1}$.

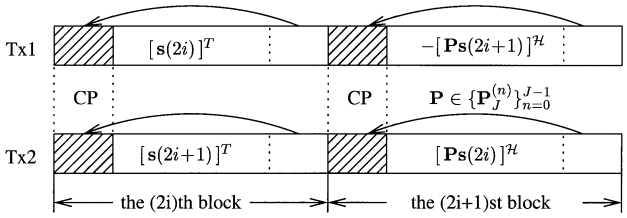

Fig. 2. The transmitted sequence for CP-based block transmissions.

At each block transmission time interval i, the blocks $\bar{\mathbf{s}}_1(i)$ and $\bar{\mathbf{s}}_2(i)$ are forwarded to the first and the second antennae, respectively. From (1), we have that

$$\bar{\mathbf{s}}_1(2i+1) = -\mathbf{P}\bar{\mathbf{s}}_2^*(2i), \qquad \bar{\mathbf{s}}_2(2i+1) = \mathbf{P}\bar{\mathbf{s}}_1^*(2i) \quad (2)$$

which shows that each transmitted block from one antenna at time slot $2i+1$ is a conjugated and permuted version of the corresponding transmitted block from the other antenna at time slot $2i$ (with a possible sign change). For flat fading channels, symbol blocking is unnecessary, i.e., $J = K = 1$ and $\mathbf{P} = 1$, and the design of (1) reduces to the well-known Alamouti ST code matrix [3]. However, for frequency-selective multipath channels, the permutation matrix $\mathbf{P}$ is necessary as will be clarified soon.

To avoid interblock interference (IBI) in the presence of frequency-selective multipath channels, we adopt the CP approach [12], [34], [44], which inserts a CP for each block before transmission. Mathematically, at each antenna $\mu \in [1, 2]$, a tall $P \times J$ transmit matrix $\mathbf{T}_{\rm cp} := [\mathbf{I}_{\rm cp}^T, \mathbf{I}_J^T]^T$, with $\mathbf{I}_{\rm cp}$ comprising the last $P-J$ rows of $\mathbf{I}_J$, is applied on $\bar{\mathbf{s}}_\mu(i)$ to obtain $P \times 1$ blocks $\mathbf{u}_\mu(i) = \mathbf{T}_{\rm cp}\bar{\mathbf{s}}_\mu(i)$. Indeed, multiplying $\mathbf{T}_{\rm cp}$ with $\bar{\mathbf{s}}_\mu(i)$ replicates the last $P-J$ entries of $\bar{\mathbf{s}}_\mu(i)$ and places them on its top. The transmitted sequences from both antennas are depicted in Fig. 2.

With symbol rate sampling,[1] let $\mathbf{h}_\mu := [h_\mu(0), \ldots, h_\mu(L)]^T$ be the equivalent discrete-time channel impulse response (that includes transmit–receive filters as well as multipath effects) between the μth transmit antenna and the single receive antenna, where L is the channel order. With the CP length at least as long as the channel order, $P-J = L$, the IBI can be avoided at the receiver by discarding the received samples corresponding to the cyclic prefix. CP insertion at the transmitter together with CP removal at the receiver yields the following channel input–output relationship in matrix–vector form (see e.g., [44] for a detailed derivation in the single-antenna scenario):

$$\mathbf{x}(i) = \sum_{\mu=1}^{2} \tilde{\mathbf{H}}_\mu \bar{\mathbf{s}}_\mu(i) + \mathbf{w}(i) \quad (3)$$

where the channel matrix $\tilde{\mathbf{H}}_\mu$ is circulant with $[\tilde{\mathbf{H}}_\mu]_{p,q} = h_\mu((p-q) \bmod J)$, and the additive Gaussian noise $\mathbf{w}(i)$ is assumed to be white with each entry having variance $\sigma_w^2 = N_0$.

The receiver will exploit the following two nice properties of circulant matrices.

[1]Extension to fractional sampling is straightforward; we here focus on symbol rate sampling for simplicity.

p1) Circulant matrices can be diagonalized by FFT operations [18, p. 202]

$$\tilde{\mathbf{H}}_\mu = \mathbf{F}_J^{\mathcal{H}} \mathbf{D}(\tilde{\mathbf{h}}_\mu)\mathbf{F}_J \quad \text{and} \quad \tilde{\mathbf{H}}_\mu^{\mathcal{H}} = \mathbf{F}_J^{\mathcal{H}} \mathbf{D}(\tilde{\mathbf{h}}_\mu^*)\mathbf{F}_J \quad (4)$$

where $\mathbf{D}(\tilde{\mathbf{h}}_\mu) := \mathrm{diag}(\tilde{\mathbf{h}}_\mu)$, and the vector

$$\tilde{\mathbf{h}}_\mu := [H_\mu(e^{j0}), \ldots, H_\mu(e^{j\frac{2\pi}{J}(J-1)})]^T$$

has the $(p+1)$st entry being the channel frequency response

$$H_\mu(z) := \sum_{l=0}^{L} h_\mu(l) z^{-l}$$

evaluated at the frequency $z = e^{j\frac{2\pi}{J}p}$.

p2) As we prove in the Appendix, pre- and postmultiplying $\tilde{\mathbf{H}}_\mu$ by $\mathbf{P}$ yields $\tilde{\mathbf{H}}_\mu^T$:

$$\mathbf{P}\tilde{\mathbf{H}}_\mu\mathbf{P} = \tilde{\mathbf{H}}_\mu^T \quad \text{and} \quad \mathbf{P}\tilde{\mathbf{H}}_\mu^*\mathbf{P} = \tilde{\mathbf{H}}_\mu^{\mathcal{H}}. \quad (5)$$

With the ST coded blocks satisfying (2), let us consider two consecutive received blocks [cf. (3)]:

$$\mathbf{x}(2i) = \tilde{\mathbf{H}}_1\bar{\mathbf{s}}_1(2i) + \tilde{\mathbf{H}}_2\bar{\mathbf{s}}_2(2i) + \mathbf{w}(2i) \quad (6)$$

$$\mathbf{x}(2i+1) = -\tilde{\mathbf{H}}_1\mathbf{P}\bar{\mathbf{s}}_2^*(2i) + \tilde{\mathbf{H}}_2\mathbf{P}\bar{\mathbf{s}}_1^*(2i) + \mathbf{w}(2i+1). \quad (7)$$

Left-multiplying (7) by $\mathbf{P}$, conjugating, and using p2), we arrive at

$$\mathbf{P}\mathbf{x}^*(2i+1) = -\tilde{\mathbf{H}}_1^{\mathcal{H}}\bar{\mathbf{s}}_2(2i) + \tilde{\mathbf{H}}_2^{\mathcal{H}}\bar{\mathbf{s}}_1(2i) + \mathbf{P}\mathbf{w}^*(2i+1). \quad (8)$$

Notice that without the permutation matrix $\mathbf{P}$ inserted at the transmitter, it would have been impossible to have the Hermitian of the channel matrices showing up in (8) that will prove instrumental for enabling multiantenna diversity gains with linear receiver processing.

We will pursue frequency-domain processing of the received blocks, which we describe by multiplying the blocks $\mathbf{x}(i)$ with the FFT matrix $\mathbf{F}_J$ that implements the J-point FFT of the entries in $\mathbf{x}(i)$. Let us define $\mathbf{y}(2i) := \mathbf{F}_J\mathbf{x}(2i)$, $\mathbf{y}^*(2i+1) := \mathbf{F}_J\mathbf{P}\mathbf{x}^*(2i+1)$, and likewise $\bar{\boldsymbol{\eta}}(2i) := \mathbf{F}_J\mathbf{w}(2i)$ and $\bar{\boldsymbol{\eta}}^*(2i+1) := \mathbf{F}_J\mathbf{P}\mathbf{w}^*(2i+1)$. For notational convenience, we also define the diagonal matrices $\boldsymbol{\mathcal{D}}_1 := \mathbf{D}(\tilde{\mathbf{h}}_1)$ and $\boldsymbol{\mathcal{D}}_2 := \mathbf{D}(\tilde{\mathbf{h}}_2)$ with the corresponding transfer function FFT samples on their diagonals. Applying the property p1) on (6) and (8), we obtain the FFT processed blocks as

$$\mathbf{y}(2i) = \boldsymbol{\mathcal{D}}_1\mathbf{F}_J\bar{\mathbf{s}}_1(2i) + \boldsymbol{\mathcal{D}}_2\mathbf{F}_J\bar{\mathbf{s}}_2(2i) + \bar{\boldsymbol{\eta}}(2i) \quad (9)$$

$$\mathbf{y}^*(2i+1) = -\boldsymbol{\mathcal{D}}_1^*\mathbf{F}_J\bar{\mathbf{s}}_2(2i) + \boldsymbol{\mathcal{D}}_2^*\mathbf{F}_J\bar{\mathbf{s}}_1(2i) + \bar{\boldsymbol{\eta}}^*(2i+1). \quad (10)$$

It is important to note at this point that permutation, conjugation, and FFT operations on the received blocks $\mathbf{x}(i)$ do not introduce any information loss, or color the additive noises in (9) and (10) that remain white. It is thus sufficient to rely only on the FFT processed blocks $\mathbf{y}(2i)$ and $\mathbf{y}^*(2i+1)$ when performing symbol detection.

After defining $\check{\mathbf{y}}(i) := [\mathbf{y}^T(2i),\ \mathbf{y}^{\mathcal{H}}(2i+1)]^T$, we can combine (9) and (10) into a single block matrix–vector form to obtain

$$\check{\mathbf{y}}(i) = \underbrace{\begin{bmatrix} \boldsymbol{\mathcal{D}}_1 & \boldsymbol{\mathcal{D}}_2 \\ \boldsymbol{\mathcal{D}}_2^* & -\boldsymbol{\mathcal{D}}_1^* \end{bmatrix}}_{:=\boldsymbol{\mathcal{D}}} \begin{bmatrix} \mathbf{F}_J \mathbf{s}(2i) \\ \mathbf{F}_J \mathbf{s}(2i+1) \end{bmatrix} + \begin{bmatrix} \overline{\boldsymbol{\eta}}(2i) \\ \overline{\boldsymbol{\eta}}^*(2i+1) \end{bmatrix} \quad (11)$$

where the identities $\bar{\mathbf{s}}_1(2i) = \mathbf{s}(2i)$ and $\bar{\mathbf{s}}_2(2i) = \mathbf{s}(2i+1)$ have been used following our design in (1).

Consider a $J \times J$ diagonal matrix $\overline{\boldsymbol{\mathcal{D}}}_{12}$ with nonnegative diagonal entries as $\overline{\boldsymbol{\mathcal{D}}}_{12} = [\boldsymbol{\mathcal{D}}_1^* \boldsymbol{\mathcal{D}}_1 + \boldsymbol{\mathcal{D}}_2^* \boldsymbol{\mathcal{D}}_2]^{1/2}$. We can verify that the matrix $\boldsymbol{\mathcal{D}}$ in (11) satisfies $\boldsymbol{\mathcal{D}}^{\mathcal{H}} \boldsymbol{\mathcal{D}} = \mathbf{I}_2 \otimes \overline{\boldsymbol{\mathcal{D}}}_{12}^2$, where $\otimes$ stands for Kronecker product. Based on $\boldsymbol{\mathcal{D}}_1$ and $\boldsymbol{\mathcal{D}}_2$, we next construct a unitary matrix $\mathbf{U}$. If $\mathbf{h}_1$ and $\mathbf{h}_2$ do not share common zeros on the FFT grid $\{e^{j\frac{2\pi}{J}n}\}_{n=0}^{J-1}$, then $\overline{\boldsymbol{\mathcal{D}}}_{12}$ is invertible, and we select $\mathbf{U}$ as $\mathbf{U} := \boldsymbol{\mathcal{D}}(\mathbf{I}_2 \otimes \overline{\boldsymbol{\mathcal{D}}}_{12}^{-1})$. If $\mathbf{h}_1$ and $\mathbf{h}_2$ happen to share common zero(s) on the FFT grid (although this event has probability zero), then we construct $\mathbf{U}$ as follows. Supposing without of loss of generality that $\mathbf{h}_1$ and $\mathbf{h}_2$ share a common zero at the first subcarrier e^{j0}, we have that

$$[\boldsymbol{\mathcal{D}}_1]_{1,1} = [\boldsymbol{\mathcal{D}}_2]_{1,1} = [\overline{\boldsymbol{\mathcal{D}}}_{12}]_{1,1} = 0.$$

We then construct a diagonal matrix $\boldsymbol{\mathcal{D}}_1'$ which differs from $\boldsymbol{\mathcal{D}}_1$ only at the first diagonal entry: $[\boldsymbol{\mathcal{D}}_1']_{1,1} = 1$. Similar to the definition of $\boldsymbol{\mathcal{D}}$ and $\overline{\boldsymbol{\mathcal{D}}}_{12}$, we construct $\boldsymbol{\mathcal{D}}'$ and $\overline{\boldsymbol{\mathcal{D}}}_{12}'$ by substituting $\boldsymbol{\mathcal{D}}_1$ with $\boldsymbol{\mathcal{D}}_1'$. Because $\overline{\boldsymbol{\mathcal{D}}}_{12}'$ is invertible, we form $\mathbf{U} := \boldsymbol{\mathcal{D}}'[\mathbf{I}_2 \otimes (\overline{\boldsymbol{\mathcal{D}}}_{12}')^{-1}]$. In summary, no matter whether $\overline{\boldsymbol{\mathcal{D}}}_{12}$ is invertible or not, we can always construct a unitary $\mathbf{U}$, which satisfies $\mathbf{U}^{\mathcal{H}}\mathbf{U} = \mathbf{I}_{2J}$ and $\mathbf{U}^{\mathcal{H}}\boldsymbol{\mathcal{D}} = \mathbf{I}_2 \otimes \overline{\boldsymbol{\mathcal{D}}}_{12}$, where the latter can be easily verified. As multiplying by unitary matrices does not incur any loss of decoding optimality in the presence of additive white Gaussian noise, (11) yields $\check{\mathbf{z}}(i) := [\mathbf{z}^T(2i),\ \mathbf{z}^T(2i+1)]^T$ as

$$\begin{aligned} \check{\mathbf{z}}(i) &= \mathbf{U}^{\mathcal{H}} \check{\mathbf{y}}(i) \\ &= \begin{bmatrix} \overline{\boldsymbol{\mathcal{D}}}_{12} \mathbf{F}_J \mathbf{s}(2i) \\ \overline{\boldsymbol{\mathcal{D}}}_{12} \mathbf{F}_J \mathbf{s}(2i+1) \end{bmatrix} + \mathbf{U}^{\mathcal{H}} \begin{bmatrix} \overline{\boldsymbol{\eta}}(2i) \\ \overline{\boldsymbol{\eta}}^*(2i+1) \end{bmatrix} \end{aligned} \quad (12)$$

where the resulting noise

$$\check{\boldsymbol{\eta}}(i) := [\boldsymbol{\eta}^T(2i),\ \boldsymbol{\eta}^T(2i+1)]^T = \mathbf{U}^{\mathcal{H}}[\overline{\boldsymbol{\eta}}^T(2i),\ \overline{\boldsymbol{\eta}}^{\mathcal{H}}(2i+1)]^T$$

is still white with each entry having variance N_0.

We infer from (12) that the blocks $\mathbf{s}(2i)$ and $\mathbf{s}(2i+1)$ can be demodulated separately without compromising the ML optimality, after linear receiver processing. Indeed, so far we applied at the receiver three linear unitary operations after the CP removal: i) permutation (via $\mathbf{P}$); ii) conjugation and FFT (via $\mathbf{F}_J$); and iii) unitary combining (via $\mathbf{U}^{\mathcal{H}}$). As a result, we only need to demodulate each information block $\mathbf{d}(i)$ separately from the following subblocks [cf. (12)]:

$$\mathbf{z}(i) = \overline{\boldsymbol{\mathcal{D}}}_{12} \mathbf{F}_J \mathbf{s}(i) + \boldsymbol{\eta}(i) = \overline{\boldsymbol{\mathcal{D}}}_{12} \mathbf{F}_J \boldsymbol{\Theta} \mathbf{d}(i) + \boldsymbol{\eta}(i). \quad (13)$$

Before specifying equalization and decoding possibilities (that will be discussed in Section III), we will first go after the benchmark performance with ML decoding. Starting from (13), we will study the diversity gains of existing and our novel unifying ST schemes in the presence of frequency-selective channels; this analysis is not available in the single-carrier approaches of [2], [22], [42].

A. Diversity Gain Analysis

Let us drop the block index i from (13), and, e.g., use $\mathbf{d}$ to denote $\mathbf{d}(i)$ for notational brevity. With perfect channel state information (CSI) at the receiver, we will consider the pairwise error probability (PEP) $P(\mathbf{d} \to \mathbf{d}'|\mathbf{h}_1, \mathbf{h}_2)$ that the symbol block $\mathbf{d}$ is transmitted, but is erroneously decoded as $\mathbf{d}' \neq \mathbf{d}$. The PEP can be approximated using the Chernoff bound as

$$P(\mathbf{s} \to \mathbf{s}'|\mathbf{h}_1, \mathbf{h}_2) \leq \exp(-d^2(\mathbf{z}, \mathbf{z}')/4N_0) \quad (14)$$

where $d(\mathbf{z}, \mathbf{z}')$ denotes the Euclidean distance between $\mathbf{z}$ and $\mathbf{z}'$.

Define the error vector as $\mathbf{e} := \mathbf{d} - \mathbf{d}'$, and a $J \times (L+1)$ Vandermonde matrix $\mathbf{V}$ with $[\mathbf{V}]_{p,q} = \exp(-j2\pi pq/J)$. The matrix $\mathbf{V}$ links the channel frequency response with the time-domain channel taps as $\tilde{\mathbf{h}}_\mu = \mathbf{V}\mathbf{h}_\mu$. Starting with (13), we then express the distance as

$$\begin{aligned} d^2(\mathbf{z}, \mathbf{z}') &= \left\|\overline{\boldsymbol{\mathcal{D}}}_{12} \mathbf{F}_J \boldsymbol{\Theta} \mathbf{e}\right\|^2 = \mathbf{e}^{\mathcal{H}} \boldsymbol{\Theta}^{\mathcal{H}} \mathbf{F}_J^{\mathcal{H}} \overline{\boldsymbol{\mathcal{D}}}_{12}^2 \mathbf{F}_J \boldsymbol{\Theta} \mathbf{e} \\ &= \sum_{\mu=1}^{2} \|\boldsymbol{\mathcal{D}}_\mu \mathbf{F}_J \boldsymbol{\Theta} \mathbf{e}\|^2 = \sum_{\mu=1}^{2} \|\mathbf{D}_e \mathbf{V} \mathbf{h}_\mu\|^2 \end{aligned} \quad (15)$$

where $\mathbf{D}_e := \mathrm{diag}(\mathbf{F}_J \boldsymbol{\Theta} \mathbf{e})$ such that

$$\boldsymbol{\mathcal{D}}_\mu \mathbf{F}_J \boldsymbol{\Theta} \mathbf{e} = \mathbf{D}_e \tilde{\mathbf{h}}_\mu = \mathbf{D}_e \mathbf{V} \mathbf{h}_\mu.$$

We focus on block quasi-static channels, i.e., channels that remain invariant over each ST coded block, but may vary from one block to the next. We further adopt the following assumption:

as0) the channels $\mathbf{h}_1$ and $\mathbf{h}_2$ are uncorrelated; and for each antenna $\mu \in [1, 2]$, the channel $\mathbf{h}_\mu$ is zero-mean, complex Gaussian distributed, with covariance matrix $\mathbf{R}_{h,\mu} := \mathrm{E}\{\mathbf{h}_\mu \mathbf{h}_\mu^{\mathcal{H}}\}$.

If the entries of $\mathbf{h}_\mu$ are independent and identically distributed (i.i.d.), then we have $\mathbf{R}_{h,\mu} = \mathbf{I}_{L+1}/(L+1)$, where the channel covariance matrix is normalized to have unit energy; i.e., $\mathrm{tr}\{\mathbf{R}_{h,\mu}\} = 1$. Because general frequency-selective multipath channels have covariance matrices with arbitrary rank, we define the "effective channel order" as $\tilde{L}_\mu = \mathrm{rank}(\mathbf{R}_{h,\mu}) - 1$. Let us consider now the following eigendecomposition:

$$\mathbf{R}_{h,\mu} = \mathbf{U}_{h,\mu} \boldsymbol{\Lambda}_{h,\mu} \mathbf{U}_{h,\mu}^{\mathcal{H}} \quad (16)$$

where $\boldsymbol{\Lambda}_{h,\mu}$ is an $(\tilde{L}_\mu + 1) \times (\tilde{L}_\mu + 1)$ diagonal matrix with the positive eigenvalues of $\mathbf{R}_{h,\mu}$ on its diagonal, and $\mathbf{U}_{h,\mu}$ is an $(L+1) \times (\tilde{L}_\mu + 1)$ matrix having orthonormal columns $\mathbf{U}_{h,\mu}^{\mathcal{H}} \mathbf{U}_{h,\mu} = \mathbf{I}_{\tilde{L}_\mu + 1}$. Defining $\overline{\mathbf{h}}_\mu = \boldsymbol{\Lambda}_{h,\mu}^{-\frac{1}{2}} \mathbf{U}_{h,\mu}^{\mathcal{H}} \mathbf{h}_\mu$, we can verify that the entries of $\overline{\mathbf{h}}_\mu$ are i.i.d. with unit variance. Since $\mathbf{h}_\mu$ and $\mathbf{U}_{h,\mu} \boldsymbol{\Lambda}_{h,\mu}^{\frac{1}{2}} \overline{\mathbf{h}}_\mu$ have identical distributions, we replace the former by the latter in the ensuing PEP analysis. A special case of interest corresponds to transmissions experiencing channels with full-rank correlation matrices; i.e., $\mathrm{rank}(\mathbf{R}_{h,\mu}) = L+1$ and $\tilde{L}_\mu = L$. As will be clear later, a rich scattering environment leads to $\mathbf{R}_{h,\mu}$'s with full rank, which is favorable in broad-band wireless applications because it is also rich in diversity.

With the aid of the whitened and normalized channel vector $\overline{\mathbf{h}}_\mu$, we can simplify (15) to

$$d^2(\mathbf{z},\mathbf{z}') = \left\|\mathbf{D}_e\mathbf{V}\mathbf{U}_{h,1}\mathbf{\Lambda}_{h,1}^{\frac{1}{2}}\overline{\mathbf{h}}_1\right\|^2 + \left\|\mathbf{D}_e\mathbf{V}\mathbf{U}_{h,2}\mathbf{\Lambda}_{h,2}^{\frac{1}{2}}\overline{\mathbf{h}}_2\right\|^2. \tag{17}$$

From the spectral decomposition of the matrix $\mathbf{A}_{e,\mu}^{\mathcal{H}}\mathbf{A}_{e,\mu}$, where $\mathbf{A}_{e,\mu} := \mathbf{D}_e\mathbf{V}\mathbf{U}_{h,\mu}\mathbf{\Lambda}_{h,\mu}^{\frac{1}{2}}$, we know that there exists a unitary matrix $\mathbf{U}_{e,\mu}$ such that $\mathbf{U}_{e,\mu}^{\mathcal{H}}\mathbf{A}_{e,\mu}^{\mathcal{H}}\mathbf{A}_{e,\mu}\mathbf{U}_{e,\mu} = \mathbf{\Lambda}_{e,\mu}$, where $\mathbf{\Lambda}_{e,\mu}$ is diagonal with nonincreasing diagonal entries collected in the vector

$$\boldsymbol{\lambda}_{e,\mu} := [\lambda_{e,\mu}(0),\, \lambda_{e,\mu}(1),\, \ldots,\, \lambda_{e,\mu}(\tilde{L}_\mu)]^T.$$

Consider now the channel vectors $\overline{\mathbf{h}}'_\mu := \mathbf{U}_{e,\mu}^{\mathcal{H}}\overline{\mathbf{h}}_\mu$, with identity correlation matrix. The vector $\overline{\mathbf{h}}'_\mu$ is clearly zero mean, complex Gaussian, with i.i.d entries. Using $\overline{\mathbf{h}}_\mu$, we can rewrite (17) as

$$\begin{aligned} d^2(\mathbf{z},\mathbf{z}') &= \sum_{\mu=1}^{2} \left(\overline{\mathbf{h}}'_\mu\right)^{\mathcal{H}} \mathbf{U}_{e,\mu}^{\mathcal{H}}\mathbf{A}_{e,\mu}^{\mathcal{H}}\mathbf{A}_{e,\mu}\mathbf{U}_{e,\mu}\overline{\mathbf{h}}'_\mu \\ &= \sum_{l=1}^{\tilde{L}_1}\lambda_{e,1}(l)\left|\overline{h}'_1(l)\right|^2 + \sum_{l=1}^{\tilde{L}_2}\lambda_{e,2}(l)\left|\overline{h}'_2(l)\right|^2. \end{aligned} \tag{18}$$

Based on (18), and by averaging (14) with respect to the i.i.d. Rayleigh random variables $|\overline{h}'_1(l)|$, $|\overline{h}'_2(l)|$, we can upper-bound the average PEP as follows:

$$P(\mathbf{s}\to\mathbf{s}') \le \prod_{l=0}^{\tilde{L}_1}\frac{1}{1+\lambda_{e,1}(l)/(4N_0)}\prod_{l=0}^{\tilde{L}_2}\frac{1}{1+\lambda_{e,2}(l)/(4N_0)}. \tag{19}$$

If $r_{e,\mu}$ is the rank of $\mathbf{A}_{e,\mu}$ (and thus the rank of $\mathbf{A}_{e,\mu}^{\mathcal{H}}\mathbf{A}_{e,\mu}$), then $\lambda_{e,\mu}(l) \neq 0$ if and only if $l \in [0,\, r_{e,\mu} - 1]$. It thus follows from (19) that

$$\begin{aligned} P(\mathbf{s}\to\mathbf{s}') \le{}& \left(\frac{1}{4N_0}\right)^{-(r_{e,1}+r_{e,2})} \\ &\times \left(\prod_{l=0}^{r_{e,1}-1}\lambda_{e,1}(l)\prod_{l=0}^{r_{e,2}-1}\lambda_{e,2}(l)\right)^{-1}. \end{aligned} \tag{20}$$

As in [38], we call $r_e := r_{e,1} + r_{e,2}$ the diversity gain $G_{d,e}$, and

$$\left[\prod_{l=0}^{r_{e,1}-1}\lambda_{e,1}(l)\prod_{l=0}^{r_{e,2}-1}\lambda_{e,2}(l)\right]^{1/r_e}$$

the coding gain $G_{c,e}$ of the system for a given symbol error vector $\mathbf{e}$. The diversity gain $G_{d,e}$ determines the slope of the averaged [with respect to (w.r.t.) the random channel] PEP (between $\mathbf{s}$ and $\mathbf{s}'$) as a function of the signal-to-noise ratio (SNR) at high SNR ($N_0 \to 0$). Correspondingly, $G_{c,e}$ determines the shift of this PEP curve in SNR relative to a benchmark error rate curve of $[1/(4N_0)]^{-r_e}$. Different from [19], [38], that relied on PEP to design (nonlinear) ST codes for flat fading channels, we here invoke PEP bounds to prove diversity properties of our proposed single-carrier block transmissions over frequency-selective channels.

Since both $G_{d,e}$ and $G_{c,e}$ depend on the choice of $\mathbf{e}$ (thus, on $\mathbf{s}$ and $\mathbf{s}'$), we define the diversity and coding gains for our system, respectively, as

$$G_d := \min_{\mathbf{e}\neq 0} G_{d,e} \qquad \text{and} \qquad G_c := \min_{\mathbf{e}\neq 0} G_{c,e}. \tag{21}$$

Based on (21), one can check both diversity and coding gains. However, in this paper, we focus only on the diversity gain. First, we observe that the matrix $\mathbf{A}_{e,\mu}^{\mathcal{H}}\mathbf{A}_{e,\mu}$ is square of size $(\tilde{L}_\mu + 1)$. Therefore, the maximum achievable diversity gain in a two transmit and one receive antennas system is $G_d = \sum_{\mu=1}^{2}(\tilde{L}_\mu + 1)$ for FIR channels with effective channel order $\tilde{L}_\mu$, $\mu = 1, 2$, while it becomes $2(L+1)$ in rich scattering environments. This maximum diversity can be easily achieved by e.g., a simple redundant transmission where each antenna transmits the same symbol followed by L zeros in two nonoverlapping time slots. We next examine the achieved diversity levels in our following proposed schemes, which certainly have much higher rate than redundant transmissions.

B. CP-Only

We term CP-only the block transmissions with no precoding: $\mathbf{\Theta} = \mathbf{I}_K$, $J = K$, and $\mathbf{s}(i) = \mathbf{d}(i)$. The word "only" emphasizes that, unlike OFDM, no IFFT is applied at the transmitter. Let us now check the diversity order achieved by CP-only. The worst case is to select $\mathbf{d} = a\mathbf{1}_{J\times 1}$ and $\mathbf{d}' = a'\mathbf{1}_{J\times 1}$ implying $\mathbf{e} = (a - a')\mathbf{1}_{J\times 1}$, where $a,\, a' \in \mathcal{A}$. Verifying that for these error events, the matrix $\mathbf{D}_e = \mathrm{diag}(\mathbf{F}_J\mathbf{e})$ has only one nonzero entry, we deduce that $r_{e,1} = r_{e,2} = 1$. Therefore, the system diversity order achieved by CP-only is $G_d = 2$. This is nothing but space diversity of order two coming from the two transmit antennas [cf. (13)]. Note that CP-only schemes (as those in [2], [42]) suffer from loss of multipath diversity.

To benefit also from the embedded multipath-induced diversity, we have to modify our transmissions.

C. Linearly Precoded CP-Only

To increase our ST system's diversity order, we will adopt the linear precoding ideas developed originally for single-antenna transmissions in [44], [45]. One can view CP-only as a special case of the linearly precoded CP-only system (denoted henceforth as LP-CP-only) with identity precoder. With $\mathbf{s}(i) = \mathbf{\Theta}\mathbf{d}(i)$ and carefully designed $\mathbf{\Theta} \neq \mathbf{I}_K$, we next show that the maximum diversity is achieved. We will discuss two cases: the first one introduces no redundancy because it uses $J = K$, while the second one is redundant and adopts $J = K + L$.

For nonredundant precoding with $J = K$, it has been established that for *any* signal constellation adhering to a finite alphabet, there always exists a $K \times K$ unitary constellation rotating (CR) matrix $\mathbf{\Theta}_{\mathrm{cr}}$ ensuring that each entry of $\mathbf{\Theta}_{\mathrm{cr}}(\mathbf{d} - \mathbf{d}')$ is nonzero *for any* pair of $(\mathbf{d},\, \mathbf{d}')$ [46]. We thus propose to construct $\mathbf{\Theta} = \mathbf{F}_K^{\mathcal{H}}\mathbf{\Theta}_{\mathrm{cr}}$ such that $\mathbf{F}_K\mathbf{\Theta} = \mathbf{\Theta}_{\mathrm{cr}}$. With this construction, $\mathbf{D}_e = \mathrm{diag}(\mathbf{\Theta}_{\mathrm{cr}}\mathbf{e})$ is guaranteed to have nonzero entries on its diagonal, and thus it has full rank. Consequently, the matrix $\mathbf{D}_e\mathbf{V}$ has full column rank $L+1$, and $\mathbf{A}_{e,\mu} = \mathbf{D}_e\mathbf{V}\mathbf{U}_{h,\mu}\mathbf{\Lambda}_{h,\mu}^{\frac{1}{2}}$ has full column rank $r_{e,\mu} = \tilde{L}_\mu + 1$. Hence, the maximum achievable diversity order is indeed achieved.

We emphasize here that the nonredundant precoder $\mathbf{\Theta}_{\rm cr}$ is constellation dependent. For commonly used binary phase shift keying (BPSK), quaternary phase shift keying (QPSK), and all quadrature amplitude modulation (QAM) constellations, and for the block size K equal to a power of 2 : $K = 2^m$, one class of $\mathbf{\Theta}_{\rm cr}$ precoders with large coding gains is found to be [9], [17], [46]

$$\mathbf{\Theta}_{\rm cr} = \mathbf{F}_K \mathbf{\Delta}(\alpha) \quad \text{and, thus,} \quad \mathbf{\Theta} = \mathbf{\Delta}(\alpha) \tag{22}$$

where

$$\mathbf{\Delta}(\alpha) := \text{diag}(1,\, \alpha,\, \ldots,\, \alpha^{K-1})$$

with the scalar $\alpha \in \{e^{j\frac{\pi}{2K}(1+4n)}\}_{n=0}^{K-1}$. For block size $K \neq 2^m$, one can construct $\mathbf{\Theta}_{\rm cr}$ by truncating a larger unitary matrix constructed as in (22) [46]. The price paid for our increased diversity gain is that LP-CP-only does not offer constant modulus transmissions, in general. However, by designing K to be a power of 2, and by choosing $\mathbf{\Theta}$ as in (22), the transmitted signals $\mathbf{s}(i) = \mathbf{\Delta}(\alpha)\mathbf{d}(i)$ are constant modulus if $\mathbf{d}(i)$ are phase shift keying (PSK) signals. Therefore, by selecting K to be a power of 2, we can increase the diversity gain without reducing power efficiency.

Alternatively, we can adopt a redundant $J \times K$ precoder $\mathbf{\Theta}$ with $J = K + L$. Our criterion for selecting such tall precoding matrices $\mathbf{\Theta}$ is to guarantee that $\mathbf{F}_J\mathbf{\Theta}$ satisfies the following property: *any* K rows of $\mathbf{F}_J\mathbf{\Theta}$ are linearly independent. One class of $\mathbf{F}_J\mathbf{\Theta}$ satisfying this property includes Vandermonde matrices $\mathbf{\Theta}_{\rm van}$ with distinct generators $[\rho_1, \ldots, \rho_J]$, defined as [44], [45]

$$\mathbf{\Theta}_{\rm van} = \frac{1}{\sqrt{J}} \begin{bmatrix} 1 & \rho_1^{-1} & \cdots & \rho_1^{-(K-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \rho_J^{-1} & \cdots & \rho_J^{-(K-1)} \end{bmatrix}$$

and thus

$$\mathbf{\Theta} = \mathbf{F}_J^{\mathcal{H}} \mathbf{\Theta}_{\rm van}. \tag{23}$$

With $\mathbf{F}_J\mathbf{\Theta} = \mathbf{\Theta}_{\rm van}$, we have that $\mathbf{\Theta}_{\rm van}\mathbf{e}$ has at least $(L+1)$ nonzero entries for any $\mathbf{e}$ regardless of the underlying signal constellation. Indeed, if $\mathbf{\Theta}_{\rm van}\mathbf{e}$ has only L nonzero entries for some $\mathbf{e}$, then it has K zero entries. Picking the corresponding K rows of $\mathbf{\Theta}_{\rm van}$ to form the truncated matrix $\overline{\mathbf{\Theta}}_{\rm van}$, we have $\overline{\mathbf{\Theta}}_{\rm van}\mathbf{e} = \mathbf{0}$, which shows that these K rows are linearly dependent, thus violating the design of the precoder $\mathbf{\Theta}_{\rm van}$. With $\mathbf{D}_e = \text{diag}(\mathbf{\Theta}_{\rm van}\mathbf{e})$ having at least $(L+1)$ nonzero entries, the matrix $\mathbf{D}_e\mathbf{V}$ has full rank because *any* $L+1$ rows of $\mathbf{V}$ are linearly independent. Thus, the maximum diversity gain is achieved with redundant precoding *irrespective* of the underlying constellation.

When $J \in [K,\, K+L]$, constellation-irrespective precoders are impossible because $\mathbf{\Theta}\mathbf{e}$ cannot have $L+1$ nonzero entries for any $\mathbf{e}$ that is unconstrained. Therefore, constellation-independent precoders are not possible for $J < K + L$. However, with some redundancy $J > K$, the design of constellation-dependent precoders may become easier. Optimal design of nonredundant or redundant precoders that also maximize coding gains within the class of our maximum diversity achieving precoders, is certainly an interesting research topic, but will not be pursued in this paper.

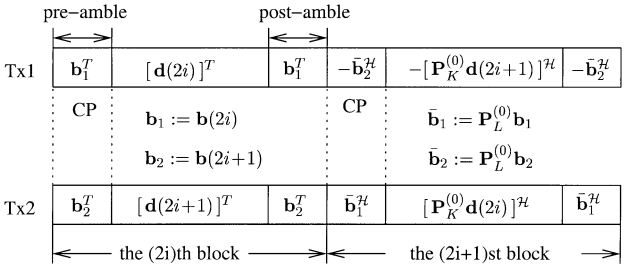

Fig. 3. Affine precoded CP-only, with $\mathbf{s}(i) = [\mathbf{d}^T(i), \mathbf{b}^T(i)]^T$ and $\mathbf{P} = \mathbf{P}_J^{(K)}$.

D. Affine Precoded CP-Only

Another interesting class of linear precoders implements an affine transformation $\mathbf{s}(i) = \mathbf{\Theta}\mathbf{d}(i) + \mathbf{\Theta}'\mathbf{b}(i)$, where $\mathbf{b}(i)$ is a known symbol vector. In this paper, we are only interested in the special form of

$$\mathbf{s}(i) = \mathbf{T}_1\mathbf{d}(i) + \mathbf{T}_2\mathbf{b}(i) = \begin{bmatrix} \mathbf{d}(i) \\ \mathbf{b}(i) \end{bmatrix} \tag{24}$$

where the precoder $\mathbf{\Theta} = \mathbf{T}_1$ is the first K columns of $\mathbf{I}_J$, the precoder $\mathbf{\Theta}' = \mathbf{T}_2$ is the last L columns of $\mathbf{I}_J$, and the known symbol vector $\mathbf{b}$ has size $L \times 1$ with entries drawn from the same alphabet $\mathcal{A}$. We henceforth term the transmission format in (24) as AP-CP-only. Notice that in this scheme, $J = K + L$ and $P = J + L$.

Although here we place $\mathbf{b}(i)$ at the bottom of $\mathbf{s}(i)$ for convenience, we could also place $\mathbf{b}(i)$ at arbitrary positions within $\mathbf{s}(i)$. As long as L consecutive symbols are known in $\mathbf{s}(i)$, all decoding schemes detailed in Section III are applicable.

Recall that the error matrix $\mathbf{D}_e = \text{diag}(\mathbf{F}_J\mathbf{T}_1\mathbf{e})$ does not contain known symbols. Since $\mathbf{F}_J\mathbf{T}_1$ is a Vandermonde matrix of the form of the matrix in Section II-C, the maximum diversity gain is achieved, as discussed in Section II-C for redundant LP-CP-only.

In the CP-based schemes depicted in Fig. 2, the CP portion of the transmitted sequence is generally unknown, because it is replicated from the unknown data blocks. However, with AP-CP-only in (24), and with the specific choice of $\mathbf{P} = \mathbf{P}_J^{(K)}$, we have

$$\mathbf{P}_J^{(K)}\mathbf{s}(i) = \left[[\mathbf{P}_K^{(0)}\mathbf{d}(i)\right]^T,\, \left[\mathbf{P}_L^{(0)}\mathbf{b}(i)]^T\right]^T$$

which implies that *both* the data block and the known symbol block are time reversed, but keep their original positions. The last L entries of $\mathbf{P}_J^{(K)}\mathbf{s}(i)$ are again known, and are then replicated as cyclic prefixes. For this *special case*, we depict the transmitted sequences in Fig. 3. In this format, the data block $\mathbf{d}(i)$ is surrounded by two known blocks that correspond to the pre-amble and post-amble in [22]. Our general design based on the CP structure includes this known pre- and post-ambles as a special case.

Notice that the pre-amble and post-amble have not been properly designed in [22]. The consequence is that "edge effects" appear for transmissions with finite block length, and an approximation on the order of $\mathcal{O}(L/J)$ has to be made in order to apply

TABLE I
SUMMARY OF SINGLE-CARRIER SCHEMES IN RICH-SCATTERING ENVIRONMENTS

	Rate R	Diversity G_d	Power Loss (dB)	Features
CP-only	$\frac{K}{K+L}\log_2\lvert\mathcal{A}\rvert$	2	$10\log_{10}\frac{K+L}{K}$	constant modulus (C-M)*
non-redundant LP-CP-only	$\frac{K}{K+L}\log_2\lvert\mathcal{A}\rvert$	$2(L+1)$	$10\log_{10}\frac{K+L}{K}$	constellation-specific precoder constant modulus
redundant LP-CP-only	$\frac{K}{K+2L}\log_2\lvert\mathcal{A}\rvert$	$2(L+1)$	$10\log_{10}\frac{K+L}{K}$	constellation-independent Not C-M in general
AP-CP-only	$\frac{K}{K+2L}\log_2\lvert\mathcal{A}\rvert$	$2(L+1)$	$10\log_{10}\frac{K+2L}{K}$	constellation-independent constant modulus
ZP-only	$\frac{K}{K+L}\log_2\lvert\mathcal{A}\rvert$	$2(L+1)$	0	constellation-independent C-M except zero guards

* only if information symbols have constant-modulus, e.g., drawn from PSK constellations.

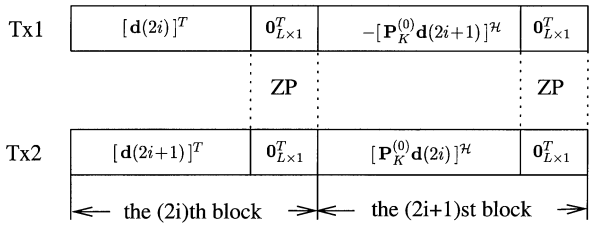

Fig. 4. The transmitted sequence for ZP-only.

Viterbi's decoding algorithm. This approximation amounts to nothing but the fact that a linear convolution can be approximated by a circular convolution when the block size is much larger than the channel order. By simply enforcing a CP structure to obtain circulant convolutions, Viterbi's algorithm can be applied to our proposed AP-CP-only with *no approximation* whatsoever, regardless of the block length and the channel order, as will be clear soon.

E. ZP-Only

Suppose now that in AP-CP-only, we let $\mathbf{b}(i)=\mathbf{0}$ instead of having known symbols drawn from the constellation alphabet, and we fix $\mathbf{P}=\mathbf{P}^{(K)}_J$. Now, the adjacent data blocks are guarded by two zero blocks, each having length L, as depicted in Fig. 3. Since the channel has only order L, presence of $2L$ zeros in the middle of two adjacent data blocks is not necessary. Keeping only a single block of L zeros corresponds to removing the CP-insertion operation at the transmitter. On the other hand, one could view that the zero block in the previous block serves as the CP for the current block, and thus all derivations done for CP-based transmission are still valid. The resulting transmission format is shown in Fig. 4, which achieves higher bandwidth efficiency than AP-CP-only. We term this scheme as ZP-only, where $J=K+L$ and $P=J$.

By mathematically viewing ZP-only as a special case of AP-CP-only with $\mathbf{b}(i)=\mathbf{0}$, it is clear that the maximum diversity is achieved. In addition to the rate improvement, ZP-only also saves the transmitted power occupied by CP and known symbols.

For convenience, we list all aforementioned schemes in Table I, assuming a rich scattering environment. Power loss induced by the cyclic prefix and the known symbols, is also considered. It certainly becomes negligible when $K\gg L$.

F. Links With Multicarrier Transmissions

In this section, we link single carrier with digital multicarrier (OFDM-based) schemes. We first examine the transmitted blocks on two consecutive time intervals. For LP-CP-only, the transmitted ST matrix is

$$\begin{bmatrix}\mathbf{u}_1(2i) & \mathbf{u}_1(2i+1)\\ \mathbf{u}_2(2i) & \mathbf{u}_2(2i+1)\end{bmatrix} = \begin{bmatrix}\mathbf{T}_{\rm cp}\mathbf{\Theta}\mathbf{d}(2i) & -\mathbf{T}_{\rm cp}\mathbf{P}\mathbf{\Theta}^*\mathbf{d}^*(2i+1)\\ \mathbf{T}_{\rm cp}\mathbf{\Theta}\mathbf{d}(2i+1) & \mathbf{T}_{\rm cp}\mathbf{P}\mathbf{\Theta}^*\mathbf{d}^*(2i)\end{bmatrix}. \quad (25)$$

If we let $\mathbf{P}=\mathbf{P}^{(1)}_J$, and $\mathbf{\Theta}=\mathbf{F}^{\mathcal{H}}_J\mathbf{\Psi}$ for a general matrix $\mathbf{\Psi}$, we obtain from (25)

$$\begin{bmatrix}\mathbf{u}_1(2i) & \mathbf{u}_1(2i+1)\\ \mathbf{u}_2(2i) & \mathbf{u}_2(2i+1)\end{bmatrix} = \begin{bmatrix}\mathbf{T}_{\rm cp}\mathbf{F}^{\mathcal{H}}_J\mathbf{\Psi}\mathbf{d}(2i) & -\mathbf{T}_{\rm cp}\mathbf{F}^{\mathcal{H}}_J\mathbf{\Psi}^*\mathbf{d}^*(2i+1)\\ \mathbf{T}_{\rm cp}\mathbf{F}^{\mathcal{H}}_J\mathbf{\Psi}\mathbf{d}(2i+1) & \mathbf{T}_{\rm cp}\mathbf{F}^{\mathcal{H}}_J\mathbf{\Psi}^*\mathbf{d}^*(2i)\end{bmatrix}. \quad (26)$$

If $\mathbf{\Psi}=\mathbf{I}_K$, then (26) corresponds to the ST block-coded OFDM proposed in [20], [27]. Designing $\mathbf{\Psi}\neq\mathbf{I}_K$ introduces linear precoding across OFDM subcarriers, as proposed in [26], [27]. Therefore, LP-CP-only includes linear precoded ST-OFDM as a special case by selecting the precoder $\mathbf{\Phi}$ and the permutation $\mathbf{P}$ appropriately. Although linear precoding was introduced [26], [27] for ST-OFDM systems, the diversity analysis was not provided. The link we introduced here reveals that the maximum diversity gain is also achieved by linearly precoded ST-OFDM with the Vandermonde precoders provided in [26], [27].

Interestingly, linearly precoded OFDM can even be converted to zero padded transmissions. Indeed, choosing $\mathbf{\Psi}$ to be the first K columns of $\mathbf{F}_J$, we obtain the transmitted block as $\mathbf{u}(i)=\mathbf{T}_{\rm cp}\mathbf{F}^{\mathcal{H}}_J\mathbf{\Psi}\mathbf{d}(i)=[\mathbf{0}^T_{L\times 1},\,\mathbf{d}^T(i),\,\mathbf{0}^T_{L\times 1}]^T$ which inserts zeros both at the top and at the bottom of each data block. By converting linearly precoded OFDM to a zero-padded transmission,

the special multicarrier multiple-access scheme in [44] offers constant modulus transmissions.

The design example presented in [7] for ST-coded OFDM happens to fall into this category, too. In the design example of [7], $J = 8$ subcarriers are used for OFDM, and the channel order is $L = 1$. Two symbols s_1 and s_2 are linearly transformed and transmitted as follows: the first antenna transmits $s_1\mathbf{f}_0 + s_2\mathbf{f}_2$ on its eight subcarriers, and the second antenna transmits $s_1\mathbf{f}_4 + s_2\mathbf{f}_6$ on its eight subcarriers, where $\mathbf{f}_p$ is the $(p+1)$st column of $\mathbf{F}_8$ with $p \in [0, 2, 4, 6]$. Carrying out the calculation for the transmitted sequence $\mathbf{u}_\mu(i)$, the first antenna transmits $0, s_1, 0, s_2, 0, 0, 0, 0, 0$, while the second antenna transmits $0, 0, 0, 0, 0, s_1, 0, s_2, 0$, which amounts to nonoverlapping transmissions from two antennas, with each antenna transmitting each symbol followed by L zeros. This simple example achieves full diversity, but suffers significant rate loss, due to the nature of repeated transmissions, and the absence of symbol blocking that is used in [44].

G. Capacity Result

We now analyze the capacity of the ST block-coding format of (1). The equivalent channel input–output relationship, after receiver processing, is described by (13) as $\mathbf{z} = \overline{\boldsymbol{\mathcal{D}}}_{12}\mathbf{F}_J\mathbf{s} + \boldsymbol{\eta}$, where we drop the block index for brevity. Let $\mathcal{I}(\mathbf{z};\mathbf{s})$ denote the mutual information between $\mathbf{z}$ and $\mathbf{s}$, and recall that $\mathcal{I}(\mathbf{z};\mathbf{s})$ is maximized when $\mathbf{s}$ is Gaussian distributed [13]. Due to the lack of channel knowledge at the transmitter, the transmission power is equally distributed among symbols, with $\mathbf{R}_s = \mathrm{E}\{\mathbf{s}\mathbf{s}^{\mathcal{H}}\} = \sigma_s^2\mathbf{I}_J$. Taking into account the CP of length L, the channel capacity, for a fixed channel realization, is thus

$$\begin{aligned} C_J &= \frac{1}{J+L}\max\mathcal{I}(\mathbf{z};\mathbf{s}) \\ &= \frac{1}{J+L}\log_2\det\left(\mathbf{I}_J + \frac{\sigma_s^2}{N_0}\overline{\boldsymbol{\mathcal{D}}}_{12}\mathbf{F}_J\mathbf{F}_J^{\mathcal{H}}\overline{\boldsymbol{\mathcal{D}}}_{12}^{\mathcal{H}}\right) \\ &= \frac{1}{J+L}\sum_{n=0}^{J-1}\log_2\left(1 + \frac{\sigma_s^2}{N_0}\left(\left|H_1\left(e^{j\frac{2\pi n}{J}}\right)\right|^2 \right.\right. \\ &\qquad \left.\left. + \left|H_2\left(e^{j\frac{2\pi n}{J}}\right)\right|^2\right)\right). \end{aligned} \tag{27}$$

Define $E_s = 2\sigma_s^2$ as the total transmitted power from two antennas per channel use. As the block size J increases, we obtain

$$C_{J\to\infty} = \int_0^1 \log_2\left(1 + \frac{E_s}{2N_0}\left(\left|H_1\left(e^{j2\pi f}\right)\right|^2 + \left|H_2\left(e^{j2\pi f}\right)\right|^2\right)\right) df. \tag{28}$$

The capacity for frequency-selective channels with multiple transmit and receive antennas can be found in, e.g., [23]. The result in (28) coincides with that of [23] when we have two transmit antennas and one receive antenna. Therefore, our proposed transmission format in (1) does not incur capacity loss in this special case. This is consistent with [16], [33], where the Alamouti coding [3] is shown to achieve capacity for frequency-flat fading channels with such an antenna configuration. To achieve capacity for systems with two transmit antennas and a single receive antenna, it thus suffices to deploy suitable one-dimensional channel codes, or scalar codes [16].

III. Equalization and Decoding

Let $\overline{\mathbf{z}}(i) := \mathbf{z}(i)$ for CP-only, LP-CP-only, ZP-only, and $\overline{\mathbf{z}}(i) := \mathbf{z}(i) - \overline{\boldsymbol{\mathcal{D}}}_{12}\mathbf{F}_J\mathbf{T}_2\mathbf{b}(i)$ for AP-CP-only. With this convention, we can unify the equivalent system output after the linear receiver processing as

$$\overline{\mathbf{z}}(i) = \overline{\boldsymbol{\mathcal{D}}}_{12}\mathbf{F}_J\boldsymbol{\Theta}\mathbf{d}(i) + \boldsymbol{\eta}(i) = \mathbf{A}\mathbf{d}(i) + \boldsymbol{\eta}(i) \tag{29}$$

where $\mathbf{A} := \overline{\boldsymbol{\mathcal{D}}}_{12}\mathbf{F}_J\boldsymbol{\Theta}$, the noise $\boldsymbol{\eta}(i)$ is white with covariance $\sigma_w^2\mathbf{I}_J$, and the corresponding $\boldsymbol{\Theta}$ is defined as in Section II.

Brute-force ML decoding applied to (29) requires $|\mathcal{A}|^K$ enumerations, which becomes certainly prohibitive as the constellation size $|\mathcal{A}|$ and/or the block length K increases. A relatively faster near-ML search is possible with the sphere decoding (SD) algorithm, which only searches for vectors that are within a sphere centered at the received symbols [41]. The theoretical complexity of SD is polynomial in K, which is lower than exponential, but still too high for $K > 16$. Only when the block size K is small, the SD equalizer can be adopted to achieve near-ML performance at a manageable complexity. The unique feature of SD is that the complexity does not depend on the constellation size. Thus, SD is suitable for systems with small block size K, but with large signal constellations.

We now turn our attention to low-complexity equalizers by trading off performance with complexity. Linear zero forcing (ZF) and minimum mean square error (MMSE) block equalizers certainly offer low complexity alternatives. The block MMSE equalizer is

$$\boldsymbol{\Gamma}_{\mathrm{mmse}} = [\mathbf{A}^{\mathcal{H}}\mathbf{A} + (\sigma_w^2/\sigma_s^2)\,\mathbf{I}_K]^{-1}\mathbf{A}^{\mathcal{H}} \tag{30}$$

where we have assumed that the symbol vectors are white with covariance matrix $\mathbf{R}_s = \mathrm{E}\{\mathbf{s}(i)\mathbf{s}^{\mathcal{H}}(i)\} = \sigma_s^2\mathbf{I}_K$. The MMSE equalizer reduces to the ZF equalizer by setting $\sigma_w^2 = 0$ in (30).

For nonredundant LP-CP-only with $\boldsymbol{\Theta} = \boldsymbol{\Delta}(\alpha)$, we further simplify (30) to

$$\boldsymbol{\Gamma}_{\mathrm{mmse}} = \boldsymbol{\Delta}(\alpha^*)\mathbf{F}_K^{\mathcal{H}}\left[\overline{\boldsymbol{\mathcal{D}}}_{12}^{\,2} + (\sigma_w^2/\sigma_s^2)\mathbf{I}_K\right]^{-1}\overline{\boldsymbol{\mathcal{D}}}_{12} \tag{31}$$

which amounts to a diagonal matrix inversion followed by an IFFT operation. The MMSE equalizer for CP-only can be simply obtained by setting $\boldsymbol{\Delta}(\alpha) = \boldsymbol{\Delta}(1) = \mathbf{I}_K$ in (31). This equalizer for CP-only is equivalent to the MMSE frequency-domain equalizer derived in [2]. Therefore, the frequency-domain MMSE equalizer is overall MMSE optimal for the CP-only investigated in [2]; this fact is not available in [2].

Capitalizing on the finite-alphabet property of source symbols, the nonlinear block decision feedback equalizer (DFE) proposed in [35] is directly applicable to the systems analyzed here, and is expected to outperform linear receivers with comparable complexity.

More interestingly, it turns out that Viterbi's algorithm is exactly applicable to AP-CP-only and ZP-only, as we detail next.

A. ML Decoding for AP-CP-Only and ZP-Only

For AP-CP-only and ZP-only, we have

$$\mathbf{z} = \overline{\boldsymbol{\mathcal{D}}}_{12}\mathbf{F}_J\mathbf{s} + \boldsymbol{\eta} \tag{32}$$

where we drop the block index i for simplicity. Distinct from other systems, AP-CP-only and ZP-only assure that $\mathbf{s}$ has the last L entries known, and the first K entries drawn from the finite alphabet $\mathcal{A}$.

In the presence of white noise, ML decoding can be expressed as

$$\begin{aligned}\hat{\mathbf{s}}_{\mathrm{ML}} &= \arg\max_{\mathbf{s}} \ln P(\mathbf{z}|\mathbf{s}) \\ &= \arg\max_{\mathbf{s}} \left\{ -\frac{1}{N_0} \left\| \mathbf{z} - \overline{\boldsymbol{\mathcal{D}}}_{12}\mathbf{F}_J\mathbf{s} \right\|^2 \right\}. \end{aligned} \tag{33}$$

Notice that the ML detection in (33) does not depend on the noise variance N_0. We retain N_0, however, in order to facilitate our subsequent extension from ML to turbo decoding, as will become clear in Section III-B. We next simplify (33), starting with

$$\begin{aligned} &- \left\| \mathbf{z} - \overline{\boldsymbol{\mathcal{D}}}_{12}\mathbf{F}_J\mathbf{s} \right\|^2 \\ &\quad = 2\mathrm{Re}\left\{ \mathbf{s}^{\mathcal{H}}\mathbf{F}_J^{\mathcal{H}}\overline{\boldsymbol{\mathcal{D}}}_{12}\mathbf{z} \right\} - \mathbf{s}\mathbf{F}_J^{\mathcal{H}}\overline{\boldsymbol{\mathcal{D}}}_{12}^2\mathbf{F}_J\mathbf{s} - \mathbf{z}^{\mathcal{H}}\mathbf{z} \\ &\quad = 2\mathrm{Re}\left\{ \mathbf{s}^{\mathcal{H}}\mathbf{r} \right\} - \sum_{\mu=1}^{2} \left\| \tilde{\mathbf{H}}_\mu \mathbf{s} \right\|^2 - \mathbf{z}^{\mathcal{H}}\mathbf{z} \end{aligned} \tag{34}$$

where $\mathbf{r} := \mathbf{F}_J^{\mathcal{H}}\overline{\boldsymbol{\mathcal{D}}}_{12}\mathbf{z}$. We let $r_n := [\mathbf{r}]_n$, and $s_n := [\mathbf{s}]_n$. Recognizing that $\tilde{\mathbf{H}}_\mu\mathbf{s}$ expresses nothing but a circular convolution between the channel $\mathbf{h}$ and $\mathbf{s}$, we have

$$[\tilde{\mathbf{H}}_\mu\mathbf{s}]_n = \sum_{l=0}^{L} h_\mu(l) s_{(n-l)\mathrm{mod}J}.$$

Hence, we obtain

$$\begin{aligned}\hat{\mathbf{s}}_{\mathrm{ML}} = \arg\max_{\mathbf{s}} \sum_{n=0}^{J-1} \Bigg\{ \frac{1}{N_0} \Bigg[& 2\,\mathrm{Re}\{s_n^* r_n\} \\ & - \sum_{\mu=1}^{2} \left| \sum_{l=0}^{L} h_\mu(l) s_{(n-l)\mathrm{mod}J} \right|^2 \Bigg] \Bigg\}. \end{aligned} \tag{35}$$

For each $n = 0, 1, \ldots, J$, let us define a sequence of state vectors as

$$\boldsymbol{\zeta}_n = [s_{(n-1)\mathrm{mod}J}, \ldots, s_{(n-L)\mathrm{mod}J}]^T$$

out of which the first and the last states are known[2] $\boldsymbol{\zeta}_0 = \boldsymbol{\zeta}_J = [s_{J-1}, \ldots, s_{J-L}]^T$. The symbol sequence $s_0, \ldots, s_{J-1}$ determines a unique path evolving from the known initial state $\boldsymbol{\zeta}_0$ to the known final state $\boldsymbol{\zeta}_J$. Thus, Viterbi's algorithm is applicable. Specifically, we have

$$\hat{\mathbf{s}}_{\mathrm{ML}} = \arg\max_{\mathbf{s}} \sum_{n=0}^{J-1} f(\boldsymbol{\zeta}_n, \boldsymbol{\zeta}_{n+1}) \tag{36}$$

[2]In CP-only, we have $\boldsymbol{\zeta}_0 = \boldsymbol{\zeta}_J$, but $\boldsymbol{\zeta}_0$ and $\boldsymbol{\zeta}_J$ are unknown. The trellis of CP-only then corresponds to that of a tail-biting convolutional code [4], [29]. Subsequently, ML and maximum *a posteriori* probability (MAP) decoders can be developed for CP-only.

where $f(\boldsymbol{\zeta}_n, \boldsymbol{\zeta}_{n+1})$ is the branch metric, that is readily obtainable from (35). The explicit recursion formula for Viterbi's algorithm is well known; see, e.g., [40, eq. (7)].

We now simplify the branch metric further. We first have

$$\sum_{\mu=1}^{2} \|\tilde{\mathbf{H}}_\mu\mathbf{s}\|^2 = \mathbf{s}^{\mathcal{H}} \sum_{\mu=1}^{2} (\tilde{\mathbf{H}}_\mu^{\mathcal{H}}\tilde{\mathbf{H}}_\mu)\mathbf{s}.$$

The matrix $\overline{\mathbf{H}} := \sum_{\mu=1}^{2} (\tilde{\mathbf{H}}_\mu^{\mathcal{H}}\tilde{\mathbf{H}}_\mu)$ has $(p+1, q+1)$st entry

$$[\overline{\mathbf{H}}]_{p,q} = \sum_{\mu=1}^{2} \sum_{n=0}^{J-1} h_\mu^*((k-p)\mathrm{mod}J) h_\mu((k-q)\mathrm{mod}J). \tag{37}$$

Let us now select $J > 2L$, and define

$$\beta_n = \sum_{\mu=1}^{2} \sum_{l=0}^{L} h_\mu^*(l) h_\mu(n+l), \qquad \forall\, n \in [0, L]. \tag{38}$$

It can be easily verified that the first column of $\overline{\mathbf{H}}$ is

$$[\beta_0, \beta_1, \ldots, \beta_L, 0, \ldots, 0, \beta_L^*, \ldots, \beta_1^*]^T.$$

Let $\check{\mathbf{H}}$ denote the circulant matrix with first column

$$[\beta_0/2, \beta_1, \ldots, \beta_L, 0, \ldots, 0]^T.$$

Because $\overline{\mathbf{H}}$ is circulant and Hermitian, $\overline{\mathbf{H}}$ can be decomposed into $\overline{\mathbf{H}} = \check{\mathbf{H}} + \check{\mathbf{H}}^{\mathcal{H}}$. We thus obtain $\mathbf{s}^{\mathcal{H}}\overline{\mathbf{H}}\mathbf{s} = 2\mathrm{Re}\{\mathbf{s}^{\mathcal{H}}\check{\mathbf{H}}\mathbf{s}\}$. Recognizing

$$[\check{\mathbf{H}}\mathbf{s}]_n = (1/2)\beta_0 s_n + \sum_{l=1}^{L} \beta_l s_{(n-l)\mathrm{mod}J}$$

and combining with (35), we obtain a simplified metric as

$$\begin{aligned} & f(\boldsymbol{\zeta}_n, \boldsymbol{\zeta}_{n+1}) \\ & = \frac{2}{N_0}\,\mathrm{Re}\left\{ s_n^* \left[r_n - \frac{1}{2}\beta_0 s_n - \sum_{l=1}^{L} \beta_l s_{(n-l)\mathrm{mod}J} \right] \right\}. \end{aligned} \tag{39}$$

The branch metric in (39) has a format analogous to the one proposed by Ungerboeck for ML sequence estimation (MLSE) receivers with *single-antenna serial* transmissions [8], [39]. For multiantenna block-coded transmissions, a similar metric has been suggested in [22, eq. (23)]. The important distinction is that [22] suffers from "edge effects" for transmissions with finite block length, resulting in an approximation on the order of $\mathcal{O}(L/J)$, while our derivation here is exact. The key is that our CP-based design assures a circular convolution, while the linear convolution in [22] approximates well a circulant convolution only when $J \gg L$. Note also that we allow for an arbitrary permutation matrix $\mathbf{P}$, which includes the time reversal in [22] as a special case. Furthermore, a known symbol vector $\mathbf{b}$ can be placed in an arbitrary position within the vector $\mathbf{s}$ for AP-CP-only. If the known symbols occupy positions $B-L, \ldots, B-1$, we just need to redefine the states as

$$\boldsymbol{\zeta}_n = [s_{(n+B-1)\mathrm{mod}J}, \ldots, s_{(n+B-L)\mathrm{mod}J}]^T.$$

Notice that for channels with order L, the complexity of Viterbi's algorithm is $\mathcal{O}(|\mathcal{A}|^L)$ per symbol; thus, ML decoding with our exact application of Viterbi's algorithm should be particularly attractive for transmissions with small constellation size, over relatively short channels.

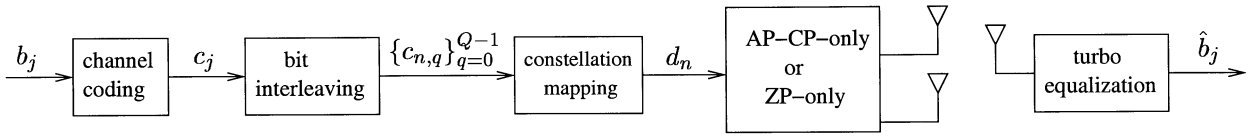

Fig. 5. Coded AP-CP-only or ZP-only with turbo equalization.

B. Turbo Equalization for Coded AP-CP-Only and ZP-Only

So far, we have only considered uncoded systems, and established that full diversity is achieved. To further improve system performance by enhancing also coding gains, conventional channel coding can be applied to our systems. For example, outer convolutional codes can be used in AP-CP-only and ZP-only, as depicted in Fig. 5. Other codes such as trellis-coded modulation (TCM) and turbo codes are applicable as well.

In the presence of frequency-selective channels, iterative (turbo) equalization is known to enhance system performance, at least for single-antenna transmissions [14]. We here derive turbo equalizers for our coded AP-CP-only and ZP-only multiantenna systems.

To enable turbo equalization, one needs to find the *a posteriori* probability on the transmitted symbols s_n based on the received vector $\mathbf{z}$. Suppose each constellation point s_n is determined by $Q = \log_2 |\mathcal{A}|$ bits $\{c_{n,0}, \ldots, c_{n,Q-1}\}$. Let us consider the log-likelihood ratio (LLR)

$$\mathcal{L}_{n,q} = \ln \frac{P(c_{n,q} = +1|\mathbf{z})}{P(c_{n,q} = -1|\mathbf{z})}, \quad \forall\, n \in [0, J-1], \quad q \in [0, Q-1]. \quad (40)$$

As detailed in [40], the LLR in (40) can be obtained by running two generalized Viterbi recursions: one in the forward direction evolving from $\boldsymbol{\zeta}_0$ to $\boldsymbol{\zeta}_J$, and the other in the backward direction going from $\boldsymbol{\zeta}_J$ to $\boldsymbol{\zeta}_0$. We refer the readers to [40, eqns. (7′), (8′), (10′)] for explicit expressions. The only required change is to modify our branch metric as follows:

$$g(\boldsymbol{\zeta}_n, \boldsymbol{\zeta}_{n+1}) = f(\boldsymbol{\zeta}_n, \boldsymbol{\zeta}_{n+1}) + \ln P(\boldsymbol{\zeta}_{n+1}|\boldsymbol{\zeta}_n). \quad (41)$$

This modification is needed to take into account the *a priori* probability $P(\boldsymbol{\zeta}_{n+1}|\boldsymbol{\zeta}_n)$, determined by the extrinsic information from the convolutional channel decoders during the turbo iteration. When the transition from $\boldsymbol{\zeta}_n$ to $\boldsymbol{\zeta}_{n+1}$ is caused by the input symbol s_n, we have $\ln P(\boldsymbol{\zeta}_{n+1}|\boldsymbol{\zeta}_n) = \ln P(s_n)$. We assume that the bit interleaver in Fig. 5 renders the symbols s_n independent and equal likely, such that

$$\ln P(s_n) = \sum_{q=0}^{Q-1} \ln P(c_{n,q})$$

which, in turn, can be determined by the LLRs for bits $\{c_{n,q}\}_{q=0}^{Q-1}$.

Finally, we remark that one could also adopt the turbo decoding algorithm of [43] that is based on MMSE equalizers. This iterative receiver is applicable not only to AP-CP-only and ZP-only, but also to CP-only and LP-CP-only systems.

C. Receiver Complexity

Omitting the complexity of permutation and diagonal matrix multiplication, the linear processing to reach (13) only requires one size-J FFT per block, which amounts to $\mathcal{O}(\log_2 J)$ per information symbol.

Channel equalization is then performed based on (13) for each block. We notice that the complexity is the same as the equalization complexity for single-antenna block transmissions over FIR channels [45]. We refer the readers to [45] for detailed complexity comparisons of the different equalization options. For coded AP-CP-only and ZP-only, the complexity of turbo equalization is again the same as that of single-antenna transmissions [14].

In summary, the overall receiver complexity for the two transmit antennas case is comparable to that of single-antenna transmissions, with only one additional FFT per data block. This nice property originates from the orthogonal ST block code design, that enables linear ML processing to collect antenna diversity. Depending desirable/affordable diversity–complexity tradeoffs, the designer is then provided with the flexibility to collect extra multipath-diversity gains.

IV. Extension to Multiple Antennas

In Section II, we focused on $N_t = 2$ transmit and $N_r = 1$ receive antennas. In this section, we will extend our system design to the general case with $N_t > 2$ and/or $N_r > 1$ antennas. For each $\mu = 1, \ldots, N_t$ and $\nu = 1, \ldots, N_r$, we denote the channel between the μth transmit and the νth receive antennas as $\mathbf{h}_{\mu\nu} := [h_{\mu\nu}(0), \ldots, h_{\mu\nu}(L)]^T$, and as before we model it as a zero-mean, complex Gaussian vector with covariance matrix $\mathbf{R}_{h,\mu\nu}$. Correspondingly, we define the effective channel order $\tilde{L}_{\mu\nu} := \text{rank}\{\mathbf{R}_{h,\mu\nu}\} - 1$, which for a sufficiently rich scattering environment becomes $\tilde{L}_{\mu\nu} = L$.

Transmit diversity with $N_t > 2$ has been addressed in [25], [26] for OFDM-based multicarrier transmissions over FIR channels by applying the orthogonal ST block codes of [36] on each OFDM subcarrier. Here, we extend the orthogonal designs to single-carrier block transmissions over frequency-selective channels.

We will review briefly generalized orthogonal designs to introduce notation, starting with the basic definitions given in the context of frequency-flat channels [36].

Definition 1: Define $\mathbf{x} := [x_1, \ldots, x_{N_s}]^T$, and let $\mathcal{G}_r(\mathbf{x})$ be an $N_d \times N_t$ matrix with entries $0, \pm x_1, \ldots, \pm x_{N_s}$. If $\mathcal{G}_r^T(\mathbf{x})\mathcal{G}_r(\mathbf{x}) = \alpha(x_1^2 + \cdots + x_{N_s}^2)\mathbf{I}_{N_t}$ with α positive, then $\mathcal{G}_r(\mathbf{x})$ is termed a generalized real orthogonal design (GROD) in variables $x_1, \ldots, x_{N_s}$ of size $N_d \times N_t$ and rate $R = N_s/N_d$.

Definition 2: Define $\mathbf{x} := [x_1, \ldots, x_{N_s}]^T$, and let $\mathcal{G}_c(\mathbf{x})$ be an $N_d \times N_t$ matrix with entries $0, \pm x_1, \pm x_1^*, \ldots, \pm x_{N_s}, \pm x_{N_s}^*$.

If $\mathcal{G}_c^{\mathcal{H}}(\mathbf{x})\mathcal{G}_c(\mathbf{x}) = \alpha(|x_1|^2 + \cdots + |x_{N_s}|^2)\mathbf{I}_{N_t}$ with α positive, then $\mathcal{G}_c(\mathbf{x})$ is termed a generalized complex orthogonal design (GCOD) in variables $x_1, \ldots, x_{N_s}$ of size $N_d \times N_t$ and rate $R = N_s/N_d$.

Explicit construction of $\mathcal{G}_r(\mathbf{x})$ with $R = 1$ was discussed in [36], where it was also proved that the highest rate for $\mathcal{G}_c(\mathbf{x})$ is $1/2$ when $N_t > 4$. When $N_t = 3, 4$, there exist some sporadic codes with rate $R = 3/4$. Although the orthogonal designs with $R = 3/4$ for $N_t = 3, 4$ have been incorporated in [26] for multicarrier transmissions, we will not consider them in our single-carrier block transmissions here;[3] we will only consider the $R = 1/2$ GCOD designs primarily because GCOD $\mathcal{G}_c(\mathbf{x})$ of $R = 1/2$ can be constructed using the following steps ($N_s = 4$ for $N_t = 3, 4$, while $N_s = 8$ for $N_t = 5, 6, 7, 8$ [36]):

s1) construct GROD $\mathcal{G}_r(\mathbf{x})$ of size $N_s \times N_t$ with $R = 1$;

s2) replace the symbols $x_1, \ldots, x_{N_s}$ in $\mathcal{G}_r(\mathbf{x})$ by their conjugates $x_1^*, \ldots, x_{N_s}^*$ to arrive at $\mathcal{G}_r(\mathbf{x}^*)$;

s3) form $\mathcal{G}_c(\mathbf{x}) = [\mathcal{G}_r^T(\mathbf{x}), \mathcal{G}_r^T(\mathbf{x}^*)]^T$.

As will be clear soon, we are explicitly taking into account the fact that all symbols from the upper part of $\mathcal{G}_c(\mathbf{x})$ are unconjugated, while all symbols from the lower part are conjugated. The rate loss can be as high as 50%, when $N_t > 2$.

With $N_t > 2$, the ST mapper takes N_s consecutive blocks to output the following $N_t J \times N_d$ ST coded matrix ($N_d = 2N_s$):

$$\overline{\mathbf{S}}(i) = \mathcal{E}\{\mathbf{s}(iN_s), \ldots, \mathbf{s}(iN_s + N_s - 1)\} = \begin{bmatrix} \overline{\mathbf{s}}_1(iN_d) & \cdots & \overline{\mathbf{s}}_1(iN_d + N_d - 1) \\ \vdots & \ddots & \vdots \\ \overline{\mathbf{s}}_{N_t}(iN_d) & \cdots & \overline{\mathbf{s}}_{N_t}(iN_d + N_d - 1) \end{bmatrix} \begin{matrix} \rightarrow \text{time} \\ \downarrow \text{space}. \end{matrix} \tag{42}$$

The design steps are summarized as follows:

d1) construct $\mathcal{G}_c$ of size $2N_s \times N_t$ in the variables $x_1, \ldots, x_{N_s}$, as in s1)–s3);

d2) replace $x_1, \ldots, x_{N_s}$ in $\mathcal{G}_c^T$ by

$$\mathbf{s}(iN_s), \ldots, \mathbf{s}(iN_s + N_s - 1);$$

d3) replace $x_1^*, \ldots, x_{N_s}^*$ in $\mathcal{G}_c^T$ by

$$\mathbf{P}\mathbf{s}^*(iN_s), \ldots, \mathbf{P}\mathbf{s}^*(iN_s + N_s - 1)$$

where $\mathbf{P}$ is taken properly for different schemes as explained in Section II.

Note that with $N_t = 2$ transmit antennas, the above procedure has actually been applied to the 2×2 GCOD design in [3].

At each block transmission slot i, $\overline{\mathbf{s}}_\mu(i)$ is forwarded to the μth antenna, and transmitted through the FIR channel after CP insertion. Each receive antenna processes blocks independently as follows. The receiver removes the CP and collects $N_d = 2N_s$ blocks $\mathbf{x}(iN_d), \ldots, \mathbf{x}(iN_d + N_d - 1)$. Then FFT is performed on the first N_s blocks $\mathbf{x}(iN_d), \ldots, \mathbf{x}(iN_d + N_s - 1)$, while permutation and conjugation is applied to the last N_s blocks $\mathbf{P}\mathbf{x}^*(iN_d + N_s), \ldots, \mathbf{P}\mathbf{x}^*(iN_d + N_d - 1)$, followed by FFT processing. Coherently combining the FFT outputs as we did for the two-antennas case to derive (13), we obtain on each antenna the equivalent output after the optimal linear processing

$$\mathbf{z}_\nu(i) = \overline{\boldsymbol{\mathcal{D}}}_\nu \mathbf{F}_J \mathbf{s}(i) + \boldsymbol{\eta}_\nu(i) \tag{43}$$

where

$$\overline{\boldsymbol{\mathcal{D}}}_\nu := \left[\sum_{\mu=1}^{N_t} \boldsymbol{\mathcal{D}}_{\mu,\nu}^* \boldsymbol{\mathcal{D}}_{\mu,\nu}\right]^{1/2}$$

and

$$\boldsymbol{\mathcal{D}}_{\mu,\nu} := \mathrm{diag}(\tilde{\mathbf{h}}_{\mu\nu}) = \mathrm{diag}(\mathbf{V}\mathbf{h}_{\mu\nu}).$$

We next stack $\mathbf{z}_\nu(i)$ to form $\overline{\mathbf{z}}(i) = [\mathbf{z}_1^T(i), \ldots, \mathbf{z}_{N_r}^T(i)]^T$ (likewise for $\overline{\boldsymbol{\eta}}(i)$), and define $\mathbf{B} := [\overline{\boldsymbol{\mathcal{D}}}_1, \ldots, \overline{\boldsymbol{\mathcal{D}}}_{N_r}]^T$, to obtain: $\overline{\mathbf{z}}(i) = \mathbf{B}\mathbf{F}_J\mathbf{s}(i) + \overline{\boldsymbol{\eta}}(i)$. Defining

$$\overline{\mathbf{B}} := \left[\sum_{\mu=1}^{N_t}\sum_{\nu=1}^{N_r} \boldsymbol{\mathcal{D}}_{\mu,\nu}^* \boldsymbol{\mathcal{D}}_{\mu,\nu}\right]^{1/2}$$

we have $\mathbf{B}^{\mathcal{H}}\mathbf{B} = \overline{\mathbf{B}}^2$. Therefore, we can construct a matrix $\mathbf{U}_b := \mathbf{B}\overline{\mathbf{B}}^{-1}$, which has orthonormal columns $\mathbf{U}_b^{\mathcal{H}}\mathbf{U}_b = \mathbf{I}_J$, and satisfies $\mathbf{U}_b^{\mathcal{H}}\mathbf{B} = \overline{\mathbf{B}}$. As $\mathbf{U}_b$ and $\mathbf{B}$ share range spaces, multiplying $\mathbf{U}_b^{\mathcal{H}}$ by $\overline{\mathbf{z}}(i)$ incurs no loss of optimality, and leads to the following equivalent block:

$$\mathbf{z}(i) := \mathbf{U}_b^{\mathcal{H}}\overline{\mathbf{z}}(i) = \overline{\mathbf{B}}\mathbf{F}_J\mathbf{s}(i) + \boldsymbol{\eta}(i) \tag{44}$$

where the noise $\boldsymbol{\eta}(i)$ is still white. Now the distance between $\mathbf{z}$ and $\mathbf{z}'$, corresponding to two different symbol blocks $\mathbf{d}$ and $\mathbf{d}'$, becomes

$$d^2(\mathbf{z}, \mathbf{z}') = \sum_{\mu=1}^{N_t}\sum_{\nu=1}^{N_r} \|\mathbf{D}_e \mathbf{V}\mathbf{h}_{\mu\nu}\|^2. \tag{45}$$

Comparing (45) with (15), the contribution now comes from $N_t N_r$ multipath channels. Following the same steps as in Section II, it is straightforward to establish the following result.

Proposition 1: The maximum achievable diversity order is $\sum_{\mu=1}^{N_t}\sum_{\nu=1}^{N_r}(\check{L}_{\mu\nu} + 1)$ with N_t transmit and N_r receive antennas, which equals $N_t N_r(L + 1)$ when the channel correlation has full rank.

1) CP-only achieves multiantenna diversity of order $N_t N_r$;

2) LP-CP-only achieves the maximum diversity gain through either nonredundant but constellation-dependent, or, redundant but constellation-independent precoding;

3) AP-CP-only and ZP-only achieve the maximum diversity gain irrespective of the underlying signal constellation. □

The linear ML processing to reach (44) requires a total of $N_d N_r = 2N_s N_r$ FFTs corresponding to each ST-coded block of (42), which amounts to $2N_r$ FFTs per information block. Channel equalization based on (44) incurs identical complexity as in single-antenna transmissions. For AP-CP-only and ZP-only, the ML estimate $\hat{\mathbf{s}}_{\mathrm{ML}} = \arg\max_{\mathbf{s}}(-\|\mathbf{z} - \overline{\mathbf{B}}\mathbf{F}\mathbf{s}\|^2/N_0)$ can be obtained via exact application of Viterbi's algorithm. Relative to the two-antenna case detailed in Section III-A, we can basically use the same expression for the branch metric

[3] We find that the rate $3/4$ code for $N_t = 3, 4$ can be incorporated in CP-only, LP-CP-only, and AP-CP-only transmissions only when $\mathbf{P} = \mathbf{P}_J^{(1)}$ is used, because $\mathbf{F}_J\mathbf{P}_J^{(1)}\mathbf{x}^*(i) = (\mathbf{F}_J\mathbf{x}(i))^*$ indicates that no actual permutation is needed at the receiver. In this case, the first step at the receiver is to take the FFT of each received block to yield $\mathbf{F}_J\mathbf{x}(i)$, and follow the processing used in [26]. We only point out this case in a footnote for brevity.

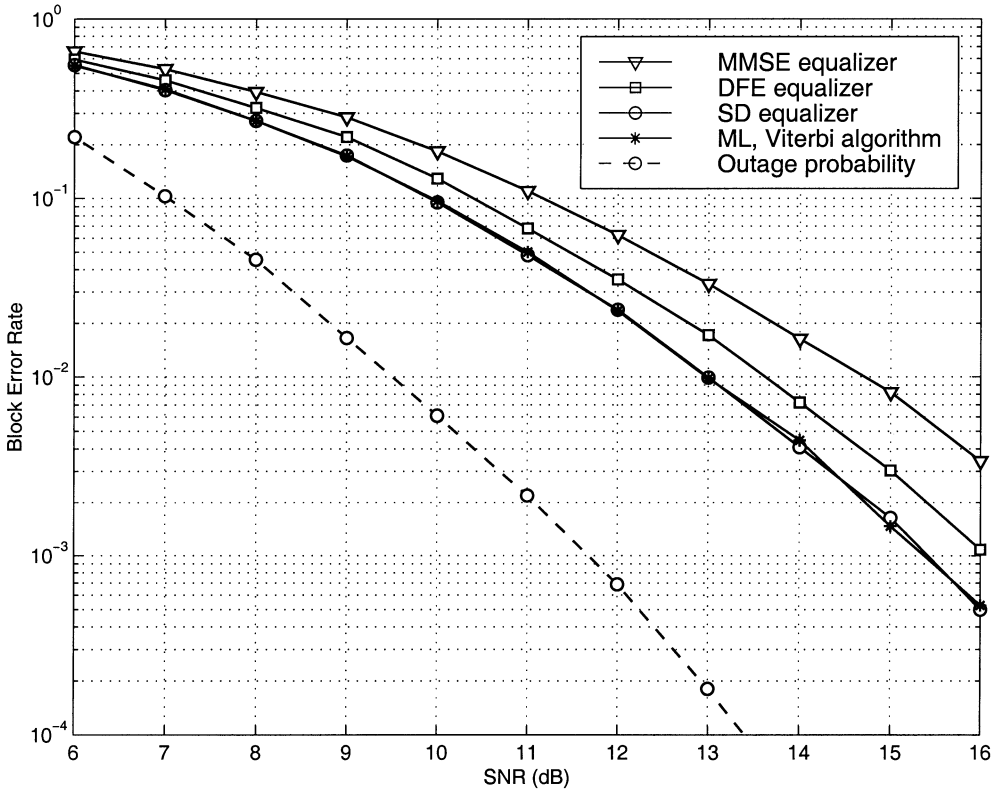

Fig. 6. Comparisons of various equalizers for uncoded ZP-only.

of (39), with two modifications, namely, $r_n = [\mathbf{r}]_n$ with $\mathbf{r} = \mathbf{F}_J^{\mathcal{H}}\overline{\mathbf{B}}\mathbf{z}$, and

$$\beta_n = \sum_{\mu=1}^{N_t}\sum_{\nu=1}^{N_r}\sum_{l=0}^{L} h_{\mu\nu}^*(l)h_{\mu\nu}(n+l), \qquad \forall\, n \in [0, L]. \quad (46)$$

We summarize the general complexity results of this section and those of Section II in the following.

Proposition 2: The proposed ST block coded CP-only, LP-CP-only, AP-CP-only, and ZP-only systems with $N_t > 2$ ($N_t = 2$) transmit and N_r receive antennas require an additional complexity of $\mathcal{O}(2N_r \log_2 J)$ (respectively, $\mathcal{O}(N_r \log_2 J)$) per information symbol, relative to their counterparts with single transmit and single receive antenna, where J is the FFT size. □

V. Simulated Performance

In this section, we present simulation results for systems with two transmit and one receive antennas. For ease in FFT processing, we always choose the block size J to be a power of 2. In all figures, we define SNR as the average received symbol energy-to-noise ratio at the receive antenna. For reference, we also depict the (outage) probability that the channel capacity is less than the desired rate, so that reliable communication at this rate is impossible. Specifically, we calculate (28) numerically, and similar to [23], we evaluate the outage probability at the targeted rate R as $P(C_{J\to\infty} < R)$ with Monte Carlo simulations.

Test Case 1 (Comparisons for Different Equalizers): We first set $L = 2$, and assume that the channels between each transmit and each receive antenna are i.i.d., Gaussian, with covariance matrix $\mathbf{I}_{L+1}/(L+1)$. We investigate the performance of ZP-only with block sizes $K = 14$ and $P = J = 16$. We adopt QPSK constellations. Fig. 6 depicts the block error rate performance corresponding to MMSE, DFE, SD, and ML equalizers. We observe that the SD equalizer indeed achieves near-ML performance, and outperforms the suboptimal block DFE as well as the block MMSE alternatives. Without channel coding, the performance of ZP-only is far away from the outage probability at rate $2K/(K+L) = 1.75$ bits per channel use.

Test Case 2 (Convolutionally Coded ZP-Only): We here adopt the channel setup of [23], [24], which consists of two i.i.d. taps per FIR channel, i.e., $L = 1$. We set the block sizes as $K = 127$, $P = J = 128$ for our ZP-only system, and use 8-PSK constellation. For convenience, we view each block of length $P = 128$ as one data frame, with the ST codes applied to two adjacent frames. Within each frame, the information bits are convolutionally coded (CC) with a 16-state rate $2/3$ encoder taken from [5, Table 11.6]. Omitting the trailing bits to terminate the CC trellis, and ignoring the rate loss induced by the CP since $L \ll K$, we obtain a transmission rate of 2 bits per channel use. The outage probability provided in Fig. 7 is thus identical to that in [23, Fig. 6].

As in [23], [24], five turbo decoding iterations are performed. With the 16-state convolutional code, the frame error rate for ZP-only is within 2.3 dB away from the outage probability. This performance is comparable to (or, slightly better than) that of ST trellis coding (STTC) [23], [24] for frequency-selective channels, that is replicated[4] in Fig. 7 as well. This result is also con-

[4]Notice that this curve is obtained with frame length of 130 symbols in [23], [24]. We just copy this curve from [23, Fig. 6].

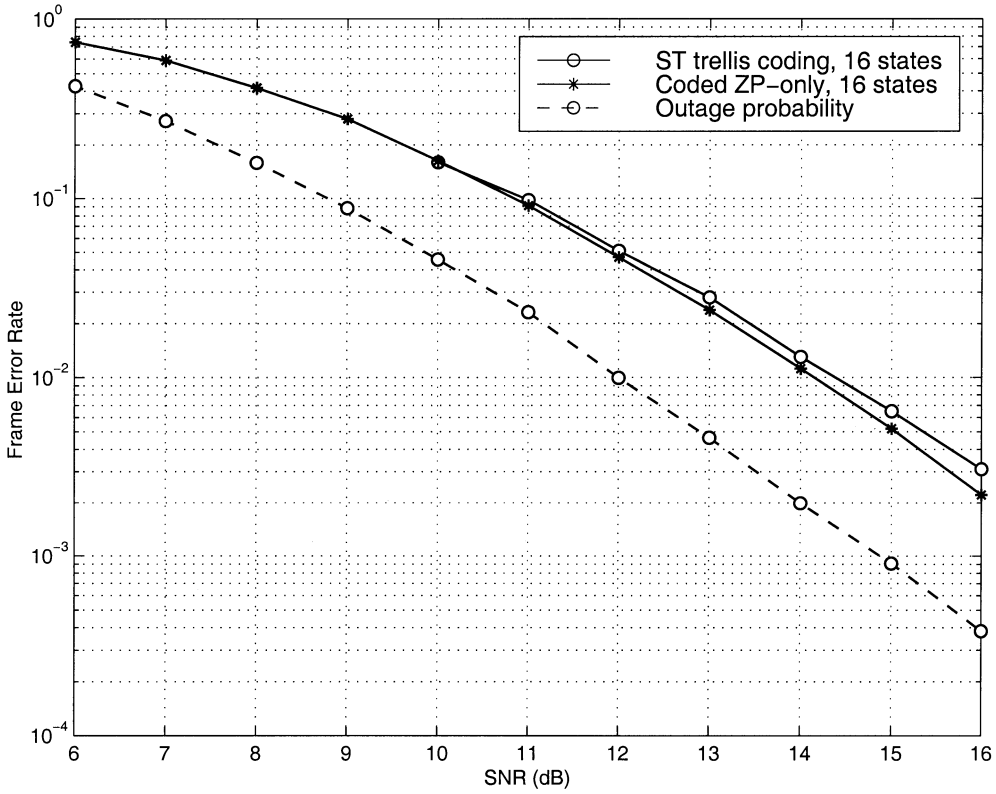

Fig. 7. Convolutionally coded ZP-only versus ST trellis coding of [23], [24]

sistent with [32], where the performance of an STBC system is shown to be better than those of known STTC systems, in flat fading channels.

Unlike 8-PSK used here, QPSK constellation is used in the STTC of [23], [24]. However, thanks to the orthogonal design, the trellis induced by the FIR channel has $|\mathcal{A}|^L = 8$ states for our ZP-only system, which is smaller than the $|\mathcal{A}'|^{2L} = 16$ states required by [23], [24]; hence, ZP-only will (at least in this case) incur lower turbo decoding complexity than [23], [24].

Existing channel codes with arbitrary number of states can be applied directly to our ZP-only system. Irrespective of the channel code and the random interleaver used, full diversity is always guaranteed. In contrast, the STTC and the random interleaver in [23], [24] should be jointly designed to ensure full diversity. And this design is certainly more complex than our ZP-only scheme.

Test Case 3 (Convolutionally Coded AP-CP-Only Over EDGE Channels): We test the Typical Urban (TU) channel with a linearized Gaussian minimum shift keying (GMSK) transmit pulse shape, and a symbol duration $T = 3.69$ μs as in the proposed third-generation time division multiple access (TDMA) cellular standard EDGE [enhance date rates for global system for mobile communications (GSM) evolution][15]. The channel has order $L = 3$ and correlated taps [2]. We use QPSK constellations, and set the block size $J = 128$. We adopt AP-CP-only that guarantees perfectly constant modulus transmissions. Within each frame of 128 symbols, the last three are known. Information bits are coded using a 16-state rate $1/2$ convolutional code taken from [5, Table 11.2]. Taking into account the known symbols, the cyclic prefix, and zero bits to terminate the CC trellis, the overall transmission rate of the proposed AP-CP-only is $(128-3-4)/(128+3) = 0.924$ bits per channel use, or 250.4 kb/s.

As shown in Fig. 8, the system with two transmit antennas significantly outperforms its counterpart with one transmit antenna. At frame error rate of 10^{-2}, about 5-dB SNR gain has been achieved. Fig. 9 depicts the performance improvement with turbo iterations, which confirms the importance of iterative over noniterative receivers. A large portion of the performance gain is achieved within three iterations.

VI. Conclusion

In this paper, we developed single-carrier ST block-coded transmissions through frequency-selective multipath channels. We proposed novel transmission formats, that correspond to orthogonal ST block codes for frequency-selective channels. We showed that a maximum diversity of order $N_tN_r(L+1)$ in rich scattering environment is achievable. Linear processing collects full antenna diversity, while the overall receiver complexity remains comparable to that of single-antenna transmissions over frequency-selective channels. Optimal ML Viterbi decoding and various suboptimal equalization alternatives were examined. Iterative (turbo) equalizers were also developed for our ST transmissions combined with channel coding. With single receive and two transmit antennas, the proposed transmission format does not incur capacity loss. For this scenario, our novel designs benchmark performance and capacity over ST frequency-selective channels, pretty much as Alamouti's code did over frequency-flat channels. Simulation results demonstrated that joint

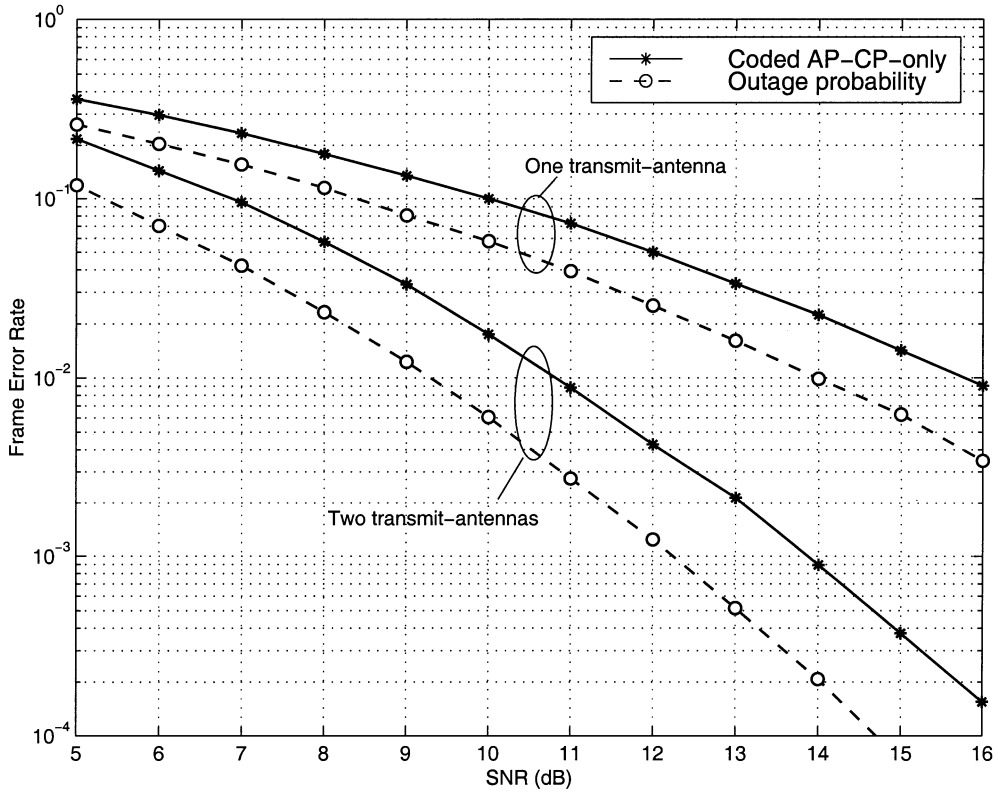

Fig. 8. Convolutionally coded AP-CP-only over EDGE channels.

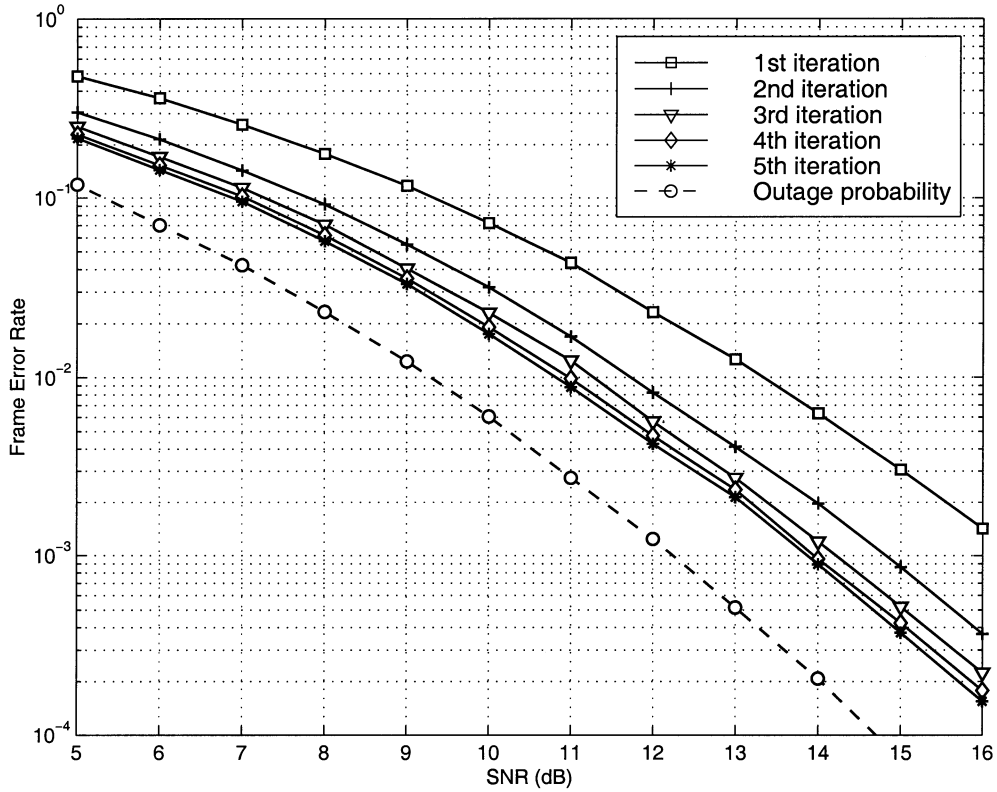

Fig. 9. Turbo iterations for coded AP-CP-only with two transmit-antennas.

exploitation of space–multipath diversity leads to significantly improved performance in the presence of frequency-selective fading channels.

APPENDIX
PROOF OF $\mathbf{P}\tilde{\mathbf{H}}_\mu\mathbf{P} = \tilde{\mathbf{H}}_\mu^T$

As defined in the text, $\tilde{\mathbf{H}}_\mu$ is a $J \times J$ circulant matrix with $(p+1,\ q+1)$th entry

$$[\tilde{\mathbf{H}}_\mu]_{p,\,q} = h_\mu(p - q \mod J)$$

where $p,\ q \in [0,\ J-1]$; and the permutation matrix $\mathbf{P}$ is drawn from the set $\{\mathbf{P}_J^{(n)}\}_{n=0}^{J-1}$. With the specific choice of $\mathbf{P} = \mathbf{P}_J^{(n)}$, the matrix $\mathbf{P}\tilde{\mathbf{H}}_\mu$ is obtained by a reverse cyclic row shifting on $\tilde{\mathbf{H}}_\mu$, moving the $(N-p+n)$th row to the $(p+1)$st row, to obtain

$$[\mathbf{P}\tilde{\mathbf{H}}_\mu]_{p,\,q} = h_\mu((N - p + n - 1) - q \mod J).$$

Likewise, post-multiplying $\mathbf{P}\tilde{\mathbf{H}}_\mu$ by $\mathbf{P}$ corresponds to reversely cyclic column-shifting by moving the $(N-q+n)$th column of $\mathbf{P}\tilde{\mathbf{H}}_\mu$ to the $(q+1)$st column. Therefore, we obtain

$$\begin{aligned}\left[\mathbf{P}\tilde{\mathbf{H}}_\mu\mathbf{P}\right]_{p,\,q} &= h_\mu((N-p+n-1)-(N-q+n-1) \mod J)\\ &= h_\mu(q-p \mod J)\\ &= \left[\tilde{\mathbf{H}}_\mu\right]_{q,\,p} \end{aligned} \tag{47}$$

which implies that $\mathbf{P}\tilde{\mathbf{H}}_\mu\mathbf{P} = \tilde{\mathbf{H}}_\mu^T$, and proves p2).

ACKNOWLEDGMENT

The authors wish to thank Dr. Naofal Al-Dhahir from AT&T Shannon Laboratory for providing the EDGE channel generator used in [2], Pengfei Xia from the University of Minnesota for his help with the simulations, Dr. Anantharam Swami from ARL for suggesting footnote 2, and anonymous reviewers for their helpful suggestions.

REFERENCES

[1] D. Agrawal, V. Tarokh, A. Naguib, and N. Seshadri, "Space–time coded OFDM for high data-rate wireless communication over wideband channels," in *Proc. Vehicular Technology Conf.*, vol. 3, Ottawa, ON, Canada, 1998, pp. 2232–2236.

[2] N. Al-Dhahir, "Single-carrier frequency-domain equalization for space–time block-coded transmissions over frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 5, pp. 304–306, July 2001.

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

[4] J. B. Anderson and S. M. Hladik, "Tailbiting MAP decoders," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 297–302, Feb. 1998.

[5] S. Benedetto and E. Biglieri, *Principles of Digital Transmission With Wireless Applications*. Norwell, MA: Kluwer, 1999.

[6] H. Bölcskei and A. J. Paulraj, "Space-frequency coded broadband OFDM systems," in *Proc. Wireless Communications and Networking Conf.*, vol. 1, Chicago, IL, Sept. 2000, pp. 1–6.

[7] ——, "Space-frequency codes for broadband fading channels," in *Proc. IEEE Int. Symp. Information Theory*, Washington, D.C., June 2001, p. 219.

[8] G. E. Bottomley and S. Chennakeshu, "Unification of MLSE receivers and extension to time-varying channels," *IEEE Trans. Commun.*, vol. 46, pp. 464–472, Apr. 1998.

[9] J. Boutros and E. Viterbo, "Signal space diversity: A power- and bandwidth-efficient diversity technique for the Rayleigh fading channel," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1453–1467, July 1998.

[10] W. Choi and J. M. Cioffi, "Multiple input/multiple output (MIMO) equalization for space–time block coding," in *Proc. IEEE Pacific Rim Conf. Communications, Computers and Signal Processing*, 1999, pp. 341–344.

[11] ——, "Space–time block codes over frequency selective Rayleigh fading channels," in *Proc. Vehicular Technology Conf.*, vol. 5, Amsterdam, The Netherlands, Fall 1999, pp. 2541–2545.

[12] M. V. Clark, "Adaptive frequency-domain equalization and diversity combining for broadband wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1385–1395, Oct. 1998.

[13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[14] C. Douillard, C. B. Jezequel, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: Turbo-equalization," *Europ. Trans. Telecommun.*, vol. 6, pp. 507–511, Sept. 1995.

[15] A. Furuskar, S. Mazur, F. Muller, and H. Olofsson, "EDGE: Enhanced data rates for GSM and TDMA/136 evolution," *IEEE Personal Commun.*, vol. 6, no. 3, pp. 56–66, June 1999.

[16] G. Ganesan and P. Stoica, "Achieving optimum coded diversity with scalar codes," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2078–2080, July 2001.

[17] X. Giraud, E. Boutillon, and J. C. Belfiore, "Algebraic tools to build modulation schemes for fading channels," *IEEE Trans. Inform. Theory*, vol. 43, pp. 938–952, May 1997.

[18] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996.

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

[20] Y. Li, J. C. Chung, and N. R. Sollenberger, "Transmitter diversity for OFDM systems and its impact on high-rate data wireless networks," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1233–1243, July 1999.

[21] Y. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 461–471, Mar. 1999.

[22] E. Lindskog and A. Paulraj, "A transmit diversity scheme for channels with intersymbol interference," in *Proc. Int. Conf. Communications*, vol. 1, New Orleans, LA, June 2000, pp. 307–311.

[23] Y. Liu, M. P. Fitz, and O. Y. Takeshita, "Space–time codes for frequency selective channel: Outage probability, performance criteria, and code design," in *Proc. 38th Annu. Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Oct. 2000.

[24] ——, "Space–time codes performance criteria and design for frequency selective fading channels," in *Proc. Int. Conf. Communications*, vol. 9, Helsinki, Finland, June 2001, pp. 2800–2804.

[25] Z. Liu and G. B. Giannakis, "Space–time coding with transmit antennas for multiple access regardless of frequency-selective multipath," in *Proc. Sensor Array and Multichannel Signal Processing Workshop*, Boston, MA, Mar. 2000, pp. 178–182.

[26] ——, "Space–time block coded multiple access through frequency-selective fading channels," *IEEE Trans. Commun.*, vol. 49, pp. 1033–1044, June 2001.

[27] Z. Liu, G. B. Giannakis, B. Muquet, and S. Zhou, "Space–time coding for broadband wireless communications," *Wireless Commun. Mobile Comput.*, vol. 1, no. 1, pp. 33–53, Jan.–Mar. 2001.

[28] B. Lu and X. Wang, "Space–time code design in OFDM systems," in *Proc. Global Telecommunications Conf.*, vol. 2, San Francisco, CA, 2000, pp. 1000–1004.

[29] H. H. Ma and J. K. Wolf, "On tail biting convolutional codes," *IEEE Trans. Commun.*, vol. 34, pp. 104–111, Feb. 1986.

[30] A. F. Naguib, N. Seshadri, and R. Calderbank, "Increasing data rates over wireless channels," *IEEE Signal Processing Mag.*, vol. 17, pp. 76–92, May 2000.

[31] G. G. Raleigh and J. M. Cioffi, "Spatio–temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, pp. 357–366, Mar. 1998.

[32] S. Sandhu, R. Heath, and A. Paulraj, "Space–time block codes versus space–time trellis codes," in *Proc. Int. Conf. Communications*, vol. 4, Helsinki, Finland, June 2001, pp. 1132–1136.

[33] S. Sandhu and A. Paulraj, "Space–time block codes: A capacity perspective," *IEEE Commun. Lett.*, vol. 4, pp. 384–386, Dec. 2000.

[34] H. Sari, G. Karam, and I. Jeanclaude, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Commun. Mag.*, vol. 33, pp. 100–109, Feb. 1995.

[35] A. Stamoulis, G. B. Giannakis, and A. Scaglione, "Block FIR decision-feedback equalizers for filterbank precoded transmissions with blind channel estimation capabilities," *IEEE Trans. Commun.*, vol. 49, pp. 69–83, Jan. 2001.

[36] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space–time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July 1999.

[37] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank, "Space–time codes for high data rate wireless communication: Performance criteria in the presence of channel estimation errors, mobility, and multiple paths," *IEEE Trans. Commun.*, vol. 47, pp. 199–207, Feb. 1999.

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

[39] G. Ungerboeck, "Adaptive maximum likelihood receiver for carrier modulated data transmission systems," *IEEE Trans. Commun.*, vol. COM-22, pp. 624–635, May 1974.

[40] A. J. Viterbi, "An intuitive justification and a simplified implementation of the MAP decoder for convolutional codes," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 260–264, Feb. 1998.

[41] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1639–1642, July 1999.

[42] F. W. Vook and T. A. Thomas, "Transmit diversity schemes for broadband mobile communication systems," in *Proc. Vehicular Technology Conf.*, vol. 6, Boston, MA, Fall 2000, pp. 2523–2529.

[43] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 46, pp. 1046–1061, July 1999.

[44] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications: Where Fourier meets Shannon," *IEEE Signal Processing Mag.*, pp. 29–48, May 2000.

[45] ——, "Linearly precoded or coded OFDM against wireless channel fades?," in *Proc. 3rd IEEE Workshop Signal Processing Advances in Wireless Communications*, Taoyuan, Taiwan, R.O.C., Mar. 2001, pp. 267–270.

[46] Y. Xin, Z. Wang, and G. B. Giannakis, "Space–time constellation-rotating codes maximizing diversity and coding gains," in *Proc. Global Telecommunications Conf.*, San Antonio, TX, Nov. 25–29, 2001, pp. 455–459.

[47] S. Zhou and G. B. Giannakis, "Space–time coded transmissions with maximum diversity gains over frequency-selective multipath fading channels," in *Proc. Global Telecommunications Conf.*, San Antonio, TX, Nov. 25–29, 2001, pp. 440–444.