

# Evolution Analysis of Iterative LMMSE-APP Detection for Coded Linear System with Cyclic Prefixes

Xiaojun Yuan, Qinghua Guo, and Li Ping, *Member IEEE*

Department of Electronic Engineering, City University of Hong Kong, Hong Kong  
Email: xjyuan@cityu.edu.hk, qh.guo@student.cityu.edu.hk, and eeliping@cityu.edu.hk

**Abstract**—This paper is concerned with the iterative detection principles for coded linear systems with cyclic prefixes. We derive a matrix-form low-cost fast Fourier transform (FFT) based iterative LMMSE-APP detector and propose an evolution technique for the performance evaluation of the proposed detector. Numerical results show a good match between simulation and evolution prediction.

**Keywords**—LMMSE-APP detection, Turbo equalization, iterative detection, SNR-variance evolution.

## I. INTRODUCTION

Iterative minimum-mean-square-error (MMSE) detection has been studied in the context of a variety of applications involving inter-symbol interference (ISI), multiple-input-multiple-output (MIMO) transmission, and multiple-access (MAC) [1-4]. It can provide performance comparable to maximum *a posteriori* (MAP) detection [5, 6] at considerably reduced cost.

In this paper, we study the linear MMSE (LMMSE) approach to systems with circulant channel matrices. Such matrices arise when cyclic prefixes are added to the transmitted signals, which is a technique considered, e.g., in 3GPP long term evolution systems [7]. We first derive a matrix form FFT-based iterative LMMSE-APP (APP for “*a posteriori* probability”) detector, and then propose an SNR-variance evolution technique for its performance evaluation. The extrinsic information transfer (EXIT) chart technique [2, 8] is a useful tool for the analysis of iterative detectors in fixed channels. However, its application becomes difficult when channels are not fixed, such as in the case of MIMO ISI quasi-static fading channels. In this case, the average performance of a system can be computed by collecting statistics for a sufficiently large number of channel realizations. For each realization, a different pre-simulated transfer function of the detector is required for the EXIT chart analysis. It is generally impractical to store the pre-simulated EXIT transfer functions for all possible channel realizations. In this paper, we therefore develop an alternative solution in which the transfer function is generated analytically (rather than pre-simulated) at very low cost for each channel realization during the evolution process. The overall system performance can be evaluated by averaging over different channel realizations.

**List of Notations:** Vectors are expressed in bold letters and are column vectors by default. Matrices are specified by bold capital letters. The linear equation in (1) (see below) can be expressed in a block form, where each block may represent a scalar, sub-vector or sub-matrix. The length of  $\mathbf{x}$ , in blocks, is called the “frame length”, and is denoted by  $J$ . The transpose and conjugate transpose are denoted by “ $T$ ” and “ $H$ ”, respectively. An identity matrix

with dimension  $n$  is denoted by  $\mathbf{I}_n$  and sometimes the subscript is omitted for simplicity.  $[\cdot]_{l,j}$  denotes the  $(l, j)$ -th entry of the matrix in the bracket. The operator  $\otimes$  denotes the Kronecker product. The operation  $\text{diag}\{\cdot\}$  returns a diagonal matrix with the elements in the brace ordered on the main diagonal, whilst  $(\cdot)^{\text{diag}}$  returns a diagonal matrix that only contains the diagonal part of the matrix in the parentheses. The operator  $E(\cdot)$  is the expectation with respect to the joint pdf of  $\mathbf{x}$  and  $\boldsymbol{\eta}$ .

## II. CHANNEL MODELS

### A. System Model

Consider a coded system, as shown in Fig.1, characterized by the following linear matrix equation

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

where  $\mathbf{r}$  is an observation vector,  $\mathbf{x}$  a transmitted signal vector,  $\mathbf{H}$  a channel matrix and  $\boldsymbol{\eta}$  an additive white Gaussian noise (AWGN) vector with zero mean and covariance  $\sigma^2\mathbf{I}$ . We assume that  $\mathbf{x}$  is generated by an encoding device (denoted by ENC in Fig.1) using forward-error-control (FEC) codes and permuted by an interleaver (marked by  $\Pi$  in Fig.1). Typical examples of (1) include ISI, MIMO and MAC channels.

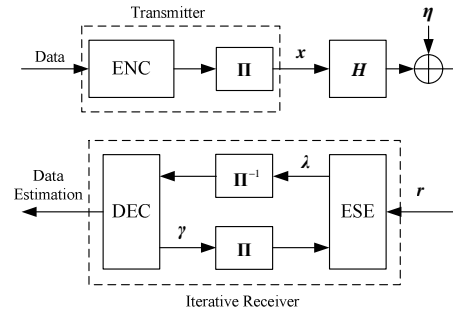


Fig. 1. The general transmitter and (iterative) receiver structures for a coded linear system with channel input  $\mathbf{x}$ ,  $\Pi$  and  $\Pi^{-1}$  denoting the interleaver and the corresponding de-interleaver, respectively.

### B. Circulant Systems

The channel matrix  $\mathbf{H}$  in (1) is a  $J \times J$  circulant matrix if  $[\mathbf{H}]_{i,j} = [\mathbf{H}]_{r,j}$  for any  $j - l = (j' - l')$  modulo  $J$ . A circulant  $\mathbf{H}$  can be represented as

$$\mathbf{H} = \begin{bmatrix} h_0 & \cdots & h_2 & h_1 \\ h_1 & h_0 & & h_2 \\ \vdots & \ddots & \ddots & \vdots \\ h_{J-1} & \cdots & h_1 & h_0 \end{bmatrix}. \quad (2)$$

Let  $r_j$  and  $x_j$  (the  $j$ th entry in  $\mathbf{r}$  and  $\mathbf{x}$ , respectively) be the received and transmitted signal at time instant  $j$ , respectively. The matrix  $\mathbf{H}$  in (2) is realizable in an ISI channel (with channel coefficients  $[h_0, \dots, h_{L-1}]$  where  $L$  denotes

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 channel memory length) by padding  $\mathbf{x}$  with a cyclic prefix (CP) [9] that consists of the last  $L-1$  entries of  $\mathbf{x}$ .

Let  $\mathbf{F}$  be the normalized discrete Fourier transform (DFT) matrix with the  $(j, l)$ -th entry given by  $J^{-1/2} \exp(-i2\pi jl/J)$ , where  $i = \sqrt{-1}$ . Let  $\{g_j\}$  be the DFT of  $\{h_j\}$  (see (2)), i.e.,

$$g_j = J^{-1/2} \sum_{l=0}^{J-1} h_l \exp(-i2\pi jl/J) \quad j = 0, \dots, J-1.$$

The following properties are well known for a circulant  $\mathbf{H}$ .

- $\mathbf{G} \equiv \mathbf{F}\mathbf{H}\mathbf{F}^H$  is diagonal;
- $\mathbf{G} = J^{1/2} \text{diag}\{g_0, g_1, \dots, g_{J-1}\}$ ;
- $\mathbf{H} = \mathbf{F}^H \mathbf{G} \mathbf{F}$ .

### C. Block Circulant Systems

A block circulant system, as a natural extension of the scalar circulant one in II.B, can be represented by

$$\begin{bmatrix} \mathbf{r}_0 \\ \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_{J-1} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_0 & \cdots & \mathbf{H}_2 & \mathbf{H}_1 \\ \mathbf{H}_1 & \mathbf{H}_0 & & \mathbf{H}_2 \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{H}_{J-1} & \cdots & \mathbf{H}_1 & \mathbf{H}_0 \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_{J-1} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\eta}_0 \\ \boldsymbol{\eta}_1 \\ \vdots \\ \boldsymbol{\eta}_{J-1} \end{bmatrix}, \quad (3a)$$

where  $\mathbf{H}$  is block circulant (i.e., each block-column is a cyclic shift of its previous block-column by one block), each  $\mathbf{H}_l$  is an  $M \times N$  sub-matrix,  $\mathbf{r}_j$  (or  $\boldsymbol{\eta}_j$ ) an  $M \times 1$  sub-vector, and  $\mathbf{x}_j$  an  $N \times 1$  sub-vector. A typical example of such a system is a MIMO ISI channel with CP padding to each transmit antenna. Let  $L$  be the channel memory length, and  $N$  (resp.,  $M$ ) be the number of transmit (resp., receive) antennas. Thus, when  $l = L, \dots, J-1$ ,  $\mathbf{H}_l = \mathbf{0}$ ; and when  $l = 0, \dots, L-1$ ,

$$[\mathbf{H}_l]_{mn} = h_{l,m,n}, \text{ for } m = 1, \dots, M, n = 1, \dots, N, \quad (3b)$$

where  $h_{l,m,n}$  denotes the  $l$ th tap coefficient between the  $n$ th transmit antenna and the  $m$ th receive antenna. The sub-vectors  $\mathbf{r}_j$ ,  $\mathbf{x}_j$  and  $\boldsymbol{\eta}_j$  are the corresponding sub-blocks in  $\mathbf{r}$ ,  $\mathbf{x}$ , and  $\boldsymbol{\eta}$  at time instant  $j$ , respectively.

Let  $\{\mathbf{G}_0, \mathbf{G}_1, \dots, \mathbf{G}_{J-1}\}$  be blocks of the same size as  $\mathbf{H}_l$ ,  $\forall l$ , as given in (3). Define the block-wise DFT of  $\{\mathbf{H}_l\}$  as  $\mathbf{G}_j = J^{-1/2} \sum_{l=0}^{J-1} \mathbf{H}_l \exp(-i2\pi jl/J)$  for  $j = 0, \dots, J-1$ .

Denote  $\mathbf{F}_M \equiv \mathbf{F} \otimes \mathbf{I}_M$ , and  $\mathbf{F}_N \equiv \mathbf{F} \otimes \mathbf{I}_N$ . Similarly to the case for a scalar circulant matrix in II.B, the following properties hold for a block circulant  $\mathbf{H}$ :

- $\mathbf{G} \equiv \mathbf{F}_M \mathbf{H} \mathbf{F}_N^H$  is block diagonal;
- $\mathbf{G} = J^{1/2} \text{diag}\{\mathbf{G}_0, \mathbf{G}_1, \dots, \mathbf{G}_{J-1}\}$ ;
- $\mathbf{H} = \mathbf{F}_M^H \mathbf{G} \mathbf{F}_N$ .

### D. Real Representation of a Complex System

The discussion in II.B and C applies to both real and complex channel models, since a complex linear system can be equivalently transformed into a real system (with doubled dimensions) by equating the real and imaginary parts separately. Hence, we only discuss real systems in this paper.

For simplicity, we assume that each element in  $\mathbf{x}$  is modulated by binary phase shift keying (BPSK) over  $\{+1, -1\}$ . Note that BPSK modulation for the converted real system model is equivalent to quadrature phase shift keying (QPSK) with Gray mapping for the original system.

## III. THE ITERATIVE LMMSE-APP DETECTION PRINCIPLES

The lower part of Fig.1 shows the iterative LMMSE-APP receiver. The ESE (for “elementary signal estimator”)

module is based on the LMMSE principle and the DEC based on the APP decoding principle, hence the name “LMMSE-APP”. These two modules are connected by the interleaver  $\boldsymbol{\Pi}$  and the corresponding de-interleaver  $\boldsymbol{\Pi}^{-1}$ , and work iteratively. Note that  $\mathbf{H}$  is assumed to be known at the receiver.

### A. The LMMSE Approach to the ESE

The ESE computes the extrinsic log-likelihood ratio (LLR) for each  $x_j$  as

$$\ln \frac{p(\mathbf{r} | x_j = +1)}{p(\mathbf{r} | x_j = -1)}, \quad j = 1, 2, \dots \quad (4)$$

with the FEC coding constraint ignored, i.e., the ESE operates as if  $\mathbf{x}$  contains un-coded bits. The exact evaluation of (4) can be realized by the MAP algorithm, but is usually prohibitively complicated. Let the diagonal matrix  $\mathbf{V}$  be the covariance of  $\mathbf{x}$ . It is shown in [4] that the LLR in (4) can be approximated by the following LMMSE estimator

$$\lambda_j = 2\mathbf{h}_j^T \mathbf{R}_j^{-1} (\mathbf{r} - \mathbf{H}\mathbf{E}(\mathbf{x}) + \mathbf{h}_j \mathbf{E}(x_j)), \quad j = 1, 2, \dots \quad (5)$$

where  $\mathbf{R}_j \equiv \mathbf{R} - v_j \mathbf{h}_j \mathbf{h}_j^T$ ,  $\mathbf{R} \equiv \mathbf{H}\mathbf{V}\mathbf{H}^T + \sigma^2 \mathbf{I}$ ,  $\mathbf{h}_j$  is the  $j$ th column of  $\mathbf{H}$ , and  $v_j$  is the variance of  $x_j$ . We adopt initial conditions  $\mathbf{E}(\mathbf{x}) = \mathbf{0}$  and  $\mathbf{V} = \mathbf{I}$  (meaning no *a priori* information). During the iterative process,  $\mathbf{E}(\mathbf{x})$  and  $\mathbf{V}$  are updated by using the feedback information from the DEC afterwards, as shown below.

### B. The DEC Operations

The DEC performs APP decoding using  $\boldsymbol{\lambda} \equiv [\lambda_0, \lambda_1, \dots, \lambda_{J-1}]^T$  as its inputs. The DEC outputs are the extrinsic LLRs given by

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

### C. The Overall Iterative Process

After the DEC operations, the ESE operations can be executed again to refine the estimates in (4) using the feedback  $\boldsymbol{\gamma} \equiv [\gamma_0, \gamma_1, \dots, \gamma_{J-1}]^T$  from the DEC. Since  $x_j$  is taken over  $\{+1, -1\}$ , we have

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

This process continues iteratively until the algorithm converges. Refer to [1-3] for details.

## IV. A FAST IMPLEMENTATION TECHNIQUE FOR THE ESE

Since the APP decoding in the DEC has been well studied, we focus on the implementation of the ESE here.

### A. ESE for Circulant $\mathbf{H}$

Applying the matrix inversion lemma to  $\mathbf{R}_j^{-1}$ , we can rewrite (5) as

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

or in a vector form as

$$\boldsymbol{\lambda} = 2(\mathbf{I} - \mathbf{V}(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})_{\text{diag}})^{-1} [\mathbf{H}^T \mathbf{R}^{-1} (\mathbf{r} - \mathbf{H}\mathbf{E}(\mathbf{x})) - (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})_{\text{diag}} \mathbf{E}(\mathbf{x})]. \quad (6)$$

Let  $\bar{v}$  be the average of  $\{v_j\}$ , i.e.,  $\bar{v} \equiv J^{-1} \sum_{j=0}^{J-1} v_j$  [2].

We update  $\mathbf{V}$  as follows.

$$\mathbf{V} \approx \bar{v} \mathbf{I} \quad (7)$$

Recall that  $\mathbf{V}$  is the variance of  $\mathbf{x}$ . Eqn. (7) implies:

- (i)  $x_i$  and  $x_j$  are uncorrelated if  $i \neq j$ , which is approximately ensured by the interleaver  $\Pi$ ; and
- (ii) All  $\{x_j\}$  have the same variance  $\bar{v}$ .

Note that (ii) is not an optimal treatment since we actually know the individual variances  $\{x_j\}$ . However, this sub-optimal treatment leads to greatly reduced complexity. Based on the properties given in II.B, we have

$$\begin{aligned} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} &= \mathbf{F}^H \mathbf{G}^H \mathbf{F} (\bar{v} \mathbf{F}^H \mathbf{G}^H \mathbf{F} \mathbf{F}^H \mathbf{G}^H \mathbf{F} + \sigma^2 \mathbf{I})^{-1} \mathbf{F}^H \mathbf{G}^H \mathbf{F} \\ &= \mathbf{F}^H \mathbf{G}^H (\bar{v} \mathbf{G} \mathbf{G}^H + \sigma^2 \mathbf{I})^{-1} \mathbf{G} \mathbf{F} \end{aligned}$$

Since  $\mathbf{G}^H (\bar{v} \mathbf{G} \mathbf{G}^H + \sigma^2 \mathbf{I})^{-1} \mathbf{G}$  is diagonal,  $\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}$  is circulant. Further, we have

$$(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})_{\text{diag}} = u \mathbf{I} \quad (8)$$

where  $u \equiv \sum_{j=0}^{J-1} |g_j|^2 (\bar{v} |g_j|^2 + \sigma^2)^{-1}$ . Then we rewrite (6) as  $\lambda = 2(1 - \bar{v}u)^{-1} (\mathbf{F}^H \mathbf{G}^H (\bar{v} \mathbf{G} \mathbf{G}^H + \sigma^2 \mathbf{I})^{-1} \mathbf{F} (\mathbf{r} - \mathbf{H}\mathbf{E}(\mathbf{x})) + u\mathbf{E}(\mathbf{x}))$  (9)

The ESE can be implemented based on (9) as follows:

- Step 1: Take the FFT of  $\mathbf{r} - \mathbf{H}\mathbf{E}(\mathbf{x})$ .
- Step 2: Multiply the results of Step 1 by a diagonal matrix  $\mathbf{G}^H (\bar{v} \mathbf{G} \mathbf{G}^H + \sigma^2 \mathbf{I})^{-1}$ .
- Step 3: Take the IFFT of the results in Step 2, then add the weighted *a priori* mean  $u\mathbf{E}(\mathbf{x})$ , and finally scale the result by  $2(1 - \bar{v}u)^{-1}$ .

The overall complexity of the above method is roughly  $O(\log_2 J)$  per entry of  $\mathbf{x}$ . Note that multiplying  $\mathbf{G}^H (\bar{v} \mathbf{G} \mathbf{G}^H + \sigma^2 \mathbf{I})^{-1}$  can be regarded as symbol-by-symbol equalization in the frequency domain. Therefore the above method is actually equivalent to the FDE method [11]. We refer to this method as frequency-domain-equalization MMSE (FDE-MMSE). For convenience of comparison, we refer to the implementation technique in [1, 2 and 4] as time-domain-equalization MMSE (TDE-MMSE). One major advantage of FDE-MMSE is that its complexity is independent of the channel memory length  $L$ , whilst that of TDE-MMSE is proportional to  $L^2$  in general.

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

Consider the block circulant  $\mathbf{H}$  in (3). Denote by  $\mathbf{V}_j$  the covariance matrix of  $\mathbf{x}_j$  (which is an  $N \times N$  diagonal matrix). Similarly to the scalar case, we approximate  $\mathbf{V}$  by

$$\mathbf{V} \approx \mathbf{I}_J \otimes \bar{\mathbf{V}} \quad (10)$$

where  $\bar{\mathbf{V}}$  is the average of  $\mathbf{V}_j$  for  $j = 0, \dots, J-1$ . Based on the properties given in II.C, it can be shown that

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

where  $\mathbf{U} \equiv (\mathbf{I}_J \otimes \mathbf{U}_{\text{sub}})_{\text{diag}}$ ,  $\mathbf{U}_{\text{sub}} \equiv \sum_{j=0}^{J-1} \mathbf{G}_j^H (\mathbf{J} \mathbf{G}_j \bar{\mathbf{V}} \mathbf{G}_j^H + \sigma^2 \mathbf{I}_M)^{-1} \mathbf{G}_j$ .

The operations in (11) can be implemented similarly to those in IV.A except that block-wise FFT should be used.

## V. EVOLUTION ANALYSIS

In this section, we outline an evolution-based technique to characterize the behavior of the iterative LMMSE-APP detector discussed previously.

### A. Evolution Analysis for Circulant $\mathbf{H}$

The evolution technique tracks a few parameters in the iterative process using pre-calculated transfer functions. We require that the parameters involved be as few as possible so as to reduce the complexity involved. At the same

time we require that these parameters accurately characterize the statistical behavior of the iterative process, so as to predict its performance. For a system with a circulant channel matrix as in (2), we find that it is most convenient to use the input variance (denoted by  $\bar{v}$ ) and output signal-to-noise ratio (denoted by  $\rho$ ) to characterize the ESE and the input SNR  $\rho$  and output  $\bar{v}$  to characterize the DEC. Thus, the density evolution of the iterative decoding process reduces to a simple recursion between  $\rho$  and  $\bar{v}$ , and the transfer functions of the ESE and the DEC, respectively, can be denoted by

$$\rho = \phi(\bar{v}) \quad \text{and} \quad \bar{v} = \psi(\rho).$$

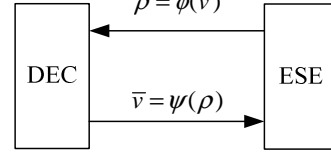


Fig. 2. An illustration of the evolution process.

Similarly to the generation of EXIT functions [8], 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 iterative process, 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 process is as follows

After the first iteration:  $\rho = \phi(1)$   
 After the second iteration:  $\rho = \phi(\psi(\phi(1)))$   
 ...

The SNR value resulting from a specified number of iterations can be used to predict the system performance.

The above principle is similar to the EXIT chart method in which mutual information is used. The EXIT chart approach [2] is for fixed channels in which the transfer functions can be pre-calculated by simulation and stored. This strategy, however, is not applicable if we want to evaluate the statistical system behavior in quasi-static fading channels, since the transfer functions involved (in particular,  $\phi(\cdot)$ ) can be different for different channel realizations. It is not practical to pre-calculate and store  $\phi(\cdot)$  and  $\psi(\cdot)$  for all possible channel realizations.

In the following, we present a solution to this difficulty that is obtained by making certain approximations. It is shown via numerical results that the proposed method can accurately characterize the behavior of the iterative receiver in Fig. 1.

First, we derive an analytical form for  $\phi(\cdot)$ . This is basically to find the output SNR  $\rho$  for the ESE with a given  $\bar{v}$ .

We rewrite  $\lambda_j$  in (5) in a signal-plus-distortion form as

$$\lambda_j = \mu x_j + \zeta_j \quad (12)$$

where  $\mu \equiv 2\mathbf{h}_j^T \bar{\mathbf{R}}_j^{-1} \mathbf{h}_j$ ,

$$\zeta_j \equiv 2\mathbf{h}_j^T \bar{\mathbf{R}}_j^{-1} (\mathbf{r} - \mathbf{H}\mathbf{E}(\mathbf{x}) - \mathbf{h}_j(x_j - \mathbf{E}(x_j))),$$

$$\bar{\mathbf{R}}_j = \bar{v} \mathbf{H} \mathbf{H}^T - \bar{v} \mathbf{h}_j \mathbf{h}_j^T + \sigma^2 \mathbf{I}.$$

Note that  $\zeta_j$  is not a function of  $x_j$  (since  $\mathbf{r} - \mathbf{h}_j x_j$  is not a function of  $x_j$ ). Also,  $\mu$  is invariant to index  $j$  since

$$\mu = 2\mathbf{h}_j^T \bar{\mathbf{R}}_j^{-1} \mathbf{h}_j = \frac{2\mathbf{h}_j^T \mathbf{R}^{-1} \mathbf{h}_j}{1 - \bar{v} \mathbf{h}_j^T \mathbf{R}^{-1} \mathbf{h}_j} = \frac{2u}{1 - \bar{v}u},$$

where the second equality follows from the matrix inversion lemma, and the last equality follows from (8). Furthermore, it can be verified that the mean of  $\zeta_j$  is zero and its variance given by  $E(\zeta_j^2) = 4\mathbf{h}_j^T \bar{\mathbf{R}}_j^{-1} \mathbf{R}_j \bar{\mathbf{R}}_j^{-1} \mathbf{h}_j$ . The average variance of  $\zeta_j$ , denoted by  $\sigma_\zeta^2$ , is given by

$$\begin{aligned} \sigma_\zeta^2 &\equiv J^{-1} \sum_{j=0}^{J-1} E(\zeta_j^2) = 4J^{-1} \sum_{j=0}^{J-1} \mathbf{h}_j^T \bar{\mathbf{R}}_j^{-1} \mathbf{R}_j \bar{\mathbf{R}}_j^{-1} \mathbf{h}_j \\ &= 4J^{-1} \left( \sum_{\substack{j=0 \\ j \neq l}}^{J-1} v_l \mathbf{h}_l^T \bar{\mathbf{R}}_l^{-1} \mathbf{h}_l \mathbf{h}_l^T \bar{\mathbf{R}}_l^{-1} \mathbf{h}_l + \sum_{j=0}^{J-1} \sigma^2 \mathbf{h}_j^T \bar{\mathbf{R}}_j^{-1} \bar{\mathbf{R}}_j^{-1} \mathbf{h}_j \right) \\ &= 4J^{-1} \left( \sum_{l=0}^{J-1} v_l \sum_{j \neq l} (\mathbf{h}_j^T \bar{\mathbf{R}}_j^{-1} \mathbf{h}_l)^2 + \sum_{j=0}^{J-1} \sigma^2 \mathbf{h}_j^T \bar{\mathbf{R}}_j^{-1} \bar{\mathbf{R}}_j^{-1} \mathbf{h}_j \right) \\ &= 4J^{-1} \sum_{j=0}^{J-1} \frac{v_j}{\bar{v}} \mathbf{h}_j^T \bar{\mathbf{R}}_j^{-1} (\bar{v} \mathbf{H} \mathbf{H}^T - \bar{v} \mathbf{h}_j \mathbf{h}_j^T + \sigma^2 \mathbf{I}) \bar{\mathbf{R}}_j^{-1} \mathbf{h}_j \\ &= 4J^{-1} \sum_{j=0}^{J-1} \frac{v_j}{\bar{v}} \mathbf{h}_j^T \bar{\mathbf{R}}_j^{-1} \mathbf{h}_j \\ &= \frac{4u}{1 - \bar{v}u}. \end{aligned}$$

It is known that the residue distortion at the output of the LMMSE estimator is approximately Gaussian distributed [10]. Thus, we introduce the following assumption.

**Assumption I:**  $\zeta_j$  is a Gaussian random variable with zero mean and variance  $\sigma_\zeta^2$ .

Under the above assumption, the ESE output in (6) can be modeled as an observation of  $x_j$  scaled by  $\mu$  (a constant) and corrupted by AWGN samples  $\{\zeta_j\}$ . A single SNR, given by

$$\rho \equiv \mu^2 / \sigma_\zeta^2 = u(1 - \bar{v}u)^{-1} \quad (13)$$

is sufficient to characterize this model. From (8),  $u$  is a function of  $\bar{v}$ . Thus (13) provides an analytical form for  $\rho = \phi(\bar{v})$ . With (13), we can instantaneously evaluate  $\phi(\cdot)$  at each step instead of pre-calculating and storing  $\phi(\cdot)$  for all possible channel realizations.

The generation of  $\psi(\cdot)$  is straightforward. Recall that  $\{\lambda_j\}$  are interleaved and delivered to DEC as its inputs. Based on (12),  $\{\lambda_j\}$  can be modeled as the observation of  $\{x_j\}$  through an AWGN channel. (The correlation among  $\{\lambda_j\}$  can be ignored due to interleaving.) Thus, we can generate  $\psi(\cdot)$  by simulating the DEC in memoryless AWGN channels with various SNRs.

### B. Evolution Analysis for Block Circulant $\mathbf{H}$

The above evolution analysis can be generalized to block systems. However, we found that the method becomes inaccurate if the ESE outputs are still characterized by a single SNR for MIMO systems. This is caused by the fact that the SNR values of the ESE outputs in a MIMO system can vary significantly because the signals transmitted via different MIMO inputs may experience different channel fading. It therefore becomes too inaccurate to use a single parameter for all of the ESE outputs. However, we observed that the following approaches are valid.

- The ESE outputs corresponding to different time instants but the same MIMO input can be characterized by a single SNR value. (Note: A MIMO input is indexed by  $n$ . See (3b).)
- If there is only one decoder in the system, then the outputs of this decoder can be characterized by a single variance value. However, if there are multiple decoders, such as for a multiple access system, multiple variance values should be used.

We will use a vector form SNR denoted by  $\boldsymbol{\rho}$ .  $\rho_n$ , denoting the  $n$ th entry in  $\boldsymbol{\rho}$ , represents the SNR value at the ESE outputs corresponding to a specific index  $n$  of the MIMO inputs. For simplicity, we assume that there is only one decoder and so one variance value is used, i.e.,  $\bar{\mathbf{V}} = \bar{v} \mathbf{I}$ . Thus we need to find the following two functions, namely  $\boldsymbol{\rho} = \phi(\bar{v})$  and  $\bar{v} = \psi(\boldsymbol{\rho})$ .

For  $\mathbf{H}$  given in (3), it can be shown that

$$\boldsymbol{\rho} = (\mathbf{U}_{sub})_{diag} (\mathbf{I} - \bar{\mathbf{V}} (\mathbf{U}_{sub})_{diag})^{-1} \mathbf{1} \equiv \phi(\bar{v}), \quad (14)$$

where  $\mathbf{1}$  is an all-one vector with length  $N$ . The derivation for (14) is similar to (13). We omit details here due to the lack of space.

Next we consider the DEC. We divide the entries of  $\boldsymbol{\lambda}$  into  $N$  groups, each characterized by an entry in  $\boldsymbol{\rho}$ . We can similarly model each  $\lambda_j$  as a scaled observation of  $x_j$  corrupted by an additive Gaussian noise, i.e.,

$$\lambda_j = \mu_j x_j + \zeta_j. \quad (15)$$

However, the scaling factor  $\mu_j$  now has  $N$  different values, as specified by  $\boldsymbol{\rho}$ . We refer to (15) as an  $N$ -state fading channel. Thus the inputs to the DEC can be treated as an LLR sequence coming from a memoryless  $N$ -state fading channel, and the corresponding transfer function can be obtained by simulating the DEC over different memoryless  $N$ -state fading channels. Note that the set of  $N$  values for  $\{\mu_j\}$  are different for different channel realizations. We need to simulate and store all possible combinations, and thus an  $N$ -dimensional table is required. For convenience, we refer to this method as the full-table (FT) method.

The FT method is impractical even for a moderate  $N$ . The following low-complexity approach is inspired by the rationale behind the EXIT chart analysis [8], i.e., given an LLR sequence  $\boldsymbol{\lambda}$  as the input of a turbo component (the ESE or the DEC), its output behavior will not be affected by replacing  $\boldsymbol{\lambda}$  with another sequence  $\boldsymbol{\lambda}'$  that contains the same amount of mutual information (with respect to the transmitted sequence  $\mathbf{x}$ ) as  $\boldsymbol{\lambda}$  does. We write this as an assumption.

**Assumption II:** The decoder behavior over a channel modeled in (15) remains the same if the amount of mutual information with respect to  $\{x_j\}$  contained in  $\{\lambda_j\}$  is a constant.

According to the analysis in [8], 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

$$I(\lambda; x) = \frac{1}{\sqrt{8\pi\rho}} \int_{-\infty}^{+\infty} \left( \log \frac{1}{1+e^{-\lambda}} \right) \cdot e^{-\frac{(\lambda-2\rho)^2}{8\rho}} d\lambda \equiv f(\rho). \quad (16)$$

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

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

We then let  $\psi(\boldsymbol{\rho}) \approx \psi_{AWGN}(\rho_{eff})$ , based on Assumption II.

Note that both  $f(\cdot)$  and  $\psi_{AWGN}(\cdot)$  can be pre-calculated and stored using one-dimensional tables. We refer to this

method as the equivalent-mutual-information (EMI) technique.

## VI. NUMERICAL RESULTS

In this section, numerical results are given to verify the proposed implementation method and the evolution analysis. Although the generic system model in Fig. 1 includes a variety of communication channels, we restrict attention to MIMO ISI channels due to space limitation.

Consider quasi-static Rayleigh fading  $M \times N$  MIMO ISI channels. The ENC contains only one encoder employing a rate-1/2  $(7, 5)_8$  convolutional code with information length 1024. The coded bits are BPSK modulated, and then are randomly interleaved and separated into two streams to form the real and imaginary part of the transmit symbol vector. Then, this vector is separated into  $N$  streams for simultaneous transmissions over  $N$  transmit antennas. A typical implementation of this MIMO system can also be found in [4].

The quasi-static Rayleigh fading channels are modeled as follows: the fading coefficients are independently identically distributed complex Gaussian random variables with zero-mean, and the average total fading gain normalized to 1, i.e.,

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

where the expectation is taken over the channel fading distribution. Thus the bit-energy-to-noise-density ratio is calculated as

$$\frac{E_b}{N_0} \equiv \frac{E(|x_j|^2) \cdot E_{\mathbf{h}}}{R_c N_0} = \frac{1}{2R_c \sigma^2},$$

where  $E_b$  denotes the bit energy,  $N_0$  denotes the single-sideband noise power spectral density, and  $R_c$  denotes the rate of  $x_j$ .

In simulation, at least 10 iterations are taken in the detection to guarantee convergence. For each point on the performance curve, at least 1000 errors are collected.

Fig. 3 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 with FDE-MMSE (due to the approximation in (10)) is not significant. Fig. 3 also includes the performance curves predicted by evolution. In Fig. 3, it can be seen that the performance of FDE-MMSE is accurately predicted by the full-table method (denoted by "FT-Evolution"). Moreover, the prediction made by the EMI method (denoted by "EMI-Evolution") deviates only about 0.1-0.2dB away from the actual performance.

In Fig. 4, we consider a quasi-static Rayleigh-fading  $4 \times 4$  MIMO ISI channels with  $L = 2$  and 4, respectively. The evolution results are based on the EMI method 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 are increased.

## REFERENCES

- [1] 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.
- [2] M. Tüchler, R. Kowtner, A. C. Singer, "Turbo equalization: Principles and new results", *IEEE Trans. Commun.*, vol. 50, pp. 754–767, May 2002.

- [3] L. Liu, Li Ping, and W. K. Leung, "Iterative detection for inter-leaver division multiple access in channels with intersymbol interference", *IEICE Trans. Commun.*, vol. E87-B, no. 11, pp. 3274–3280, Nov. 2004.
- [4] 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.
- [5] C. Douillard et al., "Iterative correction of intersymbol interference: turbo equalization," *Eur. Trans. Telecommun.*, vol. 6, pp. 507–511, Sept.–Oct., 1995.
- [6] 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.
- [7] 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).
- [8] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes", *IEEE Trans. Commun.*, vol. 49, no. 10, Oct. 2001.
- [9] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eldson, "Frequency Domain Equalization for Single-Carrier Broadband Wireless Systems". *IEEE Commun. Mag.*, pages 58–66, April 2002.
- [10] V. Poor and S. Verdú, "Probability of error in MMSE Multiuser detection," *IEEE Trans. Inform. Theory*, vol. IT-43, pp. 835–847, May 1997.
- [11] M. Tüchler, and J. Hagenauer, "Turbo equalization using frequency domain equalizers," in *Proc. Of Allerton Conference*, Monticello, IL, USA, Oct. 2000.

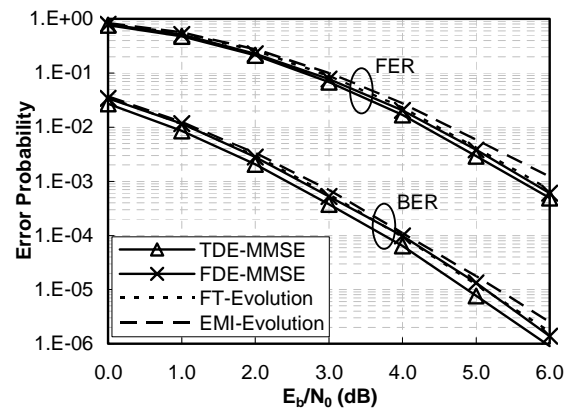


Fig. 3. BER/FER performances and evolution predictions on quasi-static Rayleigh fading  $2 \times 2$  MIMO ISI channels with  $L = 4$ .

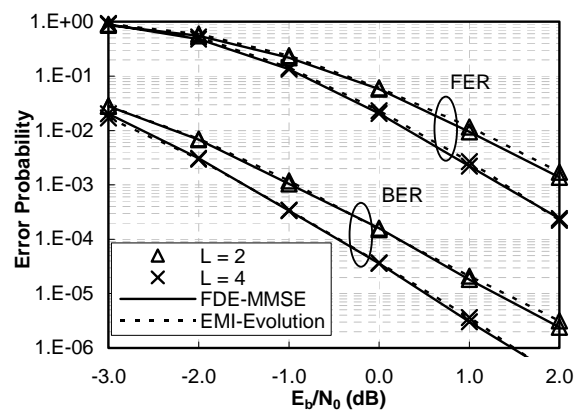


Fig. 4. BER/FER performances and evolution predictions on quasi-static Rayleigh fading  $4 \times 4$  MIMO ISI channels with different  $L$ .