

# Evolution Analysis of Low-Cost Iterative Equalization in Coded Linear Systems with Cyclic Prefixes

Xiaojun Yuan, *Student Member, IEEE*, Qinghua Guo, *Student Member, IEEE*, Xiaodong Wang, *Fellow, IEEE*, and Li Ping, *Senior Member, IEEE*

**Abstract**—This paper is concerned with the low-cost iterative equalization/detection principles for coded linear systems with cyclic prefixes. Turbo frequency-domain-equalization (FDE) is applied to systems that may contain the joint effect of multiple-access interference (MAI), cross-antenna interference (CAI) and inter-symbol interference (ISI). We develop an SNR-variance evolution technique for the performance evaluation of the proposed systems. Numerical results in various channel environments demonstrate excellent agreement between the predicted and simulated system performance.

**Index Terms**—cyclic prefix, multiuser MIMO ISI channel, Turbo FDE, SNR-variance evolution.

## I. INTRODUCTION

TURBO equalization/detection systems consist of two local operators, namely a coarse channel equalizer/detector for handling the channel effects (such as the inter-symbol interference (ISI)) and a decoder for *a posteriori* probability (APP) decoding. Such systems can achieve impressive performance gain by iteratively refining the estimates interchanged between the two operators. The coarse channel equalizer/detector can be realized by the optimal maximum *a posteriori* (MAP) algorithm [1-3] at excessively high complexity, or by the linear minimum-mean-square-error (LMMSE) technique [4-9] that provides an attractive trade-off between cost and performance. Recently, it has been shown that LMMSE based frequency domain equalization (FDE) [10-15] allows further complexity reduction without noticeably compromising performance. The FDE technique exploits the property of circulant channel matrices that can be ensured by inserting cyclic-prefixes (CP) into the transmitted signals [11][16].

In this paper, we examine a generic system framework that can cover the joint effect of multiple-access interference (MAI), cross-antenna interference (CAI) and ISI. Based on

this generic framework, we derive an efficient joint turbo equalization multiuser-detection algorithm for highly complex multiuser, multiple-input-multiple-output (MIMO), ISI channels. Our main focus is a semi-analytical method for performance evaluation of the proposed algorithm. The extrinsic information transfer (EXIT) chart technique [17] is a useful tool for the analysis of a turbo or low density parity check (LDPC) decoder involving the iteration of two local processors. The EXIT functions are generated by simulating the two local processors individually, which is a much simpler task than simulating them jointly. The former involves non-iterative processing but the latter involves iterative processing. The EXIT transfer functions are based on mutual information [17], but other variables such as signal-to-noise-ratio (SNR) or variance [18][19] can also be used alternatively. The EXIT chart method can be directly applied to turbo equalization systems [6] for fixed channel conditions. However, the situation becomes more complicated when the channel is not fixed, but characterized by a fading distribution. In this case, the average performance of the system can be computed by collecting statistics for a sufficiently large number of channel realizations. The transfer characteristic of the channel equalizer/detector is the main difficulty here, since it depends heavily on the channel realization. For each different realization, an individual pre-simulated EXIT transfer function is required, but it is very costly to cover all possibilities. In this paper, we develop an alternative solution in which the transfer function of the equalizer/detector is generated analytically online (rather than pre-simulated) at a very low cost for each channel realization during the evolution process. We show that the proposed SNR-variance evolution technique is simple, fast and accurate. We provide numerical results to demonstrate the effectiveness of the proposed method in various ISI, MIMO and multiuser environments.

## II. SYSTEM MODEL AND ITERATIVE EQUALIZATION PRINCIPLES

### A. Linear System Model

Consider a generic system (Fig. 1) that can be modeled by a linear equation

$$\mathbf{r} = \mathbf{H}\mathbf{x} + \boldsymbol{\eta} \quad (1)$$

where  $\mathbf{r}$  is an observation vector,  $\mathbf{H}$  a channel transfer matrix,  $\mathbf{x}$  a transmitted signal vector,  $\boldsymbol{\eta}$  an additive white Gaussian

Manuscript received April 1, 2007; revised October 30, 2007. This work was fully supported by a grant from the Research Grant Council of the Hong Kong SAR, China [Project No. CityU 1314/04E]. The material in this paper is presented in part at the International Symposium on Information Theory, Nice, France, 2007.

Xiaojun Yuan, Qinghua Guo and Li Ping are with the Department of Electrical Engineering, City University of Hong Kong, HK SAR (e-mail: xjyuan@cityu.edu.hk, 50007550@student.cityu.edu.hk, eeliping@cityu.edu.hk).

Xiaodong Wang is with the Department of Electrical Engineering, Columbia University, New York, NY 10027, USA (e-mail: wangx@ee.columbia.edu).

Digital Object Identifier 10.1109/JSAC.2008.080207.

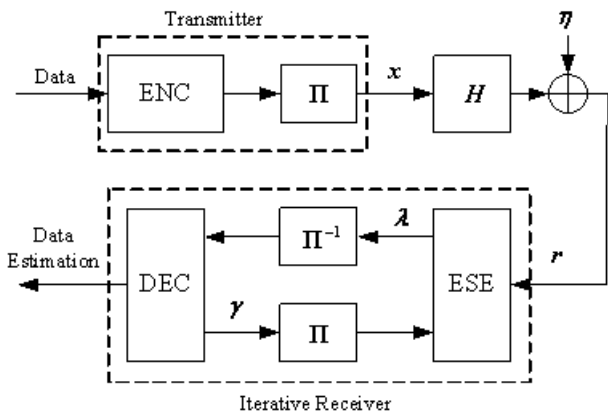


Fig. 1. The generic transmitter and (iterative) receiver structures for a coded linear system with channel input  $\mathbf{x}$ , and with  $\Pi$  and  $\Pi^{-1}$  denoting the interleaver and the corresponding de-interleaver, respectively.

noise (AWGN) vector with zero mean and covariance  $\sigma^2 \mathbf{I}$ , and  $\mathbf{I}$  an identity matrix with proper size. We assume that  $\mathbf{x}$  is generated by a generic forward-error-correcting (FEC) encoding device (labeled by ENC in Fig. 1) and permuted by a generic interleaver (labeled by  $\Pi$  in Fig. 1). Typical examples of (1) include systems involving MAI, CAI, ISI, or their combinations. The ENC may contain a single encoder (for a single user system) or multiple encoders (for a multiuser system). In the following, we will first concentrate on real systems with binary phase shift keying (BPSK) modulation over  $\{+1, -1\}$ . Later, we will extend our results to complex systems with quadrature phase shift keying (QPSK) modulation.

### B. The Overall Iterative Equalization Principles

The iterative receiver is composed of an elementary signal estimator (ESE) and a generic decoding device (DEC) separated by the interleaver  $\Pi$  and the deinterleaver  $\Pi^{-1}$ , as is shown in the lower part of Fig. 1. The ESE computes the extrinsic log-likelihood ratio (LLR) for each  $x_j$ , the  $j$ th entry of  $\mathbf{x}$ , as

$$\lambda_j \equiv \ln \frac{p(\mathbf{r}|x_j = +1)}{p(\mathbf{r}|x_j = -1)}, \quad j = 0, 1, 2, \dots \quad (2)$$

with the FEC coding constraint ignored, i.e., the ESE operates as if  $\mathbf{x}$  contains un-coded bits.

The DEC performs APP decoding using  $\boldsymbol{\lambda} \equiv [\lambda_0, \lambda_1, \dots, \lambda_j, \dots]^T$  as the input. The DEC outputs are the extrinsic LLRs given by

$$\gamma_j \equiv \ln \frac{p(\boldsymbol{\lambda}|x_j = +1)}{p(\boldsymbol{\lambda}|x_j = -1)}, \quad j = 0, 1, 2, \dots$$

Afterwards, the ESE operations can be executed again to refine the estimates in (2) using the feedback  $\boldsymbol{\gamma} \equiv [\gamma_0, \gamma_1, \dots, \gamma_j, \dots]^T$  from the DEC. Detailed discussions on this iterative process can also be found in [4-8]. Since the DEC is well-studied, we will focus on the ESE from now on.

### C. The LMMSE Approach to the ESE

Let  $E(\cdot)$  denote the expectation operation, and the superscript “ $T$ ” denote the transpose operation. Define  $\text{Cov}(\mathbf{x}, \mathbf{y}) \equiv$

$E((\mathbf{x} - E(\mathbf{x}))(\mathbf{y} - E(\mathbf{y}))^T)$ . Denote by  $E(x_j)$  and  $v_j \equiv \text{Cov}(x_j, x_j)$  the *a priori* mean and variance of  $x_j$ , respectively; and denote by  $E(\mathbf{x})$  and  $\mathbf{V} \equiv \text{Cov}(\mathbf{x}, \mathbf{x})$  the *a priori* mean and auto-covariance matrix of  $\mathbf{x}$ , respectively. Note that  $E(x_j)$  is simply the  $j$ th entry of  $E(\mathbf{x})$ , and  $v_j$  is the  $j$ th diagonal entry of  $\mathbf{V}$ .  $E(x_j)$  and  $v_j$ ,  $\forall j$  (initialized to 0 and 1, respectively, indicating that no *a priori* information is available) are computed by the feedback information from the DEC, i.e., for each  $x_j$  taken over  $\{+1, -1\}$ ,

$$E(x_j) = \frac{\exp(\gamma_j) - 1}{\exp(\gamma_j) + 1} \quad \text{and} \quad v_j = 1 - (E(x_j))^2.$$

Assuming that  $\mathbf{H}$  is known at the receiver, we focus on the estimation of a particular bit  $x_j$  and treat all the others as interference. From (1), the covariance matrix of  $\mathbf{r}$  is

$$\mathbf{R} \equiv \text{Cov}(\mathbf{r}, \mathbf{r}) = \mathbf{H}\mathbf{V}\mathbf{H}^T + \sigma^2 \mathbf{I}. \quad (3)$$

We can rewrite (1) as

$$\mathbf{r} = \mathbf{h}_j x_j + \boldsymbol{\xi}_j$$

where  $\mathbf{h}_j$  is the  $j$ th column of  $\mathbf{H}$ , and  $\boldsymbol{\xi}_j \equiv \sum_{j' \neq j} \mathbf{h}_{j'} x_{j'} + \boldsymbol{\eta}$ . It can be shown that

$$\mathbf{R}_j \equiv \text{Cov}(\boldsymbol{\xi}_j, \boldsymbol{\xi}_j) = \mathbf{R} - v_j \mathbf{h}_j \mathbf{h}_j^T.$$

We approximate  $\boldsymbol{\xi}_j$  by a jointly Gaussian random vector. This assumption can be justified by the central limit theorem, and it leads to

$$\begin{aligned} \lambda_j &= \ln \frac{\exp(-\frac{1}{2}(\mathbf{r} - \mathbf{h}_j - E(\boldsymbol{\xi}_j))^T \mathbf{R}_j^{-1} (\mathbf{r} - \mathbf{h}_j - E(\boldsymbol{\xi}_j)))}{\exp(-\frac{1}{2}(\mathbf{r} + \mathbf{h}_j - E(\boldsymbol{\xi}_j))^T \mathbf{R}_j^{-1} (\mathbf{r} + \mathbf{h}_j - E(\boldsymbol{\xi}_j)))} \\ &= 2\mathbf{h}_j^T \mathbf{R}_j^{-1} (\mathbf{r} - \mathbf{H}E(\mathbf{x}) + \mathbf{h}_j E(x_j)). \end{aligned} \quad (4a)$$

Applying the matrix inversion lemma to  $\mathbf{R}_j^{-1}$ , we can rewrite  $\lambda_j$  as

$$\lambda_j = 2 \cdot \frac{\mathbf{h}_j^T \mathbf{R}^{-1} (\mathbf{r} - \mathbf{H}E(\mathbf{x})) + \mathbf{h}_j^T \mathbf{R}^{-1} \mathbf{h}_j E(x_j)}{1 - v_j \mathbf{h}_j^T \mathbf{R}^{-1} \mathbf{h}_j} \quad (4b)$$

or in a vector form as

$$\boldsymbol{\lambda} = 2(\mathbf{I} - \mathbf{V}\mathbf{U})^{-1} (\mathbf{H}^T \mathbf{R}^{-1} (\mathbf{r} - \mathbf{H}E(\mathbf{x})) + \mathbf{U}E(\mathbf{x})) \quad (4c)$$

where

$$\mathbf{U} \equiv (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})_{\text{diag}} \quad (4d)$$

is a diagonal matrix consisting of the diagonal entries of  $\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}$ .

The above is in fact equivalent to (though far more concise than) the LMMSE method originally proposed in [4]. The proof of the equivalence can be found in [9]. For this reason, we will call  $\boldsymbol{\lambda}$  an LMMSE estimator hereafter.

### D. SNR Analysis

We rewrite (4a) in a signal-plus-distortion form as

$$\lambda_j = \mu_j x_j + \zeta_j \quad (5a)$$

where

$$\mu_j \equiv 2\mathbf{h}_j^T \mathbf{R}_j^{-1} \mathbf{h}_j \quad (5b)$$

and

$$\zeta_j \equiv 2\mathbf{h}_j^T \mathbf{R}_j^{-1} (\mathbf{r} - \mathbf{H}\mathbf{E}(\mathbf{x}) - \mathbf{h}_j(x_j - \mathbf{E}(x_j))). \quad (5c)$$

Here  $\zeta_j$  is the residual distortion (including noise and interference from  $\{x_{j'}, \forall j' \neq j\}$ ) with respect to  $x_j$  after the ESE operation. It can be shown that  $\mathbf{E}(\zeta_j) = 0$  and  $\mathbf{E}(\zeta_j^2) = 4\mathbf{h}_j^T \mathbf{R}_j^{-1} \mathbf{h}_j$ . Define  $\rho_j \equiv \mu_j^2 / \mathbf{E}(\zeta_j^2)$  that represents the SNR for  $\lambda_j$ . Thus

$$\rho_j = \mathbf{h}_j^T \mathbf{R}_j^{-1} \mathbf{h}_j = \frac{\mathbf{h}_j^T \mathbf{R}_j^{-1} \mathbf{h}_j}{1 - v_j \mathbf{h}_j^T \mathbf{R}_j^{-1} \mathbf{h}_j} \quad (6a)$$

where the last equality of (6a) follows from the matrix inversion lemma. Therefore

$$\begin{aligned} \text{diag}\{\rho_0, \rho_1, \dots\} &= \frac{1}{2} \text{diag}\{\mu_0, \mu_1, \dots\} \\ &= \frac{1}{4} \text{diag}\{\mathbf{E}(\zeta_0^2), \mathbf{E}(\zeta_1^2), \dots\} \\ &= \mathbf{U}(\mathbf{I} - \mathbf{V}\mathbf{U})^{-1} \end{aligned} \quad (6b)$$

where  $\text{diag}\{\cdot\}$  represents a diagonal matrix using the entries in the braces. Interestingly, the so-called symmetric condition  $2\mu_j = \mathbf{E}(\zeta_j^2)$  [25] can be seen from (6b). We will return to (6) later in Section IV.

### III. FAST IMPLEMENTATION TECHNIQUES

#### A. Circulant Systems

A square matrix  $\mathbf{H}$  is circulant if its  $j$ th row (or column),  $\forall j$ , is the cyclic shift of its first row (or column) by  $j$  positions. For example, a circulant  $\mathbf{H}$  can be represented as

$$\mathbf{H} = \begin{pmatrix} h_0 & \cdots & h_2 & h_1 \\ h_1 & h_0 & & \\ \vdots & \ddots & \ddots & \vdots \\ h_{J-1} & \cdots & h_1 & h_0 \end{pmatrix} \quad (7)$$

where  $J$  is the size of  $\mathbf{H}$ . Let  $\mathbf{h}$  be the first column of  $\mathbf{H}$ . Then the circular convolution between  $\mathbf{h}$  and  $\mathbf{x}$  can be expressed in a vector form as  $\mathbf{H}\mathbf{x}$  (see Chapter 3 in [20]). The matrix  $\mathbf{H}$  in (7) can be realized in a quasi-static ISI channel (with  $h_l$  denoting the  $l$ th channel coefficient) by padding  $\mathbf{x}$  with a cyclic prefix (CP) that consists of the last  $L - 1$  entries of  $\mathbf{x}$  [11][20], where  $L$  is the channel memory length.

Let  $\mathbf{F}$  be the normalized discrete Fourier transform (DFT) matrix with the  $(j, l)$ -th entry given by  $J^{-\frac{1}{2}} \exp(-i2\pi jl/J)$ , where  $i = \sqrt{-1}$ . Thus,  $\mathbf{F}\mathbf{F}^H = \mathbf{F}^H\mathbf{F} = \mathbf{I}$ , where the superscript “ $H$ ” denotes the conjugate transpose operation. The following properties are well-known for a circulant  $\mathbf{H}$ .

**Property I:** If a matrix  $\mathbf{H}$  is circulant, then  $\mathbf{G} \equiv \mathbf{F}\mathbf{H}\mathbf{F}^H$  is diagonal, i.e.,  $\mathbf{G} = J^{\frac{1}{2}} \text{diag}\{g_0, g_1, \dots, g_{J-1}\}$ ; and vice versa.

**Property II:** In the above,  $\{h_l\}$  (see (7)) and  $\{g_j\}$  form a DFT pair, i.e.,

$$\begin{aligned} g_j &= J^{-\frac{1}{2}} \sum_{l=0}^{J-1} h_l \exp(-i2\pi jl/J) \\ h_l &= J^{-\frac{1}{2}} \sum_{j=0}^{J-1} g_j \exp(i2\pi jl/J). \end{aligned}$$

#### B. The ESE Operations for Circulant $\mathbf{H}$

The ESE output  $\lambda$  in (4c) can be evaluated efficiently when  $\mathbf{H}$  is circulant. Let  $\bar{v}$  be the average of the feedback variances  $\{v_j\}$  from the DEC, i.e.,

$$\bar{v} \equiv \frac{1}{J} \sum_{j=0}^{J-1} v_j \quad (8a)$$

and approximate the *a priori* covariance matrix  $\mathbf{V}$  by  $\bar{v}\mathbf{I}$ , i.e.,

$$\mathbf{V} \approx \bar{v}\mathbf{I}. \quad (8b)$$

This approximation was first implicitly used in [21], and later adopted by many other authors [6][10][14]. Eqn. (8b) implies the following assumption.

**Assumption I:** (i)  $x_j$  and  $x_{j'}$  are uncorrelated if  $j \neq j'$ , as is approximately ensured by the interleaver  $\mathbf{\Pi}$ ; and (ii) all  $\{x_j\}$  have the same variance  $\bar{v}$ .

In above, (ii) is not an optimal treatment since we actually know the individual variances  $\{v_j\}$ . However, according to our observation on numerical results (as provided later), this sub-optimal treatment can lead to reduced complexity with only marginal performance loss in most channel conditions. For convenience and without confusion, we hereafter replace the approximation in (8b) with the equation:

$$\mathbf{V} = \bar{v}\mathbf{I}. \quad (9)$$

From (3) and (9),  $\mathbf{R}$  is also circulant if  $\mathbf{H}$  is circulant. Thus,  $\mathbf{R}$  can be diagonalized by the DFT matrices, i.e., by invoking Property I

$$\mathbf{F}\mathbf{R}\mathbf{F}^H = \mathbf{F}(\bar{v}\mathbf{H}\mathbf{H}^T + \sigma^2\mathbf{I})\mathbf{F}^H = \bar{v}\mathbf{G}\mathbf{G}^H + \sigma^2\mathbf{I}. \quad (10a)$$

Hence we have

$$\mathbf{H}^T \mathbf{R}^{-1} = \mathbf{F}^H \mathbf{G}^H (\bar{v}\mathbf{G}\mathbf{G}^H + \sigma^2\mathbf{I})^{-1} \mathbf{F} \quad (10b)$$

and

$$\mathbf{U} \equiv (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})_{\text{diag}} = u\mathbf{I} \quad (10c)$$

where

$$u \equiv \sum_{j=0}^{J-1} |g_j|^2 (\bar{v}J|g_j|^2 + \sigma^2)^{-1}. \quad (10d)$$

Then (4c) can be rewritten as

$$\begin{aligned} \lambda &= 2(\mathbf{I} - \mathbf{V}\mathbf{U})^{-1} (\mathbf{F}^H \mathbf{G}^H (\mathbf{G}\mathbf{V}\mathbf{G}^H + \sigma^2\mathbf{I})^{-1} \\ &\quad \cdot (\mathbf{F}\mathbf{r} - \mathbf{G}\mathbf{F}\mathbf{E}(\mathbf{x})) + \mathbf{U}\mathbf{E}(\mathbf{x})). \end{aligned} \quad (11)$$

This leads to the following efficient implementation technique for the ESE.

#### Algorithm I:

Step 1: Compute  $\mathbf{F}\mathbf{r} - \mathbf{G}\mathbf{F}\mathbf{E}(\mathbf{x})$ .

Step 2: Left-multiply the result of Step 1 by a diagonal matrix  $\mathbf{G}^H (\mathbf{G}\mathbf{V}\mathbf{G}^H + \sigma^2\mathbf{I})^{-1}$ .

Step 3: Left-multiply the result of Step 2 by  $\mathbf{F}^H$ , add  $\mathbf{U}\mathbf{E}(\mathbf{x})$ , and scale the result by  $2(\mathbf{I} - \mathbf{V}\mathbf{U})^{-1}$ .

In Steps 1 and 3, multiplying by  $\mathbf{F}$  and  $\mathbf{F}^H$  can be realized by the fast Fourier transform (FFT) and the inverse FFT (IFFT)

algorithms, respectively. Define the normalized cost as the number of operations per entry of  $\mathbf{x}$ , with each operation equal to one multiplication and one addition. Then the normalized cost of Algorithm I is roughly  $O(\log_2 J)$ . It can be shown that for a simple ISI channel, (11) is equivalent to the FDE algorithm first derived in [10]. However, the generic form of (11) allows extensions to a wider range of applications such as multiuser MIMO ISI channels, as will be shown next.

### C. Block-Circulant Systems

As a natural extension of the scalar circulant system in Section III.A, a linear system modeled by (1) can be described as "block-circulant" if it can be expressed in a block form as

$$\begin{pmatrix} \mathbf{r}_0 \\ \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_{J-1} \end{pmatrix} = \begin{pmatrix} \mathbf{H}_0 & \cdots & \mathbf{H}_2 & \mathbf{H}_1 \\ \mathbf{H}_1 & \mathbf{H}_0 & & \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{H}_{J-1} & \cdots & \mathbf{H}_1 & \mathbf{H}_0 \end{pmatrix} \begin{pmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_{J-1} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\eta}_0 \\ \boldsymbol{\eta}_1 \\ \vdots \\ \boldsymbol{\eta}_{J-1} \end{pmatrix}. \quad (12)$$

In above, the channel matrix  $\mathbf{H}$  is block-circulant as each block-column is a cyclic shift of its previous block-column by one block position; each  $\mathbf{H}_l$  is an  $M \times N$  sub-matrix; each  $\mathbf{r}_j$  or  $\boldsymbol{\eta}_j$  is an  $M \times 1$  sub-vector; each  $\mathbf{x}_j$  is an  $N \times 1$  sub-vector; and  $J$  specifies the size of  $\mathbf{H}$  in blocks. A typical example of such a system arises in a MIMO ISI channel with CP padding to each transmit antenna. Let  $N$  and  $M$  be, respectively, the number of transmit and receive antennas.  $\mathbf{H}_l$  is defined as  $[\mathbf{H}_l]_{mn} = h_l^{(m,n)}$ , for  $m = 1, \dots, M$ , and  $n = 1, \dots, N$ , where  $h_l^{(m,n)}$  is the  $l$ th tap coefficient between the  $n$ th transmit antenna and the  $m$ th receive antenna, and  $[\cdot]_{l,j}$  denotes the  $(l, j)$ -th entry of the matrix in the brackets. In (12),  $\mathbf{r}_j$ ,  $\mathbf{x}_j$  and  $\boldsymbol{\eta}_j$  are respectively the sub-vectors in  $\mathbf{r}$ ,  $\mathbf{x}$  and  $\boldsymbol{\eta}$  at time instant  $j$ .

A block-wise DFT matrix is defined as

$$\mathbf{F}_n \equiv \mathbf{F} \otimes \mathbf{I}_n$$

where  $\mathbf{F}$  is the normalized DFT matrix,  $\mathbf{I}_n$  is an identity matrix of size  $n \times n$  and  $\otimes$  is the Kronecker product. Note that  $\mathbf{F}_n$  is of size  $Jn \times Jn$ , and it is easy to verify that  $\mathbf{F}_n \mathbf{F}_n^H = \mathbf{F}_n^H \mathbf{F}_n = \mathbf{I}$ . The properties in Section III.A can be extended as follows.

**Property III:** If  $\mathbf{H}$  is block-circulant as given in (12), then  $\mathbf{G} \equiv \mathbf{F}_M \mathbf{H} \mathbf{F}_N^H$  is block-diagonal, i.e.,  $\mathbf{G} = J^{\frac{1}{2}} \text{diag}\{\mathbf{G}_0, \mathbf{G}_1, \dots, \mathbf{G}_{J-1}\}$ ; and vice versa.

**Property IV:**  $\{\mathbf{H}_l\}$  (see (12)) and  $\{\mathbf{G}_j\}$  form a block-wise DFT pair, i.e.,

$$\mathbf{G}_j = J^{-\frac{1}{2}} \sum_{l=0}^{J-1} \mathbf{H}_l \exp(-i2\pi jl/J)$$

$$\mathbf{H}_l = J^{-\frac{1}{2}} \sum_{j=0}^{J-1} \mathbf{G}_j \exp(i2\pi jl/J).$$

### D. The ESE Operations for Block-Circulant $\mathbf{H}$

The FDE technique studied in Section III.B can be extended to block-circulant systems straightforwardly. Consider the

system model (12) and let  $\mathbf{V}_j \equiv \text{Cov}(\mathbf{x}_j, \mathbf{x}_j)$ . Each  $\mathbf{V}_j$  is diagonal, and we can define their average over  $j$  by

$$\bar{\mathbf{v}} \equiv \frac{1}{J} \sum_{j=0}^{J-1} \mathbf{V}_j. \quad (13a)$$

Note that for a single-user system in which a uniform coding scheme is employed,  $\bar{\mathbf{v}}$  can be further reduced to  $\bar{v}\mathbf{I}$ , where  $\bar{v}$  is the average *a priori* variance (see (8a)). More generally, we allow unequal diagonal entries for  $\bar{\mathbf{v}}$  in a multiuser system (to be discussed later) and let different diagonal entries of  $\bar{\mathbf{v}}$  represent the output variance of the decoders of different users. Similarly to the scalar case, we approximate  $\mathbf{V}$  by

$$\mathbf{V} = \text{diag}\{\underbrace{\bar{\mathbf{v}}, \bar{\mathbf{v}}, \dots, \bar{\mathbf{v}}}_{J \text{ repeats}}\} = \mathbf{I} \otimes \bar{\mathbf{v}}. \quad (13b)$$

Eqn. (13b) implies the following assumption.

**Assumption II:** (i) Entries in  $\mathbf{x}$  are uncorrelated, as is approximately ensured by the interleaver  $\boldsymbol{\Pi}$ ; and (ii) all  $\{\mathbf{x}_j\}$  have the same diagonal auto-covariance matrix  $\bar{\mathbf{v}}$ .

From (3), (13b) and Property III, we have

$$\mathbf{R} = \mathbf{F}_M^H \mathbf{G} \mathbf{F}_N^H \mathbf{V} \mathbf{F}_N^H \mathbf{G}^H \mathbf{F}_M + \sigma^2 \mathbf{I} = \mathbf{F}_M^H (\mathbf{G} \mathbf{V} \mathbf{G}^H + \sigma^2 \mathbf{I}) \mathbf{F}_M$$

where the last equality is due to the fact that

$$\mathbf{F}_N \mathbf{V} \mathbf{F}_N^H = (\mathbf{F} \mathbf{I} \mathbf{F}^H) \otimes (\mathbf{I}_N \bar{\mathbf{v}} \mathbf{I}_N^H) = \mathbf{V}.$$

From Property III,  $\mathbf{R}$  is block-circulant since  $\mathbf{G} \mathbf{V} \mathbf{G}^H + \sigma^2 \mathbf{I}$  is block-diagonal. Then, (4c) can be rewritten as

$$\boldsymbol{\lambda} = 2(\mathbf{I} - \mathbf{V}\mathbf{U})^{-1} (\mathbf{F}_N^H \mathbf{G}^H (\mathbf{G} \mathbf{V} \mathbf{G}^H + \sigma^2 \mathbf{I})^{-1} \cdot (\mathbf{F}_M \mathbf{r} - \mathbf{G} \mathbf{F}_N \mathbf{E}(\mathbf{x})) + \mathbf{U} \mathbf{E}(\mathbf{x})) \quad (14a)$$

where

$$\mathbf{U} \equiv (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})_{\text{diag}} = \mathbf{I} \otimes \mathbf{u} \quad (14b)$$

and

$$\mathbf{u} \equiv \left( \sum_{j=0}^{J-1} \mathbf{G}_j^H (\mathbf{J} \mathbf{G}_j \bar{\mathbf{v}} \mathbf{G}_j^H + \sigma^2 \mathbf{I})^{-1} \mathbf{G}_j \right)_{\text{diag}}. \quad (14c)$$

Note that  $\mathbf{u}$  is the block equivalence of the scalar  $u$  defined in (10d). Eqn. (14a) can be implemented in a way similarly to (11). Multiplying a vector by  $\mathbf{F}_n$  or its conjugate can be realized by the FFTs or IFFT with complexity  $O(n \log J)$ . The cost for evaluating  $\mathbf{G} \mathbf{V} \mathbf{G}^H$ ,  $(\mathbf{G} \mathbf{V} \mathbf{G}^H + \sigma^2 \mathbf{I})^{-1}$  (by invoking the matrix inversion lemma similarly to (51-53) in [4]), and  $\mathbf{U}$  are  $O(JM^2N)$ , respectively. Thus, by noting the length of  $\mathbf{x}$  is  $NJ$ , the normalized cost involved in (14a) is roughly  $O((1 + M/N) \log J + M^2)$ .

### E. Extension to Complex Multiuser Systems

The generalization to a multiuser MIMO system is straightforward. Consider an approximately synchronized multiple-access MIMO system modeled by (1) in which the synchronization errors between users can be handled by the CP

technique. Then, (1) can still be expressed in a block-circulant form as (12) except that

$$\mathbf{H}_l = \begin{pmatrix} h_l^{(1,1,1)} \dots h_l^{(1,N,1)} \dots h_l^{(1,1,k)} \dots h_l^{(1,N,k)} \dots \\ \vdots \quad \ddots \quad \vdots \quad \ddots \quad \vdots \\ h_l^{(M,1,1)} \dots h_l^{(M,N,1)} \dots h_l^{(M,1,k)} \dots h_l^{(M,N,k)} \dots \end{pmatrix} \quad (15)$$

where  $h_l^{(m,n,k)}$  denotes the  $l$ th tap coefficient between the  $n$ th transmit antenna and the  $m$ th receive antenna for user  $k$ . Thus, Eqn. (14a) can still be applied here. Let  $K$  be the total number of users in the system. The normalized cost involved in (14a) is  $O((1+M/(NK))\log J + M^2)$ . When  $K$  is large, we usually have  $NK \gg M$  and so the normalized cost is roughly  $O(\log J + M^2)$ . This method is numerically very efficient for small or moderate  $M$ . It is also interesting to note that the normalized cost is independent of  $K$ ,  $N$  and also the ISI delay spread.

The treatment for a complex system is also straightforward. For example, if (12) is complex, it can be converted into an equivalent real form as

$$\begin{pmatrix} \vdots \\ \text{Re}(\mathbf{r}_j) \\ \text{Im}(\mathbf{r}_j) \\ \vdots \end{pmatrix} = \begin{pmatrix} \ddots & \vdots & \vdots \\ \text{Re}([\mathbf{H}]_{l,j}) & -\text{Im}([\mathbf{H}]_{l,j}) \\ \text{Im}([\mathbf{H}]_{l,j}) & \text{Re}([\mathbf{H}]_{l,j}) \\ \vdots & \vdots & \ddots \end{pmatrix} \cdot \begin{pmatrix} \vdots \\ \text{Re}(\mathbf{x}_j) \\ \text{Im}(\mathbf{x}_j) \\ \vdots \end{pmatrix} + \begin{pmatrix} \vdots \\ \text{Re}(\boldsymbol{\eta}_j) \\ \text{Im}(\boldsymbol{\eta}_j) \\ \vdots \end{pmatrix}. \quad (16)$$

The system in (16) is equivalent to a real  $2M \times 2N$  MIMO system. The real and imaginary parts of each  $\mathbf{x}_j$  in (16) may contain independent bits, which corresponds to QPSK modulation with Gray mapping. Then (14a) can again be applied directly.

#### IV. EVOLUTION ANALYSIS

##### A. Circulant Systems

The evolution technique tracks parameters in the iterative process using pre-calculated transfer functions. In order to reduce complexity, there should be as few parameters as possible. Further, these parameters should accurately characterize the statistical behavior of the iterative process, so as to predict performance. For a system with a circulant channel matrix in (7), we find that it is most convenient to use the input average variance (denoted by  $\bar{v}$ ) and output SNR (denoted by  $\rho$ ) to characterize the ESE, and the input SNR  $\rho$  and output average variance  $\bar{v}$  to characterize the DEC. Thus, the density evolution of the iterative equalization/detection process reduces to a simple recursion between  $\rho$  and  $\bar{v}$ . The transfer functions of the ESE and the DEC can be denoted by  $\rho = \phi(\bar{v})$  and  $\bar{v} = \psi(\rho)$ , respectively.

Similarly to the EXIT approach [17], both  $\phi(\cdot)$  and  $\psi(\cdot)$  can be obtained by simulation. Provided that  $\phi(\cdot)$  and  $\psi(\cdot)$  are available, we can track the evolution of  $\rho$  and  $\bar{v}$  during the iteration, as shown in Fig. 2. For example, we can start with an initial value of  $\bar{v} = 1$ . (For BPSK modulation over  $\{+1, -1\}$ ,  $\bar{v} = 1$  implies no *a priori* information.) Then the evolution

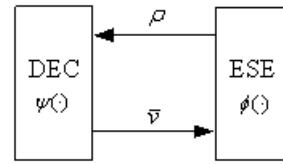


Fig. 2. An illustration of the overall evolution process.

process is as follows: after the first iteration,  $\rho = \phi(1)$ ; after the second iteration,  $\rho = \phi(\psi(\phi(1)))$ ; and so on. The SNR value resulting from a specified number of iterations can be used to predict the system performance [18].

The above principle is similar to an EXIT chart method in which mutual information is tracked. The EXIT chart approach [6] is efficient for fixed channels in which the transfer functions can be pre-calculated by simulation and stored as a table. This strategy may, however, encounter difficulties if we want to evaluate the statistical system behavior in quasi-static fading channels, since the transfer functions  $\phi(\cdot)$  and  $\psi(\cdot)$  are different for different channel realizations. It is impractical to pre-calculate and store them for all possible channel realizations.

In the following, we present a solution to this difficulty by making a number of approximations. We will demonstrate by numerical results that the proposed method can accurately characterize the behavior of the iterative receiver in Fig. 1. We first explain our technique for a circulant system in a SISO channel.

Recall the expression of the ESE output given in (5):

$$\lambda_j = \mu_j x_j + \zeta_j.$$

As the consequence of approximation (8b), substituting (9) and (10c) into (6b) leads to

$$\begin{aligned} \text{diag}\{\rho_0, \rho_1, \dots, \rho_{J-1}\} &= \frac{1}{2} \text{diag}\{\mu_0, \mu_1, \dots, \mu_{J-1}\} \\ &= \frac{1}{4} \text{diag}\{E(\zeta_0^2), E(\zeta_1^2), \dots, E(\zeta_{J-1}^2)\} \\ &= u(1 - \bar{v}u)^{-1} \mathbf{I}. \end{aligned} \quad (17)$$

Thus,  $\mu_j = 2u(1 - \bar{v}u)^{-1}$ ,  $E(\zeta_j^2) = 4u(1 - \bar{v}u)^{-1}$  and  $\rho_j = u(1 - \bar{v}u)^{-1}$  are all constant with respect to  $j$ . We can thus write

$$\rho = \rho_1 = \rho_2 = \dots = \rho_{J-1}$$

and use a single parameter  $\rho$  to characterize the ESE outputs  $\{\lambda_j\}$ . Thus the transfer function of the ESE is given by

$$\rho = \phi(\bar{v}) = u(1 - \bar{v}u)^{-1}. \quad (18)$$

With (18), we can compute  $\phi(\cdot)$  online for each channel realization instead of pre-calculating and storing  $\phi(\cdot)$  for all possible channel realizations.

Next, for a SISO channel,  $\psi(\cdot)$  can be obtained by pre-simulation, which is similar to the EXIT chart technique [6][17]. We make the following assumption.

**Assumption III:** For a circulant system,  $\{\zeta_j\}$  can be treated as independent samples of a zero-mean Gaussian random variable  $\zeta$  with  $E(\zeta^2) = 4u(1 - \bar{v}u)^{-1}$ .

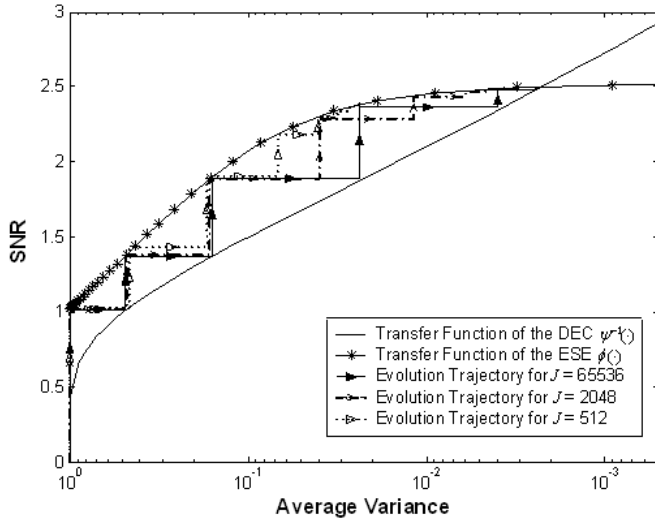


Fig. 3. SNR-variance transfer chart for the iterative equalization with convolutional code  $(23, 35)_8$  in the Proakis B channel.  $E_b/N_0 = 4\text{dB}$ .

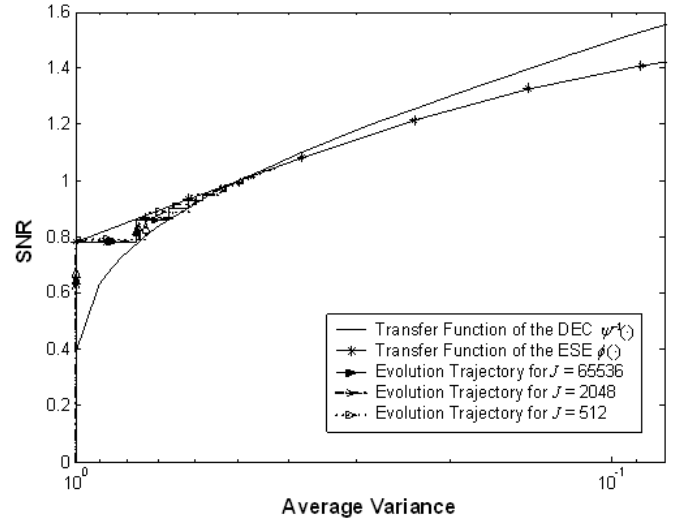


Fig. 4. SNR-variance transfer chart for the iterative equalization with convolutional code  $(23, 35)_8$  in the Proakis B channel.  $E_b/N_0 = 2\text{dB}$ .

The above assumption can be ensured by a random interleaver (see Fig. 1) with sufficient interleaving depth. Similar assumptions have been widely used in evolution analysis for LDPC or turbo codes [17][25] and for SNR analysis of LMMSE estimators [22].

Based on Assumption III, for a circulant system, the inputs of the decoder can be treated as a sequence of samples from a memoryless binary AWGN channel with  $\text{SNR} = \rho$ . We can thus generate  $\psi(\cdot)$  by simulating the DEC on AWGN channels.

### B. Block-Circulant Systems

The above evolution analysis can be generalized to block-circulant systems. We first consider single-user MIMO systems defined in (12). The extension to multiuser systems is straightforward.

For simplicity and without confusion, we will often use a set to represent a row vector or a column vector. For example, we may write

$$\mathbf{x}_j = \{x_j^{(1)}, x_j^{(2)}, \dots, x_j^{(N)}\} \quad (19)$$

even though  $\mathbf{x}_j$  is actually a column vector.

Return to (12) where each  $\mathbf{x}_j$  can be decomposed as in (19) with  $x_j^{(n)}$  denoting the symbol transmitted over the  $n$ th antenna at time instant  $j$ . Let  $\lambda_j^{(n)}$  be the output of the ESE corresponding to  $x_j^{(n)}$ . Again use the model in (5):  $\lambda_j^{(n)} = \mu_j^{(n)} x_j^{(n)} + \zeta_j^{(n)}$ . Define

$$\boldsymbol{\mu}_j \equiv \text{diag} \{ \mu_j^{(1)}, \mu_j^{(2)}, \dots, \mu_j^{(N)} \}$$

and

$$\boldsymbol{\zeta}_j \equiv \text{diag} \{ \zeta_j^{(1)}, \zeta_j^{(2)}, \dots, \zeta_j^{(N)} \}.$$

Also define the variance matrix for  $\boldsymbol{\zeta}_j$  and SNR matrix for  $\boldsymbol{\lambda}_j$  as, respectively,

$$\boldsymbol{\omega}_j \equiv \text{diag} \{ E(|\zeta_j^{(1)}|^2), E(|\zeta_j^{(2)}|^2), \dots, E(|\zeta_j^{(N)}|^2) \}$$

and

$$\boldsymbol{\rho}_j \equiv \text{diag} \left\{ \frac{(\mu_j^{(1)})^2}{E(|\zeta_j^{(1)}|^2)}, \frac{(\mu_j^{(2)})^2}{E(|\zeta_j^{(2)}|^2)}, \dots, \frac{(\mu_j^{(N)})^2}{E(|\zeta_j^{(N)}|^2)} \right\}.$$

Then, substituting (13b) and (14b) into (6b) leads to

$$\begin{aligned} \text{diag} \{ \boldsymbol{\rho}_0, \boldsymbol{\rho}_1, \dots, \boldsymbol{\rho}_{J-1} \} &= \frac{1}{2} \text{diag} \{ \boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_{J-1} \} \\ &= \frac{1}{4} \text{diag} \{ \boldsymbol{\omega}_0, \boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_{J-1} \} \\ &= \mathbf{I} \otimes (\mathbf{u}(\mathbf{I} - \bar{\mathbf{v}}\mathbf{u})^{-1}). \end{aligned} \quad (20)$$

Thus  $\boldsymbol{\mu}_j$ ,  $\boldsymbol{\omega}_j$  and  $\boldsymbol{\rho}_j$  are all constant with respect to  $j$ . We can then write

$$\boldsymbol{\rho} = \boldsymbol{\rho}_1 = \boldsymbol{\rho}_2 = \dots = \boldsymbol{\rho}_{J-1}. \quad (21)$$

From (20) and (21), the transfer function of the ESE is

$$\boldsymbol{\rho} = \phi(\bar{\mathbf{v}}) = \mathbf{u}(\mathbf{I} - \bar{\mathbf{v}}\mathbf{u})^{-1}. \quad (22)$$

We will now consider the corresponding transfer function  $\bar{\mathbf{v}} = \psi(\boldsymbol{\rho})$  for the DEC. We first make the following assumption that is a generalization of Assumption III.

**Assumption IV:** For the block-circulant system in (12),  $\{\zeta_j\}$  can be treated as independent samples of a zero mean Gaussian diagonal matrix  $\boldsymbol{\zeta}$  with  $E(\boldsymbol{\zeta}\boldsymbol{\zeta}^T) = 4\mathbf{u}(\mathbf{I} - \bar{\mathbf{v}}\mathbf{u})^{-1}$ .

The similarity and difference for Assumptions III and IV are compared below.

- An underlying intuition for Assumption III is that a circulant system (see (7)) represents a quasi-static environment in which all the signals at different time instants undergo a uniform channel condition, except for the additive noise. Therefore a single SNR value  $\rho$  is sufficient to characterize the ESE outputs at different time instants  $j$ , under the assumption that the feedback from the DEC has the same variance value  $\bar{v}$ .
- For a block circulant system,  $\mathbf{x}_j = \{x_j^{(1)}, x_j^{(2)}, \dots, x_j^{(N)}\}$  are transmitted, e.g., over  $N$  different antennas and they

can experience very different channel conditions. In this case, a single parameter is not sufficient to characterize the ESE outputs. However, we can still assume that the ESE estimation for  $\{x_0^{(n)}, x_1^{(n)}, \dots, x_j^{(n)}, \dots, x_{J-1}^{(n)}\}$  can be characterized by a single SNR value since the transmission environment is still uniform (quasi-static) for different time instants  $j$ . In this way, we need  $N$  parameters to characterize the ESE outputs.

In the single-user case,  $\bar{\mathbf{v}}$  can be reduced to  $\bar{v}\mathbf{I}$  (see the discussion below (13a)). The DEC input (i.e., the interleaved version of the ESE output) can be treated as samples with  $N$  different SNRs specified by the diagonal of  $\boldsymbol{\rho}$  in (22). We call an additive Gaussian noise channel that has its SNR randomly selected from  $N$  candidates as an  $N$ -state fading channel. Then,  $\psi(\cdot)$  can be obtained by simulating the DEC on the corresponding  $N$ -state fading channels. The set of  $N$  SNRs may differ for different channel realizations. If we simulate and store all possible combinations, an  $N$ -dimensional table is required. This method is referred to as the full-table (FT) technique, which is very costly even for a moderate  $N$ .

The following low-complexity approach is inspired by the rationale behind the EXIT chart analysis [17], i.e., given an LLR sequence  $\boldsymbol{\lambda}$  as the input to an APP decoder, its output behavior remains approximately unchanged by replacing  $\boldsymbol{\lambda}$  with another sequence  $\boldsymbol{\lambda}'$  that contains the same amount of mutual information (with respect to the transmitted  $\mathbf{x}$ ) as  $\boldsymbol{\lambda}$ . We write this as follows.

**Assumption V:** The DEC behavior remains the same if its input  $\boldsymbol{\lambda}$  is replaced by another LLR sequence  $\boldsymbol{\lambda}'$  that contains the same amount of mutual information with respect to the transmitted symbols as  $\boldsymbol{\lambda}$  does.

This assumption is true for an “ideal” decoder since such a decoder should be able to work up to the limit indicated by the mutual information. For a practical decoder, however, it is difficult to provide a rigorous proof although we have verified it in various simulation environments. The mutual information contained in an LLR  $\lambda$  from an AWGN channel with SNR equal to  $\rho$  (where  $x \in \{+1, -1\}$  is transmitted) is given by [17]

$$I(\lambda; x) = \frac{1}{\sqrt{8\pi\rho}} \int_{-\infty}^{+\infty} \left( \log \frac{1}{1+e^{-\lambda}} \right) e^{-\frac{(\lambda-2\rho)^2}{8\rho}} d\lambda \equiv f(\rho). \quad (23)$$

Let  $\bar{v} = \psi_{\text{AWGN}}(\rho)$  be the transfer function of the DEC in an AWGN channel. This function can be easily generated via simulation and stored as a one-dimensional table. Now we adopt the following simple strategy to generate  $\psi(\cdot)$ . We select such an effective SNR value  $\rho_{\text{eff}}$  that the observations on the AWGN channel with SNR =  $\rho_{\text{eff}}$  contain the same amount of mutual information as the actual sequence does, i.e.,

$$\rho_{\text{eff}} = f^{-1} \left( N^{-1} \sum_{n=1}^N f(\rho^{(n)}) \right)$$

where  $\rho^{(n)}$  denotes the  $n$ th diagonal entry of  $\boldsymbol{\rho}$ , and  $f^{-1}(\cdot)$  is the inverse of  $f(\cdot)$  that exists since  $f(\cdot)$  is monotonic. We then select  $\psi(\boldsymbol{\rho}) = \psi_{\text{AWGN}}(\rho_{\text{eff}})$ , based on Assumption V. Both

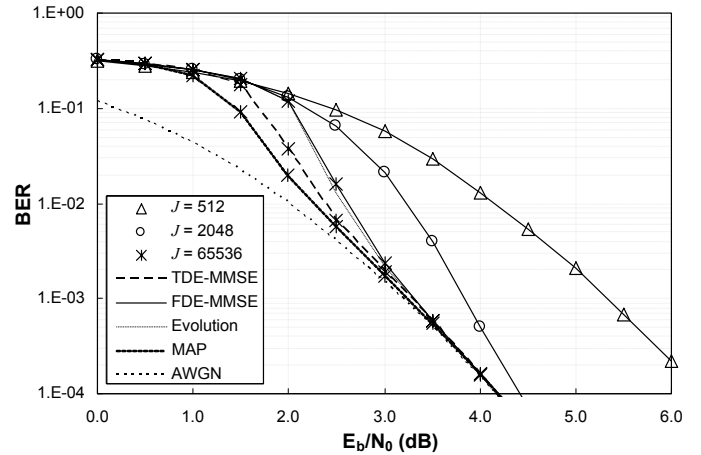


Fig. 5. BER performance with different block length  $J$  compared with prediction by evolution analysis in the Proakis B channel. Convolutional code  $(23, 35)_8$  is used.

$f(\cdot)$  and  $\psi_{\text{AWGN}}(\cdot)$  can be pre-calculated and stored as one-dimensional tables. We refer to this method as the equivalent-mutual-information (EMI) technique.

For the multiuser case, the generation of  $\phi(\cdot)$  follows the same principle as discussed above. We can decompose the transfer function  $\bar{\mathbf{v}} = \psi(\boldsymbol{\rho})$  for the overall DEC (including all the individual decoders) into user-by-user forms, with some abuse of notation, denoted by  $\bar{v}_{(k)} = \psi(\boldsymbol{\rho}_{(k)})$ , where  $\bar{v}_{(k)}$  is the output variance of the DEC for user  $k$  and  $\boldsymbol{\rho}_{(k)}$  includes the output SNRs of the ESE for user  $k$ . We can thus apply the FT or EMI technique in a user-by-user manner.

## V. NUMERICAL RESULTS

### A. Settings and Definitions

We first outline some common settings and definitions used in the following simulations. The quasi-static Rayleigh fading channels are modeled as follows: the fading coefficients are independently and identically distributed complex Gaussian random variables with zero-mean, and the average total energy normalized to 1, i.e.,

$$E_h \equiv E[|\mathbf{h}_j|^2] = 1, \quad \text{for } \forall j.$$

Thus the bit-energy-to-noise-density ratio can be calculated as

$$\frac{E_b}{N_0} = \frac{E[|x_j|^2] \cdot E_h}{R_c N_0} = \frac{1}{2R_c \sigma^2}$$

where  $E_b$  denotes the bit energy,  $N_0$  denotes the single-sideband noise power spectral density, and  $R_c$  denotes the information rate of  $\mathbf{x}$ .

In simulation, the CP insertion is always applied at the transmitter side (and CP removal at the receiver side) to ensure that the channel matrix is circulant or block-circulant. Also, the influence of CP to the power and spectral efficiency is not considered here. At least 10 iterations are taken in the iterative detection to guarantee convergence in simulation, and for each point on the performance curves at least 1000 errors are collected. For convenience of discussion, the implementation methods based on (11) and (14a) are called “FDE-MMSE”, and their counterparts [6-9], in which the ESE is realized by

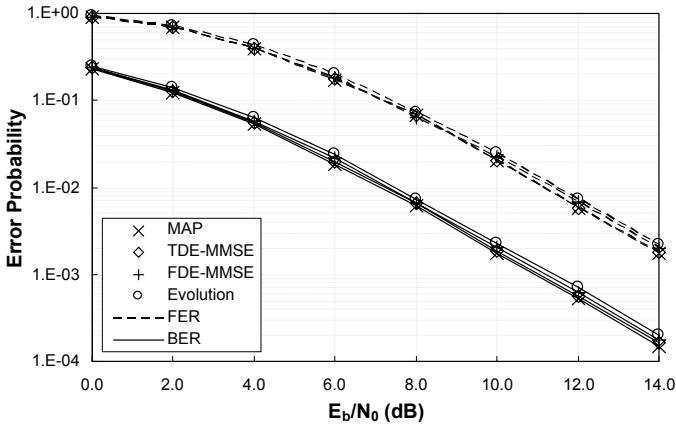


Fig. 6. FER/BER performance on the Rayleigh fading ISI channel with  $L = 3$ . Convolutional code  $(23, 35)_8$  is used. Information length = 1024.

exactly calculating (4c) in the time domain, are called “TDE-MMSE”.

### B. Examples for Fixed ISI Channels

We first consider a simple ISI channel, where the ENC in Fig. 1 contains only one encoder that is realized by a rate- $1/2$  convolutional code with generation polynomials  $(23, 35)_8$ . Proakis B Channel [23] with tap coefficients  $[0.410 \ 0.815 \ 0.410]$  is employed for demonstration. Figs. 3 and 4 show the SNR-variance transfer chart for the FDE-MMSE at  $E_b/N_0 = 4$  dB and  $E_b/N_0 = 2$  dB, respectively. The evolution trajectories are obtained by averaging over 1000 frames with information length = 256, 1024, and 32768 (or equivalently,  $J = 512, 2048$ , and 65536), respectively, and within each frame the ESE’s output SNR is calculated by

$$\text{SNR} = \mu_\lambda^2 / \sigma_\lambda^2$$

where

$$\mu_\lambda = \frac{1}{J} \sum_{j=0}^{J-1} \lambda_j x_j \quad \text{and} \quad \sigma_\lambda^2 = \frac{1}{J} \sum_{j=0}^{J-1} \lambda_j^2 - \mu_\lambda^2.$$

It is shown in Figs. 3 and 4 that the behavior of the iterative receiver, represented by the evolution trajectories, is in good agreement with the transfer characteristics of the ESE and DEC only when  $J = 65536$ . Recall that throughout this paper, we have assumed the independency of the LLRs passed between the ESE and the DEC (see Assumptions I-IV). This is a good approximation only when  $J$  is sufficiently large.

The bit-error-rate (BER) performance of the system in Figs. 3 and 4 is shown in Fig. 5. We can see that, as  $J$  increases, the simulated performance becomes closer to its asymptotical limit (as specified by the evolution analysis), which is consistent with the observations in Figs. 3 and 4. Note that Proakis B channel is a severely distorted ISI channel. As shown later, for most channel realizations, the simulated performance agrees well with the prediction at a much smaller  $J$ . For comparison, we also include in Fig. 5 the performance curves for TDE-MMSE and MAP as well as that of the system in the AWGN channel. Implementation details for the MAP algorithm can be found, for example, in [24]. Note that the performance gap between FDE-MMSE and TDE-MMSE is

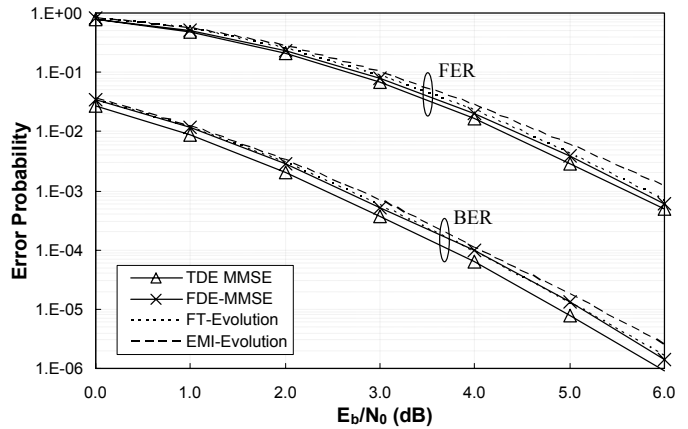


Fig. 7. BER/BER performance and evolution prediction on the quasi-static Rayleigh fading  $2 \times 2$  MIMO ISI channel with  $L = 4$ . Convolutional code  $(7, 5)_8$  is used. Information length = 1024.

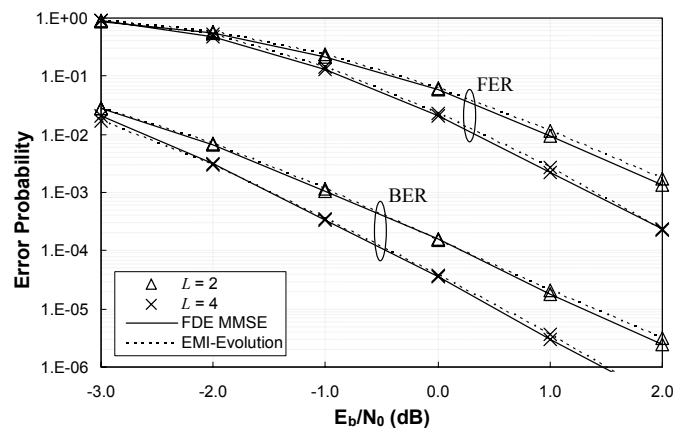


Fig. 8. BER/BER performance and evolution prediction on the quasi-static Rayleigh fading  $4 \times 4$  MIMO ISI channel with different  $L$ . Convolutional code  $(7, 5)_8$  is used. Information length = 1024.

caused by the assumption in (8b). (Otherwise, the time and frequency domain realization techniques will only differ in computation methods but not in performance.) We can see from Fig. 5 that the performance loss due to (8b) is marginal.

### C. Examples for Fading ISI Channels

We next consider quasi-static Rayleigh fading channels. The same convolutional code is employed at the encoder. QPSK modulation is applied to  $x$ . More specifically, after random interleaving, the scrambled codeword is separated into two bit streams, each BPSK modulated, to form the real and imaginary parts of  $x$ . Fig. 6 shows the BER and frame-error-rate (FER) performance of several implementation techniques, namely, MAP, FDE-MMSE and TDE-MMSE, for the ESE in quasi-static Rayleigh fading ISI channels with  $L = 3$  and  $J = 1024$ . These algorithms exhibit similar performance, although their related implementation costs differ significantly. The evolution prediction for FDE-MMSE is also included in Fig. 6 for reference. Interestingly, since FDE-MMSE performance should be the worst among these options (as it involves more approximations), the evolution prediction for FDE-MMSE can also serve as an upper-bound for TDE-MMSE or MAP.

#### D. Examples for Fading MIMO ISI Channels

Now we study MIMO systems. The ENC contains one encoder formed by a rate-1/2  $(7, 5)_8$  convolutional code with information length 1024. The coded bits are randomly interleaved and separated into two streams to form the real and imaginary parts of the transmit symbol vector. This vector is separated into  $N$  streams for simultaneous transmissions over  $N$  transmit antennas. This transmitter structure follows the interleave division multiplexing space-time (IDM-ST) coding technique discussed in [9].

Fig. 7 shows the BER/FER performance of the system in quasi-static Rayleigh-fading  $2 \times 2$  MIMO ISI channels with  $L = 4$ . Both FDE-MMSE and TDE-MMSE are considered. It can be seen that the performance loss of FDE-MMSE due to the approximation in (13b) is again not significant. Fig. 7 also includes the performance curves predicted by evolution. It is shown that the performance of FDE-MMSE corresponds with the prediction by the full-table method (denoted by "FT-Evolution"), whereas the prediction by the EMI method (denoted by "EMI-Evolution") deviates about 0.1-0.2 dB away from the actual performance.

In Fig. 8, we consider quasi-static Rayleigh-fading  $4 \times 4$  MIMO ISI channels with  $L = 2$  and 4, respectively. The evolution predictions are based on the EMI method only since the FT method becomes too complicated here. We can see that the EMI method provides better predictions as the MIMO dimension and/or channel memory length increases.

#### E. Examples for Fading Multiuser MIMO ISI Channels

We finally consider multiuser MIMO systems. The transmitter structure basically follows the interleave-division multiple-access (IDMA) principle [7]. More specifically, the transmitter for each user has a similar structure as that in Figs. 7 and 8, with the only difference being that the encoder for each user consists of a rate-1/2  $(7, 5)_8$  convolutional code with information length 1024 and a following-up length-2 spreader. The spreading sequence is  $(+1, -1)$  to remove the direct-current (DC) component.

Fig. 9 shows the BER/FER performance of this multiple-access system in quasi-static Rayleigh fading channels with  $N = 2$ ,  $M = 4$ , and  $L = 4$ . Single-user performance is also included for comparison. It is shown that the system can support a total throughput of 8 while it can still achieve almost single-user performance. Evolution analysis again demonstrates that the predicted performance is in good agreement with the simulation.

## VI. CONCLUSIONS

We have considered iterative equalization/detection in a generic coded linear system and presented a concise derivation of the LMMSE estimator for the realization of the ESE. With the aid of the CP technique, iterative FDE techniques are developed for multiuser MIMO ISI channels. We have analyzed the behavior of the resulting iterative detector, and developed the SNR-variance evolution technique for fast evaluation of system performance. Numerical results have been provided for demonstration. The proposed SNR-variance transfer functions

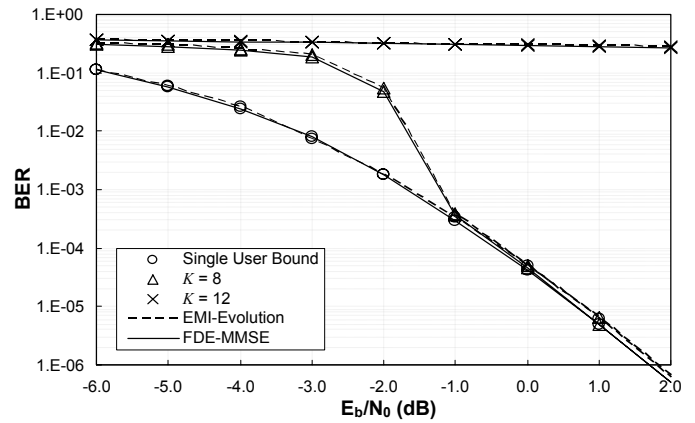


Fig. 9. Comparison between the simulated and predicted performance in quasi-static Rayleigh-fading multi-user  $2 \times 4$  MIMO ISI channels with  $L = 4$ . Convolutional code  $(7, 5)_8$  with information length 1024 is employed for each user.

can accurately describe the evolution behavior of the FDE-MMSE based iterative detection systems. We are currently looking at the extension of the proposed detection and evolution techniques in higher-order modulated systems.

## ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable suggestions and comments.

## REFERENCES

- [1] C. Douillard *et al.*, "Iterative correction of inter-symbol interference: Turbo equalization," *Eur. Trans. Telecommun.*, vol. 6, pp. 507-511, Sept.-Oct., 1995.
- [2] M. Moher, "An iterative multiuser decoder for near-capacity communications," *IEEE Trans. Commun.*, vol. 46, pp. 870-880, July 1998.
- [3] M. C. Reed, C. B. Schlegel, P. D. Alexander, and J. A. Asenstorfer, "Iterative multiuser detection for CDMA with FEC: Near-single-user performance," *IEEE Trans. Commun.*, vol. 46, pp. 1693-1699, Dec. 1998.
- [4] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, pp. 1046-1061, July 1999.
- [5] X. Wang and H. V. Poor, "Space-Time multiuser detection in multipath CDMA channels," *IEEE Trans. Signal Processing*, vol. 47, pp. 2356-2374, Sept. 1999.
- [6] M. Tüchler, R. Koetter, and A. C. Singer, "Turbo equalization: Principles and new results," *IEEE Trans. Commun.*, vol. 50, pp. 754-767, May 2002.
- [7] Li Ping, L. Liu, K. Wu, and W. K. Leung, "Interleave division multiple-access," *IEEE Trans. Wireless Commun.*, vol. 5, no. 4, pp. 938-947, Apr. 2006.
- [8] L. Liu and Li Ping, "An extending window MMSE turbo equalization algorithm," *IEEE Signal Processing Lett.*, vol. 11, no. 11, pp. 891-894, Nov 2004.
- [9] X. Yuan, K. Wu, and Li Ping, "The jointly Gaussian approach to iterative detection in MIMO systems," in *Proc. IEEE Int. Conf. on Commun., ICC'06*, Istanbul, Turkey, 11-15 June 2006.
- [10] M. Tüchler, and J. Hagenauer, "Turbo equalization using frequency domain equalizers," in *Proc. of Allerton Conference*, Monticello, IL, USA, Oct. 2000.
- [11] D. Falconer, S. L. Ariyavisitakul, A. Benyamini-Seeyar, and B. Eldson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 40, no. 4, pages 58-66, April 2002.
- [12] K. Kansanen, T. Matsumoto, C. Schneider, and R. Thoma, "Frequency domain MMSE turbo equalization of multilevel coded QAM convergence in real fields," in *Proc. IEEE Int. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC'05)*, Berlin, Germany, 11-14 Sept. 2005.

- [13] R. Visoz, A. O. Berthet, and S. Chtourou, "Frequency-domain block Turbo-equalization for single-carrier transmission over MIMO broadband wireless channel," *IEEE Trans. Commun.*, vol. 54, pp. 2144-2149, Dec 2006.
- [14] Y. Wu, X. Zhu, and A. K. Nandi, "Low complexity Turbo space-frequency equalization for single-carrier MIMO wireless communications," in *Proc. EUSIPCO'06*, Sep 2006.
- [15] Amis, Le Josse, and Laot, "Efficient frequency-domain MMSE Turbo equalization derivation and performance comparison with the time-domain counterpart," in *Proc. Third Int. Conf. on Wireless and Mobile Communications (ICWMC'07)*, pp. 65, 2007.
- [16] 3rd Generation Partnership project, Technical Specification Group Radio Access Network, Physical Layer Aspects for Evolved UTRA (Release 7), 3GPP TR25.814 V1.0.1 (2005-11).
- [17] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, Oct. 2001.
- [18] D. Divsalar, S. Dolinar, and F. Pollara, "Serial Turbo trellis coded modulation with rate-1 inner code," in *Proc. Int. Symp. Information Theory*, 2000, pp. 194.
- [19] Li Ping and L. Liu, "Analysis and design of IDMA based on SNR evolution and power allocation," in *Proc. VTC' 2004-Fall*, Los Angeles, CA, Sept. 2004, pp. 1068-1072.
- [20] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge: Cambridge University Press, 2005.
- [21] A. Glavieux, C. Laot, and J. Labat, "Turbo equalization over a frequency selective channel," in *Int. Symp. on Turbo Codes*, pp. 96-102, Brest, France, 1997.
- [22] V. Poor and S. Verdú, "Probability of error in MMSE multiuser detection," *IEEE Trans. Inform. Theory*, vol. IT-43, pp. 835-847, May 1997.
- [23] J. G. Proakis, *Digital Communications*, New York: Mcgraw-Hill Companies Inc, 2001.
- [24] J. Hagenauer, "The Turbo principle: Tutorial introduction and state of the art," in *Proc. Int. Symp. On Turbo Codes*, Brest, France, Sept. 1997, pp. 1-11.
- [25] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. 47, No. 2, Feb. 2001.



**Qinghua Guo** (S'07) received his B.E. degree in electronic engineering and M.E. degree in signal and information processing from Xidian University, China, in 2001 and 2004, respectively. He is currently working towards the Ph.D. degree at City University of Hong Kong. His research interests include statistical signal processing and broadband wireless communications.



**Xiaodong Wang** (S'98-M'98-SM'04-F'07) received the Ph.D degree in Electrical Engineering from Princeton University. He is on the faculty of the Electrical Engineering Department, Columbia University.

Dr. Wang's research interests fall in the general areas of computing, signal processing and communications. Among his publications is a book entitled "Wireless Communication Systems: Advanced Techniques for Signal Reception", published in 2003 by Prentice Hall, Upper Saddle River, NJ. His current research interests include wireless communications, statistical signal processing, and genomic signal processing. Dr. Wang received the 1999 NSF CAREER Award, and the 2001 IEEE Communications Society and Information Theory Society Joint Paper Award. He has served as an Associate Editor for the IEEE Transactions on Communications, the IEEE Transactions on Wireless Communications, the IEEE Transactions on Signal Processing, and the IEEE Transactions on Information Theory. He is a Fellow of the IEEE.



**Li Ping** (S'87-M'91-SM'06) received his Ph.D degree at Glasgow University in 1990. He lectured at Department of Electronic Engineering, Melbourne University, from 1990 to 1992, and worked as a member of research staff at Telecom Australia Research Laboratories from 1993 to 1995. He has been with the Department of Electronic Engineering, City University of Hong Kong, since January 1996, where he is now a chair professor.

His research interests are communications systems and coding theory. Dr. Li Ping was awarded a British Telecom-Royal Society Fellowship in 1986, the IEE J J Thomson premium in 1993 and a Croucher Senior Research Fellowship in 2005.



**Xiaojun Yuan** (S'04) received the B.S. degree in electronic and information systems from Shanghai Jiaotong University, and the M.S. degree in circuit and systems from Fudan University, Shanghai, China, in 2003. He is currently working towards the Ph.D. degree at the City University of Hong Kong. His research interests are communication systems and coding design.