

# A Quasi-Random Approach to Space–Time Codes

Keying Wu and Li Ping, *Senior Member, IEEE*

**Abstract**—This paper presents a quasi-random approach to space–time (ST) codes. The basic principle is to transmit randomly interleaved versions of forward error correction (FEC)-coded sequences simultaneously from all antennas in a multilayer structure. This is conceptually simple, yet still very effective. It is also flexible regarding the transmission rate, antenna numbers, and channel conditions (e.g., with intersymbol interference). It provides a unified solution to various applications where the traditional ST codes may encounter difficulties. We outline turbo-type iterative joint detection and equalization algorithms with complexity (per FEC-coded bit) growing linearly with the transmit antenna number and independently of the layer number. We develop a signal-to-noise-ratio (SNR) evolution technique and a bounding technique to assess the performance of the proposed code in fixed and quasi-static fading channels, respectively. These performance assessment techniques are very simple and reasonably accurate. Using these techniques as a searching tool, efficient power allocation strategies are examined, which can greatly enhance the system performance. Simulation results show that the proposed code can achieve near-capacity performance with both low and high rates at low decoding complexity.

**Index Terms**—Equalization, forward error correction (FEC), iterative decoding, power allocation, quasi-random code, space–time (ST) code.

## I. INTRODUCTION

THE spatial diversity provided by multiple antenna systems can be used to significantly increase the reliability and spectrum efficiency of wireless communication systems [1], [2]. Two approaches have been studied [3]–[9]. The first is space–time (ST) coding [3]–[7] that achieves high diversity gain by spreading transmitted energy over different transmit antennas. The second is spatial multiplexing [8], [9] that achieves high spectrum efficiency by transmitting independent information sequences from different antennas. Multiple receive antennas are usually required in the second approach so as to ensure reliability, such as in the vertical Bell Laboratories layered space–time (V-BLAST) architecture [8].

Transmission schemes for both high diversity gain and spectrum efficiency have been explored. The diagonal BLAST (D-BLAST) architecture [9] employs a diagonally layered structure in the ST dimensions to improve diversity gain. The turbo-BLAST scheme [10] introduces ST interleaving together

with iterative detection for further performance enhancement. Good diversity gain and spectrum efficiency have also been reported for the threaded ST code [11] that combines the sophisticated algebraic ST coding design technique with the layered ST transmission architecture, and for the linear dispersion ST code [12], [13] that employs carefully designed dispersion matrixes to optimize the mutual information between the transmitted and received signals.

Most good ST codes are designed for specific numbers of transmit antennas, and a unified approach is lacking for systems with arbitrary numbers of transmit antennas. The BLAST architectures are less effective in multiple-input–single-output (MISO) environments. Also, synchronous transmission is a standard assumption for most ST coding and BLAST systems but, in practice, intersymbol interference (ISI) and synchronization error are often inevitable, especially in *ad hoc* networks and relay systems [14].

The interleave-division-multiplexing space–time (IDM-ST) scheme proposed in [15] provides a potential solution to the above difficulties. The IDM-ST scheme is applicable for an arbitrary number (denoted by  $N$  hereafter) of transmit antennas. The related decoding complexity (per coded bit) is only  $O(N)$ , which is considerably lower than that of many other alternatives [such as trellis-based ST codes with decoding complexity  $O(2^N)$ ]. An IDM-ST code is conceptually very simple: we employ several forward error correction (FEC) codes (each referred to as a layer), and transmit the interleaved versions of their coded sequences simultaneously from all antennas. Interleaving here provides the only means to separate the concurrently transmitted signals. Our motivation is random coding, following the success of turbo and low-density parity-check (LDPC) codes [16], [17].

In this paper, we provide a comprehensive study for the IDM-ST scheme. Our main contributions are listed as follows.

- We develop a signal-to-noise-ratio (SNR) evolution technique that is closely related to the extrinsic information transfer (EXIT) chart method [18] and the related SNR evolution techniques [19]–[22]. These methods carry out performance analysis using presimulated transfer functions in additive white Gaussian noise (AWGN) channels. [**Note:** In this case, each presimulated transfer function can be stored as a one-dimensional (1-D) table.] The performance evaluation of an ST code, however, involves multiple fading coefficients (for multiple antennas). In this case, the transfer function is multidimensional (for all possible combinations of fading coefficients). It is usually impractical to presimulate and store such a transfer function. To overcome this difficulty, we develop a solution in which one of the transfer functions is generated analytically online (rather than presimulated) using a bounding technique. (The other transfer function related to the

Manuscript received September 20, 2005; revised June 19, 2007. This work was supported by the Research Grant Council of the Hong Kong SAR, China, under Grant CityU 117007.

K. Y. Wu is with the Research and Innovation Center, Alcatel Shanghai Bell Co., Ltd., Shanghai 201206, China.

L. Ping is with the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong (e-mail: eeliping@cityu.edu.hk).

Communicated by Ø. Ytrehus, Associate Editor for Coding Techniques.

Digital Object Identifier 10.1109/TIT.2007.915889

decoder is approximately independent of the channel condition.) This technique provides a simple, fast, and relatively accurate technique for performance evaluation.

- Using the fast performance assessment technique, we examine several power allocation strategies to maximize power efficiency as well as spectral efficiency. These strategies are much more efficient than the simulation based searching method discussed in [15].
- We prove in Appendix C that the IDM-ST scheme is capacity achieving provided that an ideal conventional FEC code is used. We will also show by simulation results that the practical coded IDM-ST schemes can achieve performance close to the theoretical limit.
- We develop an equalization technique for IDM-ST codes in ISI channels following the linear minimum mean square error (LMMSE) and probabilistic data association (PDA) principles [23]–[26]. Two efficient implementation techniques using Cholesky decomposition and frequency-domain equalization (FDE), respectively, are outlined. The proposed approach requires only approximate synchronization and may find useful applications in *ad hoc* and relay networks.

## II. TRANSMITTER PRINCIPLES

### A. Motivation

The motivation of this work comes from “random codes” [27]. A random code without any structure can achieve the Shannon limit but is very difficult to encode and decode in practice, so it has long been regarded as a purely theoretical concept. The recent success of turbo and LDPC codes has demonstrated the impact of the random coding concept on practical code design. Its application in ST coding is the focus of this paper.

Consider a 1-D conventional code represented by a length- $I$  sequence

$$s_1, s_2, \quad \dots \quad s_i, s_{i+1}, \quad \dots \quad s_{I-1}, s_I$$

and a two-dimensional (2-D) ST code represented by an  $N \times J$  matrix

$$\begin{matrix} s_1^{(1)} & \dots & s_j^{(1)} & \dots & s_J^{(1)} \\ \vdots & & \vdots & & \vdots \\ s_1^{(n)} & \dots & s_j^{(n)} & \dots & s_J^{(n)} \\ \vdots & & \vdots & & \vdots \\ s_1^{(N)} & \dots & s_j^{(N)} & \dots & s_J^{(N)} \end{matrix}$$

where  $N$  is the number of transmit antennas, and  $I$  and  $J$  are both code lengths. A code (either 1-D or 2-D) is said to be “random” if its elements are independent and identically distributed (i.i.d.). In many cases, the Gaussian distribution is optimal for mutual information maximization [27]. Clearly, a  $N \times J$  2-D random code can be generated by segmenting a length- $(N \times J)$  1-D random code into  $N$  rows, which suggests a simple way to form a 2-D random code from a 1-D one. This is the rationale behind the scheme presented below.

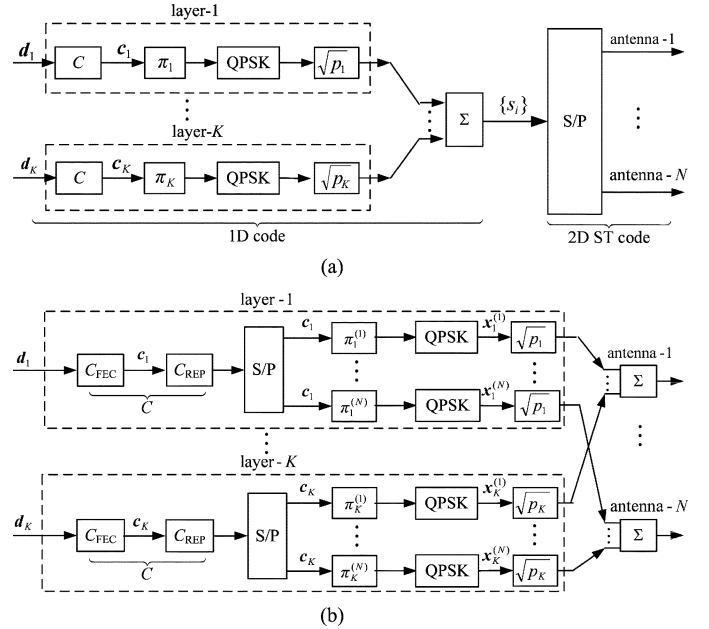


Fig. 1. (a) Transmitter structure of a superposition IDM-ST code, where  $\pi_k$  is the interleaver for layer- $k$ , “QPSK” denotes a QPSK modulator, and “S/P” denotes a serial-to-parallel converter. (b) Transmitter structure of a repetition-superposition IDM-ST code, where  $C_{\text{FEC}}$  denotes a binary FEC code and  $C_{\text{REP}}$  a length- $N$  repetition code.

### B. Transmitter Structures

The IDM-ST coding scheme outlined below is closely related to superposition coding [27], interleave-division multiple-access (IDMA) [28], and the related multiple access methods [29]–[31]. We examine two slightly different encoder structures.

1) *Superposition Encoder in Fig. 1(a)*: The inputs are  $K$  equal-length sequences  $\{\mathbf{d}_k\}$ , each encoded individually using a binary FEC code, generating  $\mathbf{c}_k$ . The signals from the same encoder are referred to as a layer. For simplicity, we use the same FEC code  $C$  for all layers. Each  $\mathbf{c}_k$  is interleaved, quaternary phase-shift keying (QPSK) modulated and scaled by an amplitude factor  $\sqrt{p_k}$ . Then,  $K$  layers are linearly superimposed and segmented in the serial-to-parallel (S/P) converter into  $N$  equal-length sections to be transmitted from  $N$  antennas simultaneously. The use of interleavers and power controllers is to facilitate the iterative detection process, which will be discussed in detail in Sections III and V. Denote by  $R_C$  the rate of  $C$  and  $R$  the overall rate. Then,  $R = 2NK R_C$ .

2) *Repetition-Superposition Encoder in Fig. 1(b)*: Fig. 1(b) is a modified form of Fig. 1(a), where the S/P converters are moved to the front of the interleavers. We assume the following: 1)  $C$  is constructed by serially concatenating a binary FEC code  $C_{\text{FEC}}$  and a length- $N$  repetition code  $C_{\text{REP}}$ ; and 2) the S/P converters are designed such that  $N$  replicas of each FEC-coded bit are transmitted from  $N$  different antennas. These two constraints introduce symmetry in the proposed scheme and make the analysis and design issues easier (see Section IV). Let the outputs of  $C_{\text{FEC}}$  for layer  $k$  be  $\mathbf{c}_k = \{c_{k,i}\}$  ( $c_{k,i} \in \{+1, -1\}$ ). The  $n$ th replica of  $\mathbf{c}_k$  is independently interleaved by an interleaver  $\pi_k^{(n)}$  and QPSK modulated, producing  $\mathbf{x}_k^{(n)} \equiv \{x_{k,i}^{(n)}\}$  to be transmitted from the  $n$ th antenna. The interleavers  $\{\pi_k^{(n)}\}$

for different layers and replicas are generated independently and randomly. Signals in  $\{\mathbf{x}_k^{(n)}, n = 1, \dots, N\}$  are scaled by a common amplitude factor  $\sqrt{p_k}$ . For the  $n$ th transmit antenna, the transmitted signal is  $\sum_{k=1}^K \sqrt{p_k} \mathbf{x}_k^{(n)}$ . Denote by  $R_C$  the rate of  $C_{\text{FEC}}$ . The overall transmission rate is  $R = 2KR_C$ .

Suppose that  $C$  in Fig. 1(a) is an ideal “random” binary code. From the central limit theorem, after superposition coding, the resultant signals are approximately Gaussian when  $K$  is sufficiently large. Thus, Fig. 1(a) is an approximate realization of a random code with Gaussian signaling [27]. It is reasonable to expect that “nearly random” binary codes, such as turbo and LDPC codes, may also be used as  $C$  in this scheme.

Although the repetition operation in Fig. 1(b) implies suboptimality in terms of coding gain, later we will show in Section IV-D that this scheme can still achieve relatively good performance, even with common codes, e.g., convolutional codes. We will also show in Appendix C that it is capacity achieving when  $R_C \rightarrow 0$  (which implies that, assuming QPSK,  $K$  should be sufficiently large if the designated overall rate  $2KR_C$  is finite).

The “layer” used above is different from that of the BLAST scheme [8], [9]. Each layer above is transmitted from the  $N$  antennas simultaneously, and therefore, there is interference among the signals in the same layer. There is no limit on the number of layers in the IDM-ST scheme. On the other hand, in the BLAST scheme, there is no interference among signals in the same layer and so the number of layers is limited by the number of antennas  $N$ . This implies that the proposed IDM-ST scheme can potentially achieve a higher throughput by increasing the number of layers, but the interference problem needs to be addressed properly.

In the following, we first discuss the scheme in Fig. 1(b), because it is easier to analyze. Later, in Section IV-E, we will demonstrate the advantage of removing repetition coding.

### III. RECEIVER PRINCIPLES

We concentrate on MISO systems with  $N$  transmit antennas and one receive antenna (an  $N \times 1$  system) in quasi-static Rayleigh fading channels, where the fading coefficients remain unchanged during one frame and vary independently from frame to frame. We will only consider single-path channels without ISI in this section. Let  $\alpha^{(n)}$  ( $n = 1, \dots, N$ ) be the fading coefficient between the  $n$ th transmit antenna and the receive antenna.  $\{\alpha^{(n)}\}$  are modeled as independent complex Gaussian random variables with zero-mean and variance  $\sigma^2 = 0.5$  per dimension. We always assume no channel state information (CSI) at the transmitter and ideal CSI at the receiver.

#### A. Overall Receiver Structure

Consider the transmitter structure in Fig. 1(b). The signal received at time  $j$  is

$$y_j = \sum_{n=1}^N \alpha^{(n)} \sum_{k=1}^K \sqrt{p_k} x_{k,j}^{(n)} + n_j \quad (1)$$

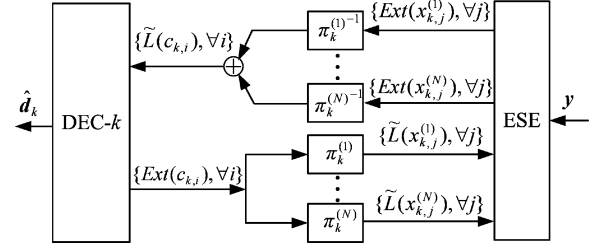


Fig. 2. Part of the receiver structure of the IDM-ST code in Fig. 1(b) for layer- $k$ , where  $\mathbf{y} = \{y_j\}$ , and  $\pi_k^{(n)}$  and  $\pi_k^{(n-1)}$  are the interleaver and deinterleaver for layer- $k$  on the  $n$ th transmit antenna.

where  $n_j$  is a sample of a complex AWGN process with variance  $\sigma^2 = N_0/2$  per dimension. We employ a suboptimal iterative receiver, which consists of an elementary signal estimator (ESE) module and  $K$  a posteriori probability (APP) decoders (DECs) operating iteratively [15], [22], [28]. The function of the ESE is to resolve the cross-antenna interference and cross-layer interference together and generate coarse estimates  $\{\text{Ext}(x_{k,j}^{(n)})\}$  of  $\{x_{k,j}^{(n)}\}$ . The outputs of the ESE are combined to form soft estimates  $\{\tilde{L}(c_{k,i})\}$  of the coded bits  $\{c_{k,i}\}$  in Fig. 1(b), which is equivalent to the APP decoding of the repetition code. The DECs carry out APP decoding of the FEC code for each layer based on  $\{\tilde{L}(c_{k,i})\}$  and generate extrinsic log-likelihood ratios (LLRs)  $\{\text{Ext}(c_{k,i})\}$  for  $\{c_{k,i}\}$  that are then used to refine the ESE output in the next iteration (see Appendix A). Fig. 2 illustrates a part of the receiver structure in which only the DEC for layer- $k$  (denoted by DEC- $k$ ) is shown. The DECs for other layers are connected to the ESE in the same way as DEC- $k$ . For the scheme in Fig. 1(a), the decoding process is similar, except that the combination operation can be skipped.

#### B. ESE Function

First, we assume BPSK modulation. Consider a particular symbol  $x_{k,j}^{(n)} \in \{+1, -1\}$ . Define  $r_j = \alpha^{(n)} y_j$ , where  $\alpha^{(n)}$  denotes the conjugate of  $\alpha^{(n)}$ . From (1), we have

$$r_j = \left| \alpha^{(n)} \right|^2 \sqrt{p_k} x_{k,j}^{(n)} + \zeta_{k,j}^{(n)} \quad (2)$$

where the distortion (interference-plus-noise) term  $\zeta_{k,j}^{(n)} = \frac{1}{\alpha^{(n)}} \left( \sum_{(n',k') \neq (n,k)} \alpha^{(n')} \sqrt{p_{k'}} x_{k',j}^{(n')} + n_j \right)$  is approximately modeled as a complex Gaussian random variable. From (2), the estimate of  $x_{k,j}^{(n)}$  is given by [15], [28]

$$\begin{aligned} \text{Ext}(x_{k,j}^{(n)}) &\equiv \log \left( \frac{p(y_j | x_{k,j}^{(n)} = +1)}{p(y_j | x_{k,j}^{(n)} = -1)} \right) \\ &= 2 \left| \alpha^{(n)} \right|^2 \sqrt{p_k} \cdot \frac{\text{Re}(r_j) - \text{E}(\text{Re}(\zeta_{k,j}^{(n)}))}{\text{Var}(\text{Re}(\zeta_{k,j}^{(n)}))} \end{aligned} \quad (3)$$

where  $\text{Re}(\cdot)$  and  $\text{Im}(\cdot)$  indicate the real and imaginary parts, and  $\text{E}(\cdot)$  and  $\text{Var}(\cdot)$  denote the mean and the variance. In Appendix A, we outline a procedure to estimate the mean and the variance of  $\zeta_{k,j}^{(n)}$  using the feedback of DECs. It is also shown in Appendix A that the complexity per FEC-coded bit  $c_{k,i}$  involved in (3) is independent of the number of layers  $K$ , and grows linearly with  $N$ .

With QPSK modulation,  $x_{k,j}^{(n)} = \text{Re}(x_{k,j}^{(n)}) + i\text{Im}(x_{k,j}^{(n)})$ , where  $\text{Re}(x_{k,j}^{(n)})$  and  $\text{Im}(x_{k,j}^{(n)})$  are two bits in  $c_k$ . Similarly to (3), we have [22], [28]

$$\text{Ext}(\text{Re}(x_{k,j}^{(n)})) = 2 \left| \alpha^{(n)} \right|^2 \sqrt{p_k} \cdot \frac{\text{Re}(r_j) - \text{E}(\text{Re}(\zeta_{k,j}^{(n)}))}{\text{Var}(\text{Re}(\zeta_{k,j}^{(n)}))}. \quad (4)$$

$\text{Ext}(\text{Im}(x_{k,j}^{(n)}))$  can be calculated by replacing “Re” with “Im” in (4).

#### IV. PERFORMANCE ANALYSIS AND OPTIMIZATION

The performance assessment of ST codes is generally a difficult issue. Simulation-based methods are time consuming. The rank and determinant criteria introduced in [4] can only provide coarse performance predictions (e.g., for the asymptotic slope). We now consider an SNR evolution approach [18]–[22]. The main challenge here is that the channel involves multidimensional fading distributions. (The work in [18]–[22] involves only one dimensional fading.) The novelty of the approach below is the bounding technique in (14) that avoids the evaluation of multidimensional functions, leading to a simple, fast, and reasonably accurate solution.

##### A. Performance Analysis With Fixed Fading Coefficients

All discussions in this section are for QPSK modulation with Gray mapping. The combination operation in Fig. 2 can be expressed as (5) shown at the bottom of the page [based on (2) and (4)], where  $S_R(c_{k,i})$  and  $S_I(c_{k,i})$  are the index sets of the  $N$  replicas of  $c_{k,i}$  in  $\{\text{Re}(x_{k,j}^{(n)}), \forall n, j\}$  and  $\{\text{Im}(x_{k,j}^{(n)}), \forall n, j\}$ , respectively. Apart from a scaling factor of 2, (5) can be regarded as a maximum ratio combining (MRC) of  $N$  independently distorted signals. Then, the SNR of  $\tilde{L}(c_{k,i})$ , denoted by  $\text{snr}_{k,i}$ , is given by [32]

$$\text{snr}_{k,i} = \sum_{(n,j) \in S_R(c_{k,i})} \text{snr}_{\text{Re-}k,j}^{(n)} + \sum_{(n,j) \in S_I(c_{k,i})} \text{snr}_{\text{Im-}k,j}^{(n)} \quad (6)$$

where  $\text{snr}_{\text{Re-}k,j}^{(n)}$  and  $\text{snr}_{\text{Im-}k,j}^{(n)}$  denote the SNRs of  $\text{Ext}(\text{Re}(x_{k,j}^{(n)}))$  and  $\text{Ext}(\text{Im}(x_{k,j}^{(n)}))$ , which can be calculated based on (4) as [22]

$$\text{snr}_{\text{Re-}k,j}^{(n)} = \frac{|\alpha^{(n)}|^4 p_k}{\text{Var}(\text{Re}(\zeta_{k,j}^{(n)}))} \quad \text{and} \quad \text{snr}_{\text{Im-}k,j}^{(n)} = \frac{|\alpha^{(n)}|^4 p_k}{\text{Var}(\text{Im}(\zeta_{k,j}^{(n)}))}. \quad (7)$$

Here,  $\{\text{Var}(\text{Re}(\zeta_{k,j}^{(n)}))\}$  and  $\{\text{Var}(\text{Im}(\zeta_{k,j}^{(n)}))\}$  are functions of the feedbacks  $\{\tilde{L}(\text{Re}(x_{k,j}^{(n)}))\}$  and  $\{\tilde{L}(\text{Im}(x_{k,j}^{(n)}))\}$  from the DEC's [see (A1), (A2), and (A8)]. We treat these feedbacks as random variables, so  $\{\text{snr}_{k,i}\}$ ,  $\{\text{Var}(\text{Re}(\zeta_{k,j}^{(n)}))\}$ , and  $\{\text{Var}(\text{Im}(\zeta_{k,j}^{(n)}))\}$  are all random variables. Define  $\text{snr}_k \equiv \text{E}(\text{snr}_{k,i})$ , where the expectation is over the distribution of the DEC feedbacks. Then

$$\begin{aligned} \text{snr}_k &\stackrel{(a)}{\geq} \sum_{(n,j) \in S_R(c_{k,i})} \frac{|\alpha^{(n)}|^4 p_k}{\text{E}(\text{Var}(\text{Re}(\zeta_{k,j}^{(n)})))} \\ &\quad + \sum_{(n,j) \in S_I(c_{k,i})} \frac{|\alpha^{(n)}|^4 p_k}{\text{E}(\text{Var}(\text{Im}(\zeta_{k,j}^{(n)})))} \\ &\stackrel{(b)}{=} \sum_{n=1}^N \frac{|\alpha^{(n)}|^2 p_k}{\sum_{(k',n') \neq (k,n)} |\alpha^{(n')}|^2 p_{k'} v_{k'} + \sigma^2} \end{aligned} \quad (8)$$

where

$$v_k \equiv \text{E}(\text{Var}(c_{k,i})) \quad (9)$$

is the average variance of  $\{c_{k,i}, \forall i\}$  calculated based on the output of DEC- $k$ . Note that  $v_k$  represents the uncertainty of signals from layer- $k$  after FEC decoding and determines the remaining interference of layer- $k$  to other layers in the current iteration. For the derivation of (8), (a) follows from the Jensen's inequality [22] and (b) is detailed in Appendix A [see (A14)]. Intuitively, the numerator in (8) denotes the desired signal power, and the denominator includes the power of the interference from other antennas and layers, plus the AWGN power. The simple expression in (8) is the consequence of (6), where the summation is over all fading coefficients, which is in tune with the consequence of the repetition coding in Fig. 1(b).

Define

$$\gamma_k \equiv \sum_{n=1}^N \frac{|\alpha^{(n)}|^2 p_k}{\sum_{(k',n') \neq (k,n)} |\alpha^{(n')}|^2 p_{k'} v_{k'} + \sigma^2}. \quad (10)$$

Clearly,  $\gamma_k$  gives a lower bound of  $\text{snr}_k$  according to (8). In the following, we use this lower bound to approximate  $\text{snr}_k$ , because it is easy to calculate. Recall that  $v_k$  is calculated based on the output of DEC- $k$  (see Appendix A), so it is a function of the average input SNR to DEC- $k$ , i.e.,

$$v_k = f(\gamma_k). \quad (11)$$

$$\begin{aligned} \tilde{L}(c_{k,i}) &= \sum_{(n,j) \in S_R(c_{k,i})} \text{Ext}(\text{Re}(x_{k,j}^{(n)})) + \sum_{(n,j) \in S_I(c_{k,i})} \text{Ext}(\text{Im}(x_{k,j}^{(n)})) \\ &= \sum_{(n,j) \in S_R(c_{k,i})} \frac{2|\alpha^{(n)}|^2 \sqrt{p_k}}{\text{Var}(\text{Re}(\zeta_{k,j}^{(n)}))} \left( |\alpha^{(n)}|^2 \sqrt{p_k} \text{Re}(x_{k,j}^{(n)}) + \text{Re}(\zeta_{k,j}^{(n)}) - \text{E}(\text{Re}(\zeta_{k,j}^{(n)})) \right) \\ &\quad + \sum_{(n,j) \in S_I(c_{k,i})} \frac{2|\alpha^{(n)}|^2 \sqrt{p_k}}{\text{Var}(\text{Im}(\zeta_{k,j}^{(n)}))} \left( |\alpha^{(n)}|^2 \sqrt{p_k} \text{Im}(x_{k,j}^{(n)}) + \text{Im}(\zeta_{k,j}^{(n)}) - \text{E}(\text{Im}(\zeta_{k,j}^{(n)})) \right) \end{aligned} \quad (5)$$

Combining (10) and (11), we have

$$\gamma_{k\_new} = \sum_{n=1}^N \frac{|\alpha^{(n)}|^2 p_k}{\sum_{(k',n') \neq (k,n)} |\alpha^{(n')}|^2 p_{k'} f(\gamma_{k'\_old}) + \sigma^2} \quad (12)$$

where  $\gamma_{k\_new}$  and  $\gamma_{k\_old}$  are the values of  $\gamma_k$  after and before one iteration, with initialization  $f(\gamma_{k\_old}) = 1$  for  $\forall k$ , implying no feedback from the DECS. Repeating (12), we can track the SNR evolution in the iterative process with any given  $\boldsymbol{\alpha} \equiv [\alpha^{(1)} \dots \alpha^{(N)}]$ . After the final iteration, we can estimate the frame-error-rate performance of layer- $k$  using  $\gamma_{k\_final}$ . If the code length is infinite, the performance estimated in this way is an upper bound of the true performance as  $\gamma_k$  is a lower bound of  $\text{snr}_k$ . However, for a finite code length, it can be only treated as an approximation due to the correlation problem in the iterative decoding process.

### B. Performance Analysis in Quasi-Static Rayleigh Fading Channels

The previous discussion is for fixed fading coefficients  $\boldsymbol{\alpha}$ . We now consider a quasi-static Rayleigh fading channel where  $\boldsymbol{\alpha}$  remains unchanged during one frame and varies independently from frame to frame. Considering all possible  $\boldsymbol{\alpha}$ , the average frame-error-rate of layer- $k$  in an IDM-ST code can be estimated as

$$\text{FER}_{k\_fading} = \int g(\gamma_{k\_final}) p(\boldsymbol{\alpha}) d\boldsymbol{\alpha} \quad (13)$$

where  $g(\cdot)$  denotes the relationship between frame-error-rate and  $\gamma_{k\_final}$ , and  $p(\boldsymbol{\alpha}) = p(\alpha^{(1)}) \dots p(\alpha^{(N)})$  is the joint probability density function (pdf) of  $\{\alpha^{(1)}, \dots, \alpha^{(N)}\}$ . (**Note:**  $\gamma_{k\_final}$  is a function of  $\boldsymbol{\alpha}$ .) The computation of (13) is relatively difficult since it involves an  $N$ -fold multiple integral. To avoid this difficulty, we introduce a bounding technique. Denote  $\lambda \equiv \sum_n |\alpha^{(n)}|^2 = \boldsymbol{\alpha} \boldsymbol{\alpha}^H$ . In Appendix B, we show that  $\gamma_k$  in (10) can be lower bounded as

$$\gamma_k \geq \frac{\lambda p_k}{\lambda \sum_{k'} p_{k'} v_{k'} - \lambda p_k v_k / N + \sigma^2} \quad (14)$$

where the equality holds when  $|\alpha^{(1)}|^2 = \dots = |\alpha^{(N)}|^2 = \lambda/N$ , corresponding to the uniform-fading situation when all  $\{\alpha^{(n)}\}$  have the same power gain.

It can be verified that  $g(\cdot)$  in (13) is a monotonically decreasing function, so  $\text{FER}_{k\_fading}$  in (13) can be upper bounded as

$$\text{FER}_{k\_fading} \leq \int g(\bar{\gamma}_{k\_final}) p(\lambda) d\lambda \quad (15)$$

where  $p(\lambda) = \lambda^{N-1} e^{-\lambda} / (N-1)!$  is the pdf of  $\lambda$  and  $\bar{\gamma}_{k\_final}$  is calculated using the following iteration:

$$\bar{\gamma}_{k\_new} = \frac{\lambda p_k}{\lambda \sum_{k'} p_{k'} f(\bar{\gamma}_{k'\_old}) - \lambda p_k f(\bar{\gamma}_{k\_old}) / N + \sigma^2} \quad (16)$$

with  $f(\bar{\gamma}_{k\_old})$  initialized to 1 for  $\forall k$ .

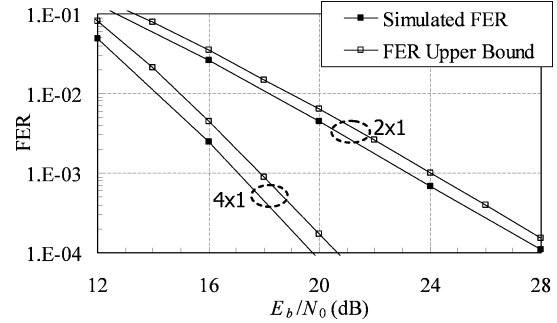


Fig. 3. Comparison between the upper bound and simulation results for  $2 \times 1$  and  $4 \times 1$  IDM-ST codes based on Fig. 1(b) with  $R = 4$  bits per channel use.

Compared with (13), the upper bound in (15) is much easier to calculate as it involves only a onefold integral.

Note that several approximations are used above, so it is necessary to examine the usefulness of the result. Fig. 3 compares the simulation and upper bound results for  $2 \times 1$  and  $4 \times 1$  IDM-ST codes based on Fig. 1(b) with  $R = 4$  bits per channel use. The detailed parameters will be explained in Section IV-D. As we can see, the upper bounds are reasonably tight.

### C. Power Allocation

We now apply the previous SNR evolution technique to IDM-ST code design. It can be verified that, for any fixed  $\boldsymbol{\alpha}$  and  $P \equiv \sum_k p_k$ , the outcome of (13) depends on the distribution of  $\{p_1, \dots, p_K\}$ , so proper power allocation can be employed to optimize IDM-ST codes. The intuition is that, with a high probability, strong signals can be correctly detected first so their interference to weak signals can be canceled, which in turn facilitates the detection of weak signals. Similar methods have been studied for other ST and BLAST schemes [33], [34].

With power allocation, we search for  $\{p_k\}$  that optimizes the system performance at fixed total power  $P = \sum_k p_k$ . The simulation-based methods discussed in [15] are usually very time consuming. Instead, we carry out searching using the fast evaluation technique derived previously.

### D. Numerical Examples

We now provide some numerical results to demonstrate the performance advantage gained by optimizing power allocation. Let  $N_{\text{info}}$  be the number of information bits per layer in a frame,  $R$  the overall transmission rate,  $R_C$  the rate of the FEC code, and  $It$  the iteration number. We apply different rotations  $\{0, \pi/2K, \dots, (K-1)\pi/2K\}$  to signals from different layers to make the interference from other layers more Gaussian-like. We call each  $\mathbf{d}_k$  a frame and  $\{\mathbf{d}_k, k = 1, \dots, K\}$  a superframe. The frame error rate (FER) and superframe error rate (SFER) are defined accordingly. Clearly,  $\text{SFER} \geq \text{FER}$ .

Consider IDM-ST coding systems in Fig. 1(b) with  $N = 2$  and 4,  $N_{\text{info}} = 4096$ ,  $R =$  (a) 2 and (b) 4 bits per channel use, corresponding to  $K = 3$  and 6, respectively. A rate-1/3 turbo code with generator  $G(x) = (1+x+x^3)/(1+x^2+x^3)$  is used for all layers, so the frame length  $= N_{\text{info}} / (2 \times R_C) = 6144$  symbols.

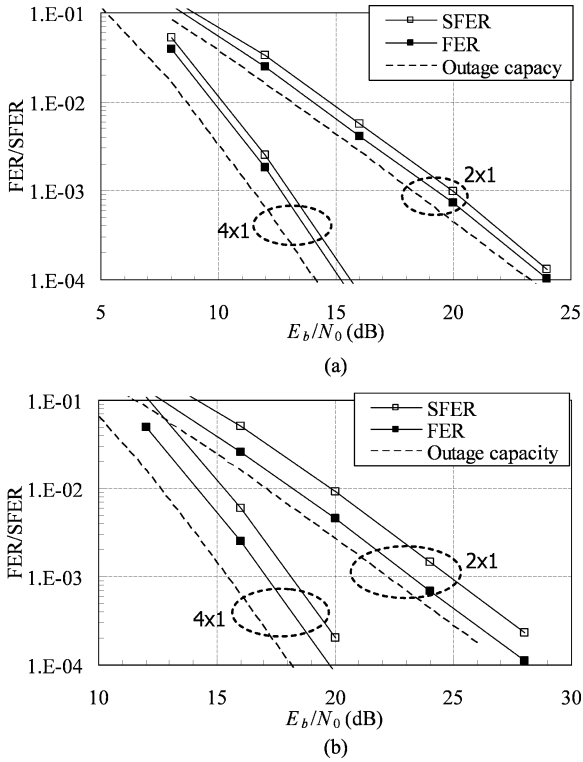


Fig. 4. FER and SFER performance of  $2 \times 1$  and  $4 \times 1$  IDM-ST coding systems in Fig. 1(b) with (a)  $R = 2$  and (b)  $R = 4$  bits per channel use over a quasi-static Rayleigh fading channel.  $N_{\text{info}} = 4096$ ,  $It = 40$ . A rate-1/3 turbo code with generator  $G(x) = (1 + x + x^3)/(1 + x^2 + x^3)$  is used for all layers.

TABLE I  
POWER LEVELS FOR TURBO-CODED IDM-ST SCHEMES

$(N, R, K)$	Power Levels
(2, 2, 3)	0.1617P, 0.3003P, 0.538P
(4, 2, 3)	0.1737P, 0.2983P, 0.528P
(2, 4, 6)	0.0242P, 0.045P, 0.0806P, 0.1439P, 0.2543P, 0.452P
(4, 4, 6)	0.0268P, 0.0461P, 0.0809P, 0.1421P, 0.2521P, 0.452P

The layer-by-layer search technique introduced in [15] is always used. Table I lists the power levels obtained. Fig. 4 shows the simulated performance. Corresponding outage capacity [1], [2], [35] [see (C13) in Appendix C] are included for reference, which serve as theoretical limits for the SFER performance. We can see that the FER performance of the IDM-ST codes is quite close to the corresponding outage capacity (within 1.5 dB). SFER curves are 0.5 ~ 1.5 dB worse than the FER curves. In practice, FER can be a more useful performance measure than SFER, as in case of error it is only necessary to discard erroneous frames, instead a complete superframe. Experimentally, we observed that the layers with lower power levels have higher FER.

It is also interesting to examine the effects of different FEC codes on the performance of the IDM-ST scheme. Fig. 5 shows a performance comparison between IDM-ST coding systems based on turbo and convolutional codes with  $R = 4$  bits per channel use. The turbo-coded systems have the same parameters as in Fig. 4(b). The convolutionally coded systems have

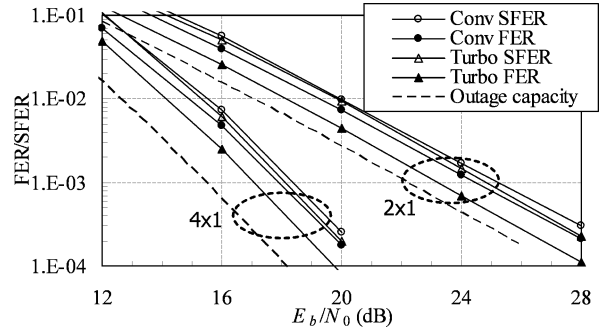


Fig. 5. Performance comparison between the turbo-coded and convolutionally coded IDM-ST coding systems with  $R = 4$  bits per channel use.

TABLE II  
POWER LEVELS FOR THE CONVOLUTIONALLY CODED IDM-ST SCHEMES

$(N, R, K)$	Power Levels
(2, 4, 4)	0.0701P, 0.101P, 0.2769P, 0.552P
(4, 4, 4)	0.0764P, 0.1056P, 0.264P, 0.554P

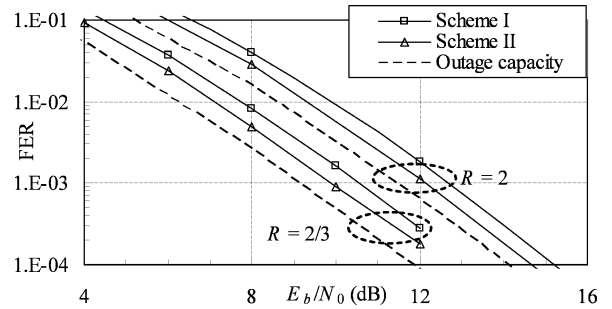


Fig. 6. Performance comparison between the IDM-ST systems based on Fig. 1(a) and (b) over quasi-static Rayleigh fading channels with  $N = 4$ ,  $N_{\text{info}} = 4096$ ,  $It = 40$ ,  $R = 2/3$  and 2 bits per channel use.

$N_{\text{info}} = 1024$ ,  $It = 20$ , and  $K = 4$ , and employ a rate-1/2 systematic convolutional code with generator  $G(x) = (1 + x + x^2 + x^4)/(1 + x^3 + x^4)$  for all layers. The power levels for the convolutionally coded systems are listed in Table II. As we can see, the performance of the two systems is quite close, while the turbo-coded systems demonstrate an advantage of about 1 dB in FER performance, mainly attributing to its higher coding gain.

E. Comparison Between the Two Transmitter Structures in Fig. 1(a) and (b)

The transmitter structure in Fig. 1(a) allows the use of more sophisticated low-rate codes with higher coding gain. We now compare the two alternatives in Fig. 1(a) and (b). Fig. 6 shows the FER performance for IDM-ST schemes based on Fig. 1(a) and (b) with  $N_{\text{info}} = 4096$ ,  $It = 40$ ,  $N = 4$ , and  $R = (a) 2/3$  and (b) 2 bits per channel use over quasi-static Rayleigh fading channels. The parameters for the FEC codes used in the two schemes are as follows.

1) *Repetition-Superposition IDM-ST Code in Fig. 1(b) (Scheme I)*: We employ the same rate-1/3 turbo code as that in Fig. 4.  $K = 1$  and 3 for  $R = 2/3$  and 2 bits per channel use, respectively. The power levels used for  $K = 3$  are listed in Table I. (Power allocation is unnecessary for  $K = 1$ .)

2) *Superposition IDM-ST Code in Fig. 1(a) (Scheme II)*: We employ a turbo-Hadamard code [36] with  $R_C = 1/12$ ,  $M = 3$ ,  $r = 4$ , and  $G(x) = 1/(1+x)$ , where  $M$  is the number of component codes,  $r$  is the order of the Hadamard code used, and  $G(x)$  is the generator of the convolutional code involved.  $K = 1$  and  $3$  for  $R = 2/3$  and  $2$  bits per channel use, respectively. The power levels used for  $K = 3$  are  $0.18P$ ,  $0.2823P$ , and  $0.5377P$ , which are obtained from a simulation-based search.

From Fig. 6, we can see that the scheme in Fig. 1(a) can further narrow the gap between the performance of IDM-ST codes and the outage capacity, thanks to the additional coding gain provided by the better low-rate FEC codes. However, the use of low-rate codes without repetition complicates the performance analysis issue (see Section IV-A), and makes power allocation a more difficult task. The power levels used in Fig. 6 are obtained from a simulation-based search. With a large number of layers, such a search method becomes too time consuming. Efficient solutions to this problem are still under investigation.

### V. IDM-ST CODES IN MULTIPATH CHANNELS

We now proceed to consider the performance of IDM-ST codes in multipath channels with ISI. The ESE function derived below is equivalent to the LMMSE [23] and PDA estimators [24], [25]. These methods all employ the Gaussian approximation [26] at different stages.

Our approach below is similar to (2), in which the Gaussian approximation is directly applied to the received signal. This is a more concise treatment. We will outline two efficient implementation techniques based on Cholesky and fast Fourier transform (FFT), respectively.

#### A. Channel Model

Consider an  $N \times 1$  system in a multipath quasi-static Rayleigh fading channel with  $L$  tap coefficients. Denote by  $\{\alpha_0^{(n)}, \dots, \alpha_l^{(n)}, \dots, \alpha_{L-1}^{(n)}\}$  the fading coefficients between the  $n$ th transmit antenna and the receive antenna.  $\{\alpha_l^{(n)}, n = 1, \dots, N, l = 0, \dots, L-1\}$  are modeled as i.i.d. complex Gaussian random variables with zero mean. We assume that the fading coefficients in any antenna link have a unit average power gain:  $E\left(\sum_{l=0}^{L-1} |\alpha_l^{(n)}|^2\right) = 1$  for  $\forall n$ . After sampling at the symbol rate, the signal received at time  $j$  can be represented by

$$y_j = \sum_{l=0}^{L-1} \left( \sum_{n=1}^N \alpha_l^{(n)} \sum_{k=1}^K \sqrt{p_k} x_{k,j-l} \right) + n_j. \quad (17a)$$

Let

$$\mathbf{\alpha}_l = [\alpha_l^{(1)} \cdots \alpha_l^{(n)} \cdots \alpha_l^{(N)}]$$

and

$$x_{k,j-l} = [x_{k,j-l}^{(1)} \cdots x_{k,j-l}^{(n)} \cdots x_{k,j-l}^{(N)}]^T.$$

Equation (17a) can be written in a more compact form as

$$y_j = \sum_{l=0}^{L-1} \mathbf{\alpha}_l \sum_{k=1}^K \sqrt{p_k} \mathbf{x}_{k,j-l} + n_j. \quad (17b)$$

Equation (17) can be further expressed in a matrix form as

$$\begin{bmatrix} y_1 \\ \vdots \\ y_j \\ \vdots \\ y_J \end{bmatrix} = \sum_{k=1}^K \sqrt{p_k} \begin{bmatrix} \mathbf{a}_0 & & & & \\ \vdots & \mathbf{a}_0 & & & 0 \\ \mathbf{a}_{L-1} & \vdots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ 0 & & \mathbf{a}_{L-1} & \cdots & \mathbf{a}_0 \end{bmatrix} \begin{bmatrix} \mathbf{x}_{k,1} \\ \vdots \\ \mathbf{x}_{k,j} \\ \vdots \\ \mathbf{x}_{k,J} \end{bmatrix} + \begin{bmatrix} n_1 \\ \vdots \\ n_j \\ \vdots \\ n_J \end{bmatrix} \quad (18)$$

where  $J$  is the frame length. Define

$$\mathbf{y}_j = \begin{bmatrix} \text{Re}(y_j) \\ \text{Im}(y_j) \end{bmatrix}, \quad \mathbf{x}_{k,j} = \begin{bmatrix} \text{Re}(\mathbf{x}_{k,j}) \\ \text{Im}(\mathbf{x}_{k,j}) \end{bmatrix},$$

$$\mathbf{A}_l = \begin{bmatrix} \text{Re}(\mathbf{\alpha}_l) & -\text{Im}(\mathbf{\alpha}_l) \\ \text{Im}(\mathbf{\alpha}_l) & \text{Re}(\mathbf{\alpha}_l) \end{bmatrix}, \quad \mathbf{n}_j = \begin{bmatrix} \text{Re}(n_j) \\ \text{Im}(n_j) \end{bmatrix}.$$

The complex channel model in (18) can be equivalently expressed by a real matrix equation as

$$\underbrace{\begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_j \\ \vdots \\ \mathbf{y}_J \end{bmatrix}}_{\tilde{\mathbf{y}}} = \sum_{k=1}^K \sqrt{p_k} \underbrace{\begin{bmatrix} \mathbf{A}_0 & & & & \\ \vdots & \mathbf{A}_0 & & & 0 \\ \mathbf{A}_{L-1} & \vdots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ 0 & & \mathbf{A}_{L-1} & \cdots & \mathbf{A}_0 \end{bmatrix}}_{\tilde{\mathbf{A}}} \underbrace{\begin{bmatrix} \mathbf{x}_{k,1} \\ \vdots \\ \mathbf{x}_{k,j} \\ \vdots \\ \mathbf{x}_{k,J} \end{bmatrix}}_{\tilde{\mathbf{x}}_k} + \underbrace{\begin{bmatrix} \mathbf{n}_1 \\ \vdots \\ \mathbf{n}_j \\ \vdots \\ \mathbf{n}_J \end{bmatrix}}_{\tilde{\mathbf{n}}} \quad (19a)$$

or in a more compact form as

$$\tilde{\mathbf{y}} = \sum_{k=1}^K \sqrt{p_k} \tilde{\mathbf{A}} \tilde{\mathbf{x}}_k + \tilde{\mathbf{n}}. \quad (19b)$$

Note that all variables in (19) are real.

#### B. ESE Function in Multipath Channels

Denote by  $x$  an entry of  $\tilde{\mathbf{x}}_k$  and  $\tilde{\alpha}$  the corresponding fading coefficients that form a column of  $\tilde{\mathbf{A}}$ . Concentrate on a particular  $x$ . Equation (19b) can be rewritten as

$$\tilde{\mathbf{y}} = \sqrt{p_k} \tilde{\alpha} \mathbf{x} + \zeta \quad (20a)$$

where

$$\boldsymbol{\zeta} = \sum_{k'=1}^K \sqrt{p_{k'}} \tilde{\mathbf{A}} \tilde{\mathbf{x}}_{k'} - \sqrt{p_k} \tilde{\boldsymbol{\alpha}} \mathbf{x} + \tilde{\mathbf{n}} \quad (20b)$$

is the distortion term in  $\tilde{\mathbf{y}}$  with respect to  $x$ . Similarly to (2), we again apply the central limit theorem. Now, considering the correlation introduced by ISI, we approximate  $\boldsymbol{\zeta}$  by a vector of joint multidimensional Gaussian random variables with mean vector and covariance matrix as

$$\mathbf{E}(\boldsymbol{\zeta}) = \mathbf{E}(\tilde{\mathbf{y}}) - \mathbf{E}(x) \sqrt{p_k} \tilde{\boldsymbol{\alpha}} \quad (21a)$$

$$\text{Cov}(\boldsymbol{\zeta}) = \mathbf{E}(\boldsymbol{\zeta} \boldsymbol{\zeta}^T) - \mathbf{E}(\boldsymbol{\zeta}) \mathbf{E}(\boldsymbol{\zeta})^T = \mathbf{R} - \text{Var}(x) p_k \tilde{\boldsymbol{\alpha}} \tilde{\boldsymbol{\alpha}}^T \quad (21b)$$

where

$$\mathbf{E}(\tilde{\mathbf{y}}) = \sum_{k=1}^K \sqrt{p_k} \tilde{\mathbf{A}} \mathbf{E}(\tilde{\mathbf{x}}_k) \quad (22a)$$

and

$$\begin{aligned} \mathbf{R} &\equiv \text{Cov}(\tilde{\mathbf{y}}) = \mathbf{E}(\tilde{\mathbf{y}} \tilde{\mathbf{y}}^T) - \mathbf{E}(\tilde{\mathbf{y}}) \mathbf{E}(\tilde{\mathbf{y}})^T \\ &= \sum_{k=1}^K p_k \tilde{\mathbf{A}} \text{Cov}(\tilde{\mathbf{x}}_k) \tilde{\mathbf{A}}^T + \sigma^2 \mathbf{I}. \end{aligned} \quad (22b)$$

In (22b),  $\text{Cov}(\tilde{\mathbf{x}}_k)$  is a diagonal matrix whose diagonal elements are the variances of the entries of  $\tilde{\mathbf{x}}_k$ . Similar to (3) and using the assumption that  $\boldsymbol{\zeta}$  is jointly Gaussian, the estimate of  $x$  can be calculated from (20) as (23), shown at the bottom of the page. The calculation of (23) is relatively complicated because of the inverse operation of  $\mathbf{R}^{-1}$ . In the following, we outline two fast implementation techniques.

1) *Cholesky Factorization Approach:* Rewrite (23) as

$$\text{Ext}(x) = 2\sqrt{p_k} \frac{\mathbf{g}^T \mathbf{f} + \mathbf{E}(x) \sqrt{p_k} \mathbf{g}^T \mathbf{g}}{1 - \text{Var}(x) p_k \mathbf{g}^T \mathbf{g}} \quad (24)$$

where  $\mathbf{g} \equiv \mathbf{L}^{-1} \tilde{\boldsymbol{\alpha}}$ ,  $\mathbf{f} \equiv \mathbf{L}^{-1} (\tilde{\mathbf{y}} - \mathbf{E}(\tilde{\mathbf{y}}))$ , and  $\mathbf{L}$  is a lower triangular band matrix obtained through the Cholesky factorization  $\mathbf{R} = \mathbf{L} \mathbf{L}^T$ . It is easy to verify that  $\mathbf{R}$  is a symmetric band matrix with bandwidth  $4L - 1$ , so the bandwidth of  $\mathbf{L}$  is  $2L$ . The cost for the Cholesky factorization is  $O(JL^2)$ , so the normalized cost per entry of  $\tilde{\mathbf{x}}_k$  is  $O(L^2)$ . The evaluation of  $\mathbf{g}$  and  $\mathbf{f}$  is equivalent to solving the following two equations:

$$\mathbf{L} \mathbf{g} = \tilde{\boldsymbol{\alpha}} \quad (25a)$$

$$\mathbf{L} \mathbf{f} = \tilde{\mathbf{y}} - \mathbf{E}(\tilde{\mathbf{y}}). \quad (25b)$$

The cost involved in (25b) is  $O(L)$  per entry of  $\tilde{\mathbf{x}}_k$ . The cost involved in (25a) is quite high because it should be carried out for every entry of  $\tilde{\mathbf{x}}_k$ . To reduce complexity, we adopt the following approximation [37]. Suppose that  $x$  is an element of  $\mathbf{x}_{k,j}$ , i.e.,  $x$  is transmitted at time  $j$ . Thus, the first  $2(j-1)$  entries of  $\tilde{\boldsymbol{\alpha}}$  are all zeros, and consequently, the first  $2(j-1)$  entries of  $\mathbf{g}$  are also zeros, because  $\mathbf{L}$  is lower triangular. Denote  $\mathbf{g} = [g_1, g_2, \dots, g_{2J}]^T$  and  $\mathbf{f} = [f_1, f_2, \dots, f_{2J}]^T$ . We introduce two approximations

$$\mathbf{g}^T \mathbf{f} = \sum_{j'=2j-1}^{2J} g_{j'} f_{j'} \approx \sum_{j'=2j-1}^{2(j+D)} g_{j'} f_{j'} \quad (26a)$$

and

$$\mathbf{g}^T \mathbf{g} = \sum_{j'=2j-1}^{2J} g_{j'}^2 \approx \sum_{j'=2j-1}^{2(j+D)} g_{j'}^2 \quad (26b)$$

where  $D$  is a properly selected positive integer. Then, we only need to calculate  $2(D+1)$  entries of  $\mathbf{g}$  [from  $g_{2j-1}$  to  $g_{2(j+D)}$ ] using (25a), with cost  $O(LD)$  per entry of  $\tilde{\mathbf{x}}_k$ . We have observed that  $D$  can be fairly small (say,  $D = 3L \sim 5L$ ), so that the complexity involved in (24) is approximately  $O(L^2)$  per entry of  $\tilde{\mathbf{x}}_k$ . For each FEC-coded bit (i.e., each  $c_{k,i}$ ), the overall complexity in (24) is  $O(NL^2)$  considering  $N$  replicas for each  $c_{k,i}$  in Fig. 1(b).

2) *FDE Approach:* Assume that synchronization error among the signals from different antennas can be treated together with multipath delay. We pad each  $\tilde{\mathbf{x}}_k$  with a cyclic prefix consisting of the last  $2 \cdot N \cdot (L-1)$  entries  $\{\mathbf{x}_{k,J-L+2}, \dots, \mathbf{x}_{k,J}\}$  of  $\tilde{\mathbf{x}}_k$ , which converts the channel coefficient matrices in (18) into circulant ones [38]. Then, an FFT-based evaluation technique can be applied to (23), which essentially leads to FDE. The complexity of this technique is independent of the channel memory length  $L$ . This is particularly advantageous when  $L$  is large. For details, see [39].

From the previous discussions, we can see that strict synchronization is not necessary for IDM-ST codes. This indicates the potential application of IDM-ST codes in *ad hoc* networks and relay systems, where strict frame synchronization is difficult to achieve [14].

### C. Numerical Results

We now use simulation results to demonstrate the performance of IDM-ST codes in multipath channels. Fig. 7 shows the FER performance of IDM-ST coding systems based on Fig. 1(b) with  $N = 2$ ,  $R =$  (a)  $2/3$  and (b) 2 bits per channel use over multipath quasi-static Rayleigh fading channels with  $L = 1, 2$ ,

$$\begin{aligned} \text{Ext}(x) &= \log \left( \frac{\exp \left( -\frac{1}{2} (\tilde{\mathbf{y}} - \mathbf{E}(\boldsymbol{\zeta}) - \sqrt{p_k} \tilde{\boldsymbol{\alpha}})^T (\text{Cov}(\boldsymbol{\zeta}))^{-1} (\tilde{\mathbf{y}} - \mathbf{E}(\boldsymbol{\zeta}) - \sqrt{p_k} \tilde{\boldsymbol{\alpha}}) \right)}{\exp \left( -\frac{1}{2} (\tilde{\mathbf{y}} - \mathbf{E}(\boldsymbol{\zeta}) + \sqrt{p_k} \tilde{\boldsymbol{\alpha}})^T (\text{Cov}(\boldsymbol{\zeta}))^{-1} (\tilde{\mathbf{y}} - \mathbf{E}(\boldsymbol{\zeta}) + \sqrt{p_k} \tilde{\boldsymbol{\alpha}}) \right)} \right) \\ &= 2\sqrt{p_k} \frac{\tilde{\boldsymbol{\alpha}}^T \mathbf{R}^{-1} (\tilde{\mathbf{y}} - \mathbf{E}(\tilde{\mathbf{y}})) + \mathbf{E}(x) \sqrt{p_k} \tilde{\boldsymbol{\alpha}}^T \mathbf{R}^{-1} \tilde{\boldsymbol{\alpha}}}{1 - \text{Var}(x) p_k \tilde{\boldsymbol{\alpha}}^T \mathbf{R}^{-1} \tilde{\boldsymbol{\alpha}}}. \end{aligned} \quad (23)$$

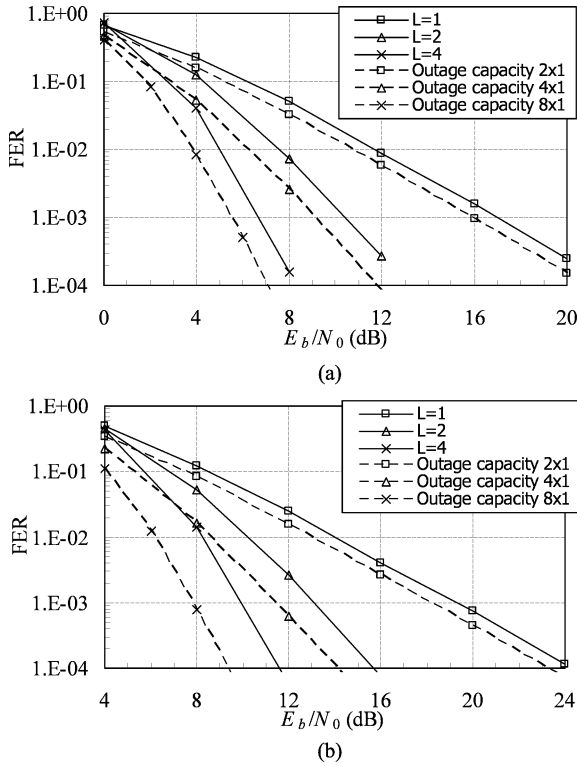


Fig. 7. Simulated FER performance of  $2 \times 1$  IDM-ST coding systems based on Fig. 1(b) with (a)  $R = 2/3$  and (b)  $R = 2$  bits per channel use in multipath quasi-static Rayleigh fading channels with  $L = 1, 2$ , and 4. The same FEC code as that in Fig. 4 is used.  $N_{\text{info}} = 4096$  and  $It = 40$ . The outage capacity curves for single-path channels are also included for reference.

and 4. All  $L$  tap coefficients have the same variance. The same turbo code as that in Fig. 4 is used, so  $K = 1$  and 3 for  $R = 2/3$  and 2 bits per channel use, respectively.  $N_{\text{info}} = 4096$ ,  $It = 40$ , and  $D = 4L$ . Power allocation is carried out in the same way as for single-path channels. Recall that the asymptotic slopes of FER curves show the degree of diversity achieved by the corresponding ST codes. We can see from Fig. 7 that the degree of diversity achieved by the IDM-ST code in a  $L$ -tap multipath channel is  $L$  times that in a single-path channel with the same number of antennas. For example, the slopes of the FER curves of  $2 \times 1$  IDM-ST codes in two-tap ISI channels are the same as those of the outage capacity curves of  $4 \times 1$  single-path channels. (The latter is shown by dashed lines in Fig. 7.) This indicates that the IDM-ST code can efficiently exploit the extra diversity provided by ISI channels.

### VI. CONCLUSION

In this paper, we studied a family of high-rate multilayer ST codes, referred to as the IDM-ST code. The design of IDM-ST codes takes a quasi-random approach by using random interleaving as the only means to separate signals from different layers and antennas. The resultant code bears a resemblance to the random ST code with i.i.d. Gaussian elements. The motivation of this work comes from the success of quasi-random FEC codes, such as turbo codes and LDPC codes.

We have demonstrated a number of advantages of IDM-ST codes, including the following:

- simple design procedure;

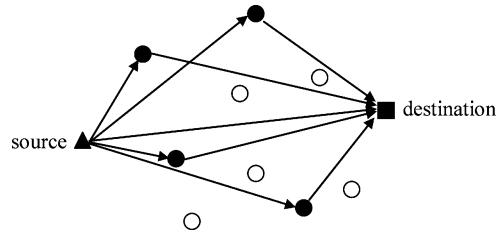


Fig. 8. Relay channel model, where black circles denote active relays and white circles denote inactive relays.

- low receiver complexity;
- applicable to arbitrary numbers of transmit antennas;
- easy to analyze and optimize;
- applicable to ISI channels and asynchronous environments;
- near-capacity performance at both high and low rates.

The IDM-ST code provides a promising and unified solution to ST coding for high-rate application with arbitrary numbers of transmit antennas. It also has good potential for application in *ad hoc* networks and relay channels. For example, consider a relay channel with a set of transmitting terminals. Each relay may become active or inactive depending on the channel condition without informing the source terminal, so the number of active relays may not be fixed in advance. Suppose that we want to apply ST coding techniques in such a scenario to improve diversity gain [14]. The destination terminal will detect simultaneously relayed signals. Strict synchronization is hard to achieve when relays are distributed in a large area. In this case, a traditional ST transmission scheme designed for a specific transmit antenna number in a frame-synchronous channel will encounter difficulties. For IDM-ST codes, however, the number of transmit antennas can be set freely, and frame synchronization error can be tolerated. This clearly indicates the advantages of the proposed IDM-ST code.

### APPENDIX A

#### GENERATION OF MEANS AND VARIANCES

In this appendix, we consider the calculation of the *a priori* mean and variance of the distortion term  $\zeta_{k,j}^{(n)}$  in (2), and the average distortion variance in (8), based on the feedback LLRs from the DEC. We focus on QPSK modulation and start with the *a priori* means and variances of the transmitted symbols  $\{x_{k,j}^{(n)}\}$ .

*A Priori Mean and Variance:* For simplicity, we only consider the real parts of  $\{x_{k,j}^{(n)}\}$ . The treatment for the imaginary parts is similar. The *a priori* means and variances of  $\{\text{Re}(x_{k,j}^{(n)})\}$  are calculated from their *a priori* information, denoted by  $\{\tilde{L}(\text{Re}(x_{k,j}^{(n)}))\}$ , as

$$\begin{aligned} E(\text{Re}(x_{k,j}^{(n)})) &\approx \frac{\exp(\tilde{L}(\text{Re}(x_{k,j}^{(n)}))) - 1}{\exp(\tilde{L}(\text{Re}(x_{k,j}^{(n)}))) + 1} \\ &= \tanh(\tilde{L}(\text{Re}(x_{k,j}^{(n)}))/2) \end{aligned} \quad (\text{A1})$$

and

$$\text{Var}(\text{Re}(x_{k,j}^{(n)})) = 1 - \left( \text{E}(\text{Re}(x_{k,j}^{(n)})) \right)^2. \quad (\text{A2})$$

At the start of decoding, we initialize  $\tilde{L}(\text{Re}(x_{k,j}^{(n)})) = 0$  for  $\forall k, n, j$ , indicating no *a priori* information about  $\{\text{Re}(x_{k,j}^{(n)})\}$ . During the iterative decoding process,  $\{\tilde{L}(\text{Re}(x_{k,j}^{(n)}))\}$  are updated as [assuming  $(n, j) \in S_R(c_{k,i})$ ]

$$\tilde{L}(\text{Re}(x_{k,j}^{(n)})) = \text{Ext}(c_{k,i}) \quad (\text{A3})$$

where  $\{\text{Ext}(c_{k,i})\}$  are the outputs of DECs.

*Distortion Mean:* The computation of  $\text{E}(\zeta_{k,j}^{(n)}) \equiv \text{E}(\text{Re}(\zeta_{k,j}^{(n)})) + i\text{E}(\text{Im}(\zeta_{k,j}^{(n)}))$  is quite straightforward from (2) as

$$\text{E}(\zeta_{k,j}^{(n)}) = \overline{\alpha^{(n)}} \left( \text{E}(y_j) - \alpha^{(n)} \sqrt{p_k} \text{E}(x_{k,j}^{(n)}) \right) \quad (\text{A4a})$$

where [from (1)]

$$\text{E}(y_j) = \sum_{n', k'} \alpha^{(n')} \sqrt{p_{k'}} \text{E}(x_{k',j}^{(n')}) \quad (\text{A4b})$$

with

$$\text{E}(x_{k,j}^{(n)}) = \text{E}(\text{Re}(x_{k,j}^{(n)})) + i\text{E}(\text{Im}(x_{k,j}^{(n)})).$$

We can take the real and imaginary parts to obtain  $\text{E}(\text{Re}(\zeta_{k,j}^{(n)}))$  and  $\text{E}(\text{Im}(\zeta_{k,j}^{(n)}))$ , respectively.

*Distortion Variance:* We define the covariance matrix for a complex random variable  $a$  as

$$\mathbf{Cov}(a) = \begin{pmatrix} \text{Var}(\text{Re}(a)) & \text{Cov}(\text{Re}(a), \text{Im}(a)) \\ \text{Cov}(\text{Re}(a), \text{Im}(a)) & \text{Var}(\text{Im}(a)) \end{pmatrix} \quad (\text{A5a})$$

where

$$\text{Cov}(b, c) = \text{E}(b \cdot c) - \text{E}(b)\text{E}(c). \quad (\text{A5b})$$

We also write  $\mathbf{Cov}(y_j)$ ,  $\mathbf{Cov}(x_{k,j}^{(n)})$ , and  $\mathbf{Cov}(\zeta_{k,j}^{(n)})$  depending on whether  $a$  is a received signal, transmitted signal, or distortion term. Then, it can be shown that

$$\mathbf{Cov}(y_j) = \sum_{n=1}^N \mathbf{R}^{(n)} \left( \sum_{k=1}^K p_k \mathbf{Cov}(x_{k,j}^{(n)}) \right) (\mathbf{R}^{(n)})^T + \sigma^2 \mathbf{I} \quad (\text{A6})$$

where

$$\mathbf{R}^{(n)} = \begin{pmatrix} \text{Re}(\alpha^{(n)}) & -\text{Im}(\alpha^{(n)}) \\ \text{Im}(\alpha^{(n)}) & \text{Re}(\alpha^{(n)}) \end{pmatrix}. \quad (\text{A7})$$

Then, from (2) and (A6), we have

$$\mathbf{Cov}(\zeta_{k,j}^{(n)}) = (\mathbf{R}^{(n)})^T \left( \mathbf{Cov}(y_j) - \mathbf{R}^{(n)} p_k \mathbf{Cov}(x_{k,j}^{(n)}) (\mathbf{R}^{(n)})^T \right) \mathbf{R}^{(n)}. \quad (\text{A8})$$

The two diagonal elements of  $\mathbf{Cov}(\zeta_{k,j}^{(n)})$  are  $\text{Var}(\text{Re}(\zeta_{k,j}^{(n)}))$  and  $\text{Var}(\text{Im}(\zeta_{k,j}^{(n)}))$ .

*Computational Cost:* The summations in (A4b) and (A6) are shared by all the layers and antennas. The computational cost per FEC-coded bit involved in the ESE [based on (4)] for the

scheme in Fig. 1(b) increases only linearly with antenna number  $N$  and is independent from the layer number  $K$ , considering  $N$  replicas for each FEC-coded bit. (**Note:** We refer to each  $c_{k,i}$  as an FEC-coded bit.)

*Average Distortion Variance:* We now proceed to calculate the average distortion variances  $\{\text{E}(\text{Var}(\text{Re}(\zeta_{k,j}^{(n)})))\}$  and  $\{\text{E}(\text{Var}(\text{Im}(\zeta_{k,j}^{(n)})))\}$  used in (8). Substituting (A6) into (A8), we have

$$\begin{aligned} \mathbf{Cov}(\zeta_{k,j}^{(n)}) &= (\mathbf{R}^{(n)})^T \left( \sum_{(k', n') \neq (k, n)} \mathbf{R}^{(n')} p_{k'} \right. \\ &\quad \left. \times \mathbf{Cov}(x_{k',j}^{(n')}) (\mathbf{R}^{(n')})^T + \sigma^2 \mathbf{I} \right) \mathbf{R}^{(n)} \\ &= (\mathbf{R}^{(n)})^T \sum_{(k', n') \neq (k, n)} \\ &\quad \left( \mathbf{R}^{(n')} p_{k'} \mathbf{Cov}(x_{k',j}^{(n')}) (\mathbf{R}^{(n')})^T \right) \mathbf{R}^{(n)} + |\alpha^{(n)}|^2 \sigma^2 \mathbf{I}. \end{aligned} \quad (\text{A9})$$

The second equation in (A9) is derived from the fact that  $(\mathbf{R}^{(n)})^T \mathbf{R}^{(n)} = |\alpha^{(n)}|^2 \mathbf{I}$ . Denote

$$v_k \equiv \frac{1}{I} \sum_{i=1}^I \text{Var}(c_{k,i}) \quad (\text{A10a})$$

where  $I$  is the codeword length of the FEC code, and  $\text{Var}(c_{k,i})$  is the variance of  $c_{k,i}$  computed from the output of DEC- $k$  as

$$\text{Var}(c_{k,i}) = 1 - \tanh^2(\text{Ext}(c_{k,i})/2). \quad (\text{A10b})$$

From (A1)–(A3), we can see that the variances of the real and imaginary parts of  $\{x_{k,j}^{(n)}, \forall n, j\}$  are a repeated and interleaved version of the variances of  $\{c_{k,i}, \forall i\}$ , so

$$\text{E}(\mathbf{Cov}(x_{k,j}^{(n)})) = v_k \mathbf{I} \quad (\text{A11})$$

where the expectation is taken over  $j$ . The average of  $\mathbf{Cov}(\zeta_{k,j}^{(n)})$  using (A9) and (A11) is

$$\begin{aligned} \text{E}(\mathbf{Cov}(\zeta_{k,j}^{(n)})) &= (\mathbf{R}^{(n)})^T \sum_{(k', n') \neq (k, n)} \left( \mathbf{R}^{(n')} p_{k'} \text{E}(\mathbf{Cov}(x_{k',j}^{(n')})) (\mathbf{R}^{(n')})^T \right) \mathbf{R}^{(n)} \\ &\quad + |\alpha^{(n)}|^2 \sigma^2 \mathbf{I} \\ &= (\mathbf{R}^{(n)})^T \sum_{(k', n') \neq (k, n)} \left( p_{k'} v_{k'} \mathbf{R}^{(n')} (\mathbf{R}^{(n')})^T \right) \mathbf{R}^{(n)} + |\alpha^{(n)}|^2 \sigma^2 \mathbf{I} \\ &= |\alpha^{(n)}|^2 \sum_{(k', n') \neq (k, n)} \left( p_{k'} v_{k'} |\alpha^{(n')}|^2 \mathbf{I} \right) + |\alpha^{(n)}|^2 \sigma^2 \mathbf{I}. \end{aligned} \quad (\text{A12})$$

Therefore, we have

$$\begin{aligned} \text{E}(\text{Var}(\text{Re}(\zeta_{k,j}^{(n)}))) &= \text{E}(\text{Var}(\text{Im}(\zeta_{k,j}^{(n)}))) \\ &= |\alpha^{(n)}|^2 \left( \sum_{k', n'} p_{k'} |\alpha^{(n')}|^2 v_{k'} + \sigma^2 \right) - |\alpha^{(n)}|^4 p_k v_k. \end{aligned} \quad (\text{A13})$$

As a final remark, we observed that the following updating rule

$$\begin{aligned} \tilde{L}(\text{Re}(x_{k,j}^{(n)})) &= \text{Ext}(c_{k,i}) + \sum_{\substack{(n',j') \in S_R(c_{k,i}) \\ (n',j') \neq (n,j)}} \text{Ext}(\text{Re}(x_{k,j'}^{(n')})) \\ &+ \sum_{(n',j') \in S_I(c_{k,i})} \text{Ext}(\text{Im}(x_{k,j'}^{(n')})) \quad (\text{A14}) \end{aligned}$$

provides slightly better performance than (A3). The last two terms in (A14) represent the information generated by the ESE for other replicas of  $c_{k,i}$  that can be regarded as independent observations of  $y_j$  (assuming sufficiently long random interleavers). However, the analysis of (A14) is a complicated issue and will not be discussed in detail here.

#### APPENDIX B

##### PROOF OF THE INEQUALITY IN (14)

Substituting  $\sum_n |\alpha^{(n)}|^2 = \lambda$  into (10), we have

$$\begin{aligned} \gamma_k &= \sum_{n=1}^N \frac{|\alpha^{(n)}|^2 p_k}{\sum_{k',n'} |\alpha^{(n')}|^2 p_{k'} v_{k'} - |\alpha^{(n)}|^2 p_k v_k + \sigma^2} \\ &= \frac{1}{v_k} \sum_{n=1}^N \frac{|\alpha^{(n)}|^2 p_k}{\lambda \sum_{k'} p_{k'} v_{k'} + \sigma^2} \\ &= \frac{1}{v_k} \sum_{n=1}^N \frac{|\alpha^{(n)}|^2 p_k}{A - |\alpha^{(n)}|^2 p_k} = \frac{1}{v_k} \sum_{n=1}^N \left( \frac{A}{A - |\alpha^{(n)}|^2 p_k} - 1 \right) \quad (\text{B1}) \end{aligned}$$

where  $A \equiv \frac{\lambda \sum_{k'} p_{k'} v_{k'} + \sigma^2}{v_k}$ . Multiplying both sides of (B1) by  $v_k \sum_{n=1}^N (A - |\alpha^{(n)}|^2 p_k)$ , we have (using Schwartz inequality)

$$\begin{aligned} &\gamma_k v_k \sum_{n=1}^N (A - |\alpha^{(n)}|^2 p_k) \\ &= \sum_{n=1}^N \frac{A}{A - |\alpha^{(n)}|^2 p_k} \sum_{n=1}^N (A - |\alpha^{(n)}|^2 p_k) \\ &\quad - N \sum_{n=1}^N (A - |\alpha^{(n)}|^2 p_k) \\ &\geq \left( \sum_{n=1}^N \sqrt{\frac{A}{A - |\alpha^{(n)}|^2 p_k}} \sqrt{A - |\alpha^{(n)}|^2 p_k} \right)^2 - N(NA - \lambda p_k) \\ &= N^2 A - N^2 A + N \lambda p_k = N \lambda p_k \quad (\text{B2}) \end{aligned}$$

where the equality holds when  $|\alpha^{(1)}|^2 = \dots = |\alpha^{(N)}|^2 = \lambda/N$ . Equation (B2) leads to

$$\begin{aligned} \gamma_k &\geq \frac{N \lambda p_k}{v_k \sum_{n=1}^N (A - |\alpha^{(n)}|^2 p_k)} = \frac{N \lambda p_k}{(NA - \lambda p_k) v_k} \\ &= \frac{\lambda p_k}{\lambda \sum_{k'} p_{k'} v_{k'} - \lambda p_k v_k / N + \sigma^2}. \quad (\text{B3}) \end{aligned}$$

#### APPENDIX C

##### POWER ALLOCATION WITH IDEAL BINARY CODING AND SUCCESSIVE CANCELLATION FOR THE IDM-ST CODE

In this appendix, we consider power allocation for the IDM-ST code in Fig. 1(b) with ideal binary FEC coding. We will show that, theoretically, channel capacity can be achieved using ideal low-rate binary FEC codes.

With ideal FEC coding, we employ the successive cancellation technique [40], [41] at the receiver. For each layer- $k$ , assume that the contributions from layers  $k' > k$  have been removed from the received signal, and treat the interference from layers  $k' < k$  in the same way as the additive noise. This is reasonable due to the random interleaving used. We now show how to determine  $p_k$  so that signals of layer- $k$  can be successively decoded and its interference to lower layers can be removed. We initialize  $v_k = 1$  for  $\forall k$  at the first iteration. (As we will see, only one iteration is required with ideal coding [40], [41].) Decoding starts from the top layer. During the decoding for layer- $k$ , we have  $v_{k'} = 0$  for  $k' > k$  and  $v_{k'} = 1$  for  $k' < k$ . Then, we have from (14) (**Note:** In this case,  $\gamma_k = \text{snr}_k$  for  $\forall k$  because  $\text{Var}(x_{k,j}^{(n)}) = \text{Var}(c_{k,i}) = v_k$  for  $\forall k, n, i, j$ .)

$$\gamma_k \geq \frac{\lambda p_k}{\lambda \sum_{k' < k} p_{k'} + \lambda(N-1)p_k/N + \sigma^2}. \quad (\text{C1})$$

Consider a real AWGN channel with binary inputs, and denote by  $\rho_2$  the minimum SNR required for reliable communication at rate  $R_C$ . Let  $A \equiv \sqrt{\rho_2}$ . It can be shown that [42]

$$R_C = \left( - \int_{-\infty}^{+\infty} p(y) \ln(p(y)) dy - \frac{1}{2} \ln(2\pi e) \right) / \ln(2) \quad (\text{C2a})$$

where

$$\begin{aligned} p(y) &= \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y-A)^2}{2}\right) \\ &\quad + \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y+A)^2}{2}\right). \quad (\text{C2b}) \end{aligned}$$

First, we assume that  $\lambda$  is known *a priori* at the transmitter. Recall that the FEC code is a binary code, and the input to each DEC- $k$  is treated approximately as the LLRs of the coded bits  $\{c_{k,i}, \forall i\}$  obtained from the outputs of a real AWGN channel with SNR given in (C1) (see Section IV-A). Assume that  $\{p_1, \dots, p_{k-1}\}$  have been determined. We choose  $p_k$  such that

$$\frac{\lambda p_k}{\lambda \sum_{k' < k} p_{k'} + \lambda(N-1)p_k/N + \sigma^2} = \rho_2. \quad (\text{C3})$$

If the FEC code is an ideal binary code, we can successfully decode layer- $k$ . We then proceed to determine  $p_{k+1}$  for layer- $(k+1)$  and so on. Starting from  $\frac{\lambda p_1}{\lambda(N-1)p_1/N + \sigma^2} = \rho_2$ , we can generate  $p_1, \dots, p_K$  recursively according to (C3). The total power required can be calculated as

$$P_{\text{IDM}} = \sum_k p_k = \frac{\sigma^2}{\lambda} \left( \left( \frac{\rho_2 + N}{N - \rho_2(N-1)} \right)^{\frac{R}{R_C}} - 1 \right) \quad (\text{C4})$$

where  $R = 2KR_C$  is the total rate.

Recall that the capacity of an  $N \times 1$  system for any given  $\alpha$  is [1], [2]

$$CAP = \log_2(1 + \alpha \alpha^H \rho / N) = \log_2(1 + \lambda P / \sigma^2) \text{ bits/symbol} \quad (C5)$$

where

$$\rho = NP / \sigma^2 \quad (C6)$$

is the average SNR per receive antenna. For any given transmission rate  $R$ , the minimum required value of  $P$  for reliable transmission, denoted by  $P_{\min}$ , can be calculated from (C5) by setting  $CAP = R$  as

$$P_{\min} = (2^R - 1) \sigma^2 / \lambda. \quad (C7)$$

In general, there is a gap between  $P_{\text{IDM}}$  and  $P_{\min}$  even when the FEC code is ideal. However, the following theorem shows that using a very low-rate binary FEC code can eliminate this gap. Hence the IDM-ST code in Fig. 1(b) can indeed achieve capacity provided that an ideal low-rate binary FEC code is used.

*Theorem:*

$$\lim_{R_C \rightarrow 0} (P_{\text{IDM}} - P_{\min}) = 0. \quad (C8)$$

*Proof:* From (C4) and (C7), we have

$$P_{\text{IDM}} - P_{\min} = \frac{\sigma^2}{\lambda} \left( \left( \frac{\rho_2 + N}{N - \rho_2(N-1)} \right)^{\frac{R}{2R_C}} - 2^R \right).$$

Thus, the theorem holds if

$$\lim_{R_C \rightarrow 0} \frac{R}{2R_C} \log_2 \left( \frac{\rho_2 + N}{N - \rho_2(N-1)} \right) = R. \quad (C9)$$

From (C2), it is easy to see that when  $A \rightarrow 0$ ,  $R_C \rightarrow 0$ . The left-hand side of (C9) can be rewritten as (C10), shown at the bottom of the page. (**Note:**  $R_C$  is a function of  $A$ , as shown in (C2). The second equation in (C10) is obtained using L'Hospital's rule.) From (C2a), we have

$$\frac{d^2 R_C}{dA} = -\frac{1}{\ln(2)} \int_{-\infty}^{+\infty} \left( p''(y) (1 + \ln(p(y))) + \frac{(p'(y))^2}{p(y)} \right) dy \quad (C11)$$

where [from (C2b)]

$$p'(y) = \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y-A)^2}{2}\right) \cdot (y-A) - \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y+A)^2}{2}\right) \cdot (y+A) \quad (C12a)$$

and

$$p''(y) = \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}} \left( \exp\left(-\frac{(y-A)^2}{2}\right) \cdot (y-A)^2 - \exp\left(-\frac{(y-A)^2}{2}\right) + \exp\left(-\frac{(y+A)^2}{2}\right) \cdot (y+A)^2 - \exp\left(-\frac{(y+A)^2}{2}\right) \right). \quad (C12b)$$

Substituting (C11) and (C12) into (C10) results in (C9), i.e., (C8) holds.  $\square$

Next, we consider the situation when  $\lambda$  is a random variable unknown at the transmitter but the distribution of  $\lambda$  is known at the transmitter. In this case, we use the so-called outage capacity  $\beta_{\text{out}}$  [1], [2], [35] as the performance criterion, where

$$\beta_{\text{out}} \equiv \Pr(CAP < R). \quad (C13)$$

Given  $\beta_{\text{out}}$  and  $R$ , we find  $\lambda_R$  using the distribution of  $\lambda$  such that

$$\Pr(\lambda < \lambda_R) = \beta_{\text{out}}. \quad (C14)$$

Theoretically, the minimum value of  $P$  to achieve any given  $\beta_{\text{out}}$  at rate  $R$ , again denoted by  $P_{\min}$ , can be found by solving the following capacity equation [obtained by substituting (C5) and (C14) into (C13)]:

$$R = \log_2(1 + \lambda_R P_{\min} / \sigma^2). \quad (C15)$$

Now, suppose that we allocate power for the IDM-ST code by setting  $\lambda = \lambda_R$  in (C3) and obtain  $P_{\text{IDM}}$  using (C4). When the FEC code is an ideal binary code, decoding fails only when  $\lambda < \lambda_R$  so the decoding failure probability (i.e., FER) is given

$$\begin{aligned} \lim_{R_C \rightarrow 0} \frac{R}{2R_C} \log_2 \left( \frac{A^2 + N}{N - A^2(N-1)} \right) &= \lim_{A \rightarrow 0} \frac{\log_2 \left( \frac{A^2 + N}{N - A^2(N-1)} \right)}{2R_C/R} \\ &= \lim_{A \rightarrow 0} \frac{\left( -\frac{4A^2}{(A^2 + N)^2} + \frac{2}{A^2 + N} + \frac{4A^2(N-1)^2}{(N - A^2(N-1))^2} + \frac{2(N-1)}{N - A^2(N-1)} \right) / \ln(2)}{\frac{2}{R} \cdot \frac{d^2 R_C}{dA}} \\ &= \frac{1}{\ln(2)} \lim_{A \rightarrow 0} \frac{R}{\frac{d^2 R_C}{dA}}. \end{aligned} \quad (C10)$$

by  $\beta_{\text{out}}$ . Because (C8) holds, i.e.,  $P_{\text{IDM}} \rightarrow P_{\text{min}}$ , when  $R_C \rightarrow 0$ , the IDM-ST code in Fig. 1(b) can approach the theoretical limit in quasi-static fading channels with ideal low-rate binary FEC coding.

## REFERENCES

- [1] G. J. Foschini and M. Gans, "On the limits of wireless communication in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 311–335, Mar. 1998.
- [2] E. Telatar, "Capacity of multi-antenna Gaussian channels," *AT&T-Bell Labs Int. Tech. Memo* Jun. 1995 [Online]. Available: mars.bell-labs.com
- [3] A. R. Hammons Jr. and H. El Gamal, "On the theory of space-time codes for PSK modulation," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 524–542, Mar. 2000.
- [4] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance analysis and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.
- [5] A. F. Naguib, V. Tarokh, N. Seshadri, and A. R. Calderbank, "A space-time coding modem for high-data-rate wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1459–1478, Oct. 1998.
- [6] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.
- [7] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.
- [8] G. D. Golden, C. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Inst. Electr. Eng. Electron. Lett.*, vol. 35, no. 1, pp. 14–16, Jan. 1999.
- [9] G. J. Foschini, "Layered space-time architecture for wireless communication in fading environment when using multiple antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.
- [10] M. Sellathurai and S. Haykin, "Turbo-BLAST for wireless communications: Theory and experiments," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2538–2546, Oct. 2002.
- [11] H. El Gamal and A. R. Hammons, "A new approach to layered space-time coding and signal processing," *IEEE Trans. Inf. Theory*, vol. 47, no. 9, pp. 2321–2334, Sep. 2001.
- [12] B. Hassibi and B. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1804–1824, Jul. 2002.
- [13] R. W. Heath and A. J. Paulraj, "Linear dispersion codes for MIMO systems based on frame theory," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2429–2441, Oct. 2002.
- [14] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2415–2424, Oct. 2003.
- [15] K. Y. Wu and L. Ping, "Multilayer turbo space-time codes," *IEEE Commun. Lett.*, vol. 9, no. 1, pp. 55–57, Jan. 2005.
- [16] C. Berrou and A. Glavieux, "Near Shannon limit error correcting coding and decoding: Turbo-codes," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261–1271, Oct. 1996.
- [17] T. Richardson and R. Urbanke, "The capacity of low density parity check codes under message passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001.
- [18] D. Divsalar, S. Dolinar, and F. Pollara, "Iterative turbo decoder analysis based on density evolution," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 5, pp. 891–907, May 2001.
- [19] G. Caire, R. Muller, and T. Tanaka, "Iterative multiuser joint decoding: Optimal power allocation and low-complexity implementation," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1950–1973, Sep. 2004.
- [20] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.
- [21] M. V. Burnashev, C. B. Schlegel, W. A. Krzymien, and Z. Shi, "Analysis of the dynamics of iterative interference cancellation in iterative decoding," *Probl. Inf. Trans.*, vol. 40, no. 4, pp. 297–317, Oct. 2004.
- [22] L. Liu, J. Tong, and L. Ping, "Analysis and optimization of CDMA systems with chip-level interleavers," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 1, pp. 141–150, Jan. 2006.
- [23] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1046–1061, Jul. 1999.
- [24] J. Luo, K. R. Pattipati, P. K. Willett, and F. Hasegawa, "Near-optimal multiuser detection in synchronous CDMA using probabilistic data association," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 361–363, Sep. 2001.
- [25] P. H. Tan and L. K. Rasmussen, "Asymptotically optimal nonlinear MMSE multiuser detection based on multivariate Gaussian approximation," *IEEE Trans. Commun.*, vol. 54, no. 8, pp. 1427–1438, Aug. 2006.
- [26] J. Fricke, M. Sandell, J. Mietzner, and P. Hoehner, "Impact of the Gaussian approximation on the performance of probabilistic data association MIMO detector," *EURASIP J. Wireless Commun.*, pp. 796–800, May 2005.
- [27] T. M. Cover and J. M. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [28] L. Ping, L. Liu, K. Y. Wu, and W. K. Leung, "Interleave division multiple-access," *IEEE Trans. Wireless Commun.*, vol. 5, no. 4, pp. 938–947, Apr. 2006.
- [29] M. Moher, "An iterative multiuser decoder for near-capacity communications," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 870–880, Jul. 1998.
- [30] F. N. Brannstrom, T. M. Aulin, and L. K. Rasmussen, "Iterative decoders for trellis code multiple-access," *IEEE Trans. Commun.*, vol. 50, no. 9, pp. 1478–1485, Sep. 2002.
- [31] S. Brück, U. Sorger, S. Gligorevic, and N. Stolte, "Interleaving for outer convolutional codes in DS-SS systems," *IEEE Trans. Commun.*, vol. 48, no. 7, pp. 1100–1107, Jul. 2000.
- [32] T. S. Rappaport, *Wireless Communications Principle and Practice*. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [33] L. H. C. Janson, M. Tao, and R. S. Cheng, "Optimal power allocation scheme on generalized layer space-time coding systems," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2001, pp. 1706–1710.
- [34] I. M. Kim and V. Tarokh, "Variable-rate space-time block codes in M-ary PSK systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 4, pp. 362–373, Apr. 2003.
- [35] D. Cui and A. M. Haimovich, "Performance of parallel concatenated space-time codes," *IEEE Commun. Lett.*, vol. 5, no. 6, pp. 236–238, Jun. 2001.
- [36] L. Ping, W. K. Leung, and K. Y. Wu, "Low-rate turbo-Hadamard codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3213–3224, Dec. 2003.
- [37] L. Liu and L. Ping, "An extending window MMSE turbo equalization algorithm," *IEEE Signal Process. Lett.*, vol. 11, no. 11, pp. 891–894, Nov. 2004.
- [38] Y. Wu, X. Zhu, and A. K. Nandi, "Low complexity turbo space-frequency equalization for single-carrier MIMO wireless communications," in *Proc. EUSIPCO*, Sep. 2006 [Online]. Available: www.eurasip.org
- [39] X. J. Yuan, Q. H. Guo, and L. Ping, "Low-complexity iterative detection in multiuser MIMO ISI channels," *IEEE Signal Process. Lett.*, vol. 15, pp. 25–28, 2008.
- [40] N. Chatay and S. Shamai, "Convergence properties of iterative soft onion peeling," in *Proc. Inf. Theory Workshop*, Kruger National Park, South Africa, Jun. 1999, pp. 9–11.
- [41] G. Caire, S. Guemghar, A. Roumy, and S. Verdu, "Maximizing the spectral efficiency of coded CDMA under successive decoding," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 152–164, Jan. 2004.
- [42] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.