

# Performance Analysis of Multi-ary Systems with Iterative Linear Minimum-Mean-Square-Error Detection

Li Ping, Jun Tong, Xiaojun Yuan and Qinghua Guo

Department of Electronic Engineering, City University of Hong Kong, Hong Kong SAR, China,  
Telephone: (852) 2788-9574 E-mail: eeliping@cityu.edu.hk

**Abstract**—This paper is concerned with coded multi-ary systems over linear channels. Based on a semi-analytical evolution technique, the impact of signaling schemes on the performance of low-cost iterative linear minimum-mean-square-error (LMMSE) detection is studied. It is shown that superposition coded modulation (SCM) maximizes the output signal-to-noise ratio (SNR) of LMMSE detectors. Consequently, SCM may potentially outperform other conventional signaling schemes when LMMSE detectors are used. Numerical examples are provided to verify the theoretical analysis.

## I. INTRODUCTION

Iteratively decoded bit-interleaved coded modulation (BICM-ID) with multi-ary signaling is an attractive scheme for high-rate transmissions. Its performance depends heavily on the signaling schemes employed [1]. Signaling design for BICM-ID has been extensively studied for memoryless scalar channels [1]-[3], where the optimal maximum *a posteriori* (MAP) detector is usually assumed. However, as we will show below, careful study is still required when a linear minimum-mean-square-error (LMMSE) detector [4]-[11] is used at the receiver.

In this paper, we establish a connection between signaling schemes and the iterative LMMSE detection performance. Maximizing the signal-to-noise ratio (SNR) of the LMMSE detector outputs is adopted as the criterion for designing signaling schemes. Under this criterion, we demonstrate the advantages of superposition coded modulation (SCM) [12], [13] over other traditional signaling schemes. In addition, we show that quadrature amplitude modulation (QAM) with Gray mapping, which was regarded as a “poor” option for BICM-ID, turns out to be advantageous under the SNR maximization criterion. This implies that, interestingly, although Gray mapping may result in relatively poor performance in memoryless scalar channels where the high-cost MAP detection is affordable, it may be a good option in more complicated channels where low-cost iterative LMMSE detection has to be used, since other options (such as the modified set-partitioning (MSP) signaling [1]) may not function well in the latter case. The analytical study is confirmed by simulation examples.

In this paper, lower case letters (e.g.,  $x$ ) denote scalars, bold lower case letters (e.g.,  $\mathbf{x}$ ) denote column vectors, bold upper case letters (e.g.,  $\mathbf{X}$ ) denote matrices, and  $\mathbf{I}$  denote the identity matrix with proper size. The superscript “ $T$ ”,

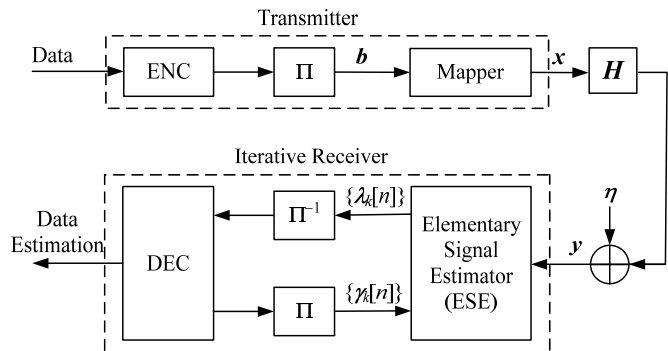


Fig. 1. The transmitter and receiver structure of a multi-ary system over linear channels.  $\Pi$  denotes the interleaver and  $\Pi^{-1}$  the de-interleaver.

“ $H$ ” and “ $-1$ ” denote the transpose, conjugate transpose and inverse operations, respectively.

## II. ITERATIVE DETECTION

### A. System Model

The transmitter scheme follows the principles of BICM-ID [1], as shown in Fig. 1. The source data is first encoded by the encoder (ENC) using a binary forward-error-control (FEC) code, and permuted by a random interleaver (marked by  $\Pi$ ) to produce a bit sequence  $\mathbf{b}$ . Then  $\mathbf{b}$  is segmented into  $N$  sub-blocks  $\mathbf{b} = \{b[0], b[1], \dots, b[N-1]\}$  where each  $b[n]$  is a sub-block of  $K$  bits  $b[n] = \{b_1[n], b_2[n], \dots, b_K[n]\}$ . The mapper then maps each  $b[n]$  onto a signaling point  $x[n]$  in a constellation  $\mathcal{S}$  of size  $2^K$ . Denote by  $\mathcal{B}$  the set of  $b[n]$  and by  $\Phi$  the mapping rule from  $\mathcal{B}$  to  $\mathcal{S}$ .

Let matrix  $\mathbf{H}$  represent the multiplicative effect of the linear channel. The received signal is given by

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

where  $\mathbf{x} = [x[0], x[1], \dots, x[N-1]]^T$  is the transmitted signal vector and  $\boldsymbol{\eta}$  is a vector of additive white Gaussian noise (AWGN) with mean vector  $\mathbf{0}$  and covariance matrix  $\sigma^2\mathbf{I}$ . Note that the generic model (1) can represent several different types of channels, such as the multipath channel, multiple-input multiple-output (MIMO) channel, and multiple access channel (MAC) [9]. We always assume that  $\mathbf{H}$  is known perfectly at the receiver.

### B. Overall Iterative Detection Principles

The generic iterative receiver structure is shown in the lower part of Fig. 1. The elementary signal estimator (ESE) computes the *extrinsic* log-likelihood ratio (LLR)  $\lambda_k[n]$  for each  $b_k[n]$

$$\lambda_k[n] \equiv \ln \left( \frac{\Pr(b_k[n] = 0 | \mathbf{y})}{\Pr(b_k[n] = 1 | \mathbf{y})} \right) - \ln \left( \frac{\Pr(b_k[n] = 0)}{\Pr(b_k[n] = 1)} \right) \quad (2)$$

with the FEC coding constraint ignored, i.e., the ESE operates as if  $b_k[n]$  is an un-coded bit. The decoder (DEC) performs the *a posteriori* probability (APP) decoding using  $\{\lambda_k[n]\}$  as inputs, producing the *extrinsic* LLRs

$$\gamma_k[n] \equiv \ln \left( \frac{\Pr(b_k[n] = 0 | \{\lambda_k[n]\})}{\Pr(b_k[n] = 1 | \{\lambda_k[n]\})} \right) - \ln \left( \frac{\Pr(b_k[n] = 0)}{\Pr(b_k[n] = 1)} \right). \quad (3)$$

After decoding, the ESE operations can be executed again to refine the estimates  $\{\lambda_k[n]\}$  using the feedbacks  $\{\gamma_k[n]\}$ . This process continues iteratively after a preset number of iterations. Hard decisions are made after the final iteration to produce the data estimates. Related discussions on such iterative detection process can be found in [4]-[9]. Since the APP decoding is a standard function, we focus on the ESE function in what follows.

### C. ESE Function

The optimal solution of the ESE function in (2) is usually prohibitively high, since, after the linear transform  $\mathbf{H}$ , the constellation of  $\mathbf{H}\mathbf{x}$  is usually significantly expanded. The iterative LMMSE detector is a low-cost, sub-optimal alternative. The detection process in the ESE can be divided into three steps as below.

*Step 1. Soft Mapping:* In this step, the means  $\{E[x[n]]\}$  and variances  $\{\text{Var}[x[n]]\}$  of the entries of  $\mathbf{x}$  are generated using the feedback LLRs  $\{\gamma_k[n]\}$  from the DEC. This is a preparation stage for the LMMSE estimation. We will discuss the details involved in this step in the next subsection.

*Step 2. LMMSE Estimation:* Define the mean of  $\mathbf{x}$  as

$$E[\mathbf{x}] = [E[x[0]], E[x[1]], \dots, E[x[N-1]]]^T. \quad (4)$$

Following [9], we assume that the entries of  $\mathbf{x}$  are independent, and the covariance matrix of  $\mathbf{x}$  is

$$\mathbf{V} = v\mathbf{I}, \quad (5)$$

where

$$v = \frac{1}{N} \sum_{n=0}^{N-1} \text{Var}[x[n]]. \quad (6)$$

The LMMSE estimate of  $\mathbf{x}$  is [14]

$$\hat{\mathbf{x}} \equiv E[\mathbf{x} | \mathbf{y}] = E[\mathbf{x}] + \mathbf{V} \mathbf{H}^H \mathbf{R}^{-1} (\mathbf{y} - E[\mathbf{y}]) \quad (7)$$

where  $E[\mathbf{y}] \equiv \mathbf{H} E[\mathbf{x}]$ , and  $\mathbf{R}$  is the covariance matrix of  $\mathbf{y}$ :

$$\begin{aligned} \mathbf{R} &\equiv E[(\mathbf{y} - E[\mathbf{y}])(\mathbf{y} - E[\mathbf{y}])^H] \\ &= \mathbf{H} \mathbf{V} \mathbf{H}^H + \sigma^2 \mathbf{I}. \end{aligned} \quad (8)$$

*Step 3. Soft De-Mapping:* Finally, we consider the LLR  $\{\lambda_k[n], \forall k\}$  defined in (2). Following [5], [6], [9], we can write  $\hat{x}[n]$ , the  $n$ th entry of  $\hat{\mathbf{x}}$ , as

$$\hat{x}[n] = \phi[n]x[n] + \xi[n] \quad (9)$$

where  $\xi[n]$  is modeled as a Gaussian noise and  $\phi[n]$  is selected so as to ensure that  $x[n]$  and  $\xi[n]$  are uncorrelated. Let  $\mathbf{h}[n]$  be the  $n$ th column of  $\mathbf{H}$ . Then

$$\phi[n] \equiv v(\mathbf{h}[n])^H \mathbf{R}^{-1} \mathbf{h}[n]. \quad (10)$$

Treating (9) as a memoryless system with channel gain  $\phi[n]$  and noise  $\xi[n]$ , we can evaluate (2) as [1]:

$$\lambda_k[n] = \ln \frac{\sum_{s \in \mathcal{S}_k^{(0)}} \exp \left( -\frac{|\hat{x}[n] - \phi[n]s - E[\xi[n]]|^2}{\text{Var}[\xi[n]]} \right) \Pr(s)}{\sum_{s \in \mathcal{S}_k^{(1)}} \exp \left( -\frac{|\hat{x}[n] - \phi[n]s - E[\xi[n]]|^2}{\text{Var}[\xi[n]]} \right) \Pr(s)} - \gamma_k[n] \quad (11)$$

where  $\mathcal{S}_k^{(0)}$  and  $\mathcal{S}_k^{(1)}$  are the sets of constellation points whose  $k$ th bit position carries 0 and 1, respectively,  $E[\xi[n]]$ ,  $\text{Var}[\xi[n]]$  and  $\Pr(s)$  are computed using the *a priori* LLRs  $\{\gamma_k[n]\}$  [7], [9].

### D. Details of the Soft Mapper

The following are some details of the soft mapping in Step 1. Let  $b_k[n]$  be a bit related to  $x_k[n]$ . Based on the DEC feedback  $\gamma_k[n]$ , the *a priori* probabilities of  $b_k[n]$  are

$$\Pr(b_k[n] = 0) = \frac{\exp(\gamma_k[n])}{1 + \exp(\gamma_k[n])}, \quad (12a)$$

$$\Pr(b_k[n] = 1) = 1 - \Pr(b_k[n] = 0). \quad (12b)$$

Let  $s$  be a point in the signaling constellation. The *a priori* probability that  $s$  is the transmitted signal is computed as  $\Pr(s) = \prod_{k=1}^K \Pr(b_k[n])$ , where  $\Pr(b_k[n])$  is either  $\Pr(b_k[n] = 0)$  or  $\Pr(b_k[n] = 1)$ , depending on the mapping rule. Then, the mean and variance of  $x[n]$  are, respectively,

$$E[x[n]] = \sum_{s \in \mathcal{S}} s \Pr(s) \quad (13a)$$

$$\text{Var}[x[n]] = \sum_{s \in \mathcal{S}} |s - E[x[n]]|^2 \Pr(s). \quad (13b)$$

In this way, we generate the mean  $E[\mathbf{x}]$  and covariance matrix  $\mathbf{V}$ .

### E. Complexity Analysis

The above LMMSE procedure is a low-cost alternative to the more complicated MAP approach. The main problem for the MAP method is that the signal constellation expands after transmission over a linear channel characterized by (1). To see this, let  $Q_y$  and  $Q_x$  be the constellation sizes of  $\mathbf{H}\mathbf{x}$  and  $\mathbf{x}$  in (1), respectively. Then  $Q_x = 2^K$  and  $Q_y = 2^{KL}$  if there are  $L$  non-zero entries in each row of  $\mathbf{H}$  (since each entry in  $\mathbf{H}\mathbf{x}$  is the summation of  $L$  entries in  $\mathbf{x}$ ). In the worst case,  $L = N$ , where  $N$  is the number of columns of  $\mathbf{H}$ . The complexity of the MAP method is proportional to  $Q_y$ , i.e.,  $O(2^{KL})$  per symbol, which is usually extremely large.

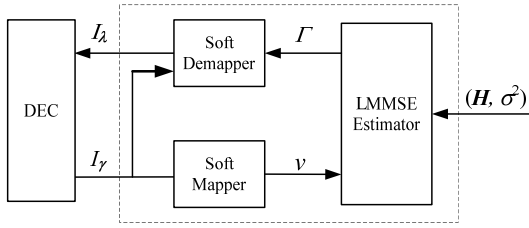


Fig. 2. Evolution model of the iterative LMMSE receiver.

On the other hand, the complexity of the LMMSE approach is  $O(L^2 + 2^K)$  per symbol. (The complexity is  $O(2^K)$  for soft mapping/de-mapping since we work on the constellation of  $x[n]$  only, and the complexity of the LMMSE estimation in Step 2 is  $O(L^2)$  per symbol.) This is greatly reduced from the MAP approach. Furthermore, the LMMSE estimation in Step 2 can be implemented by a fast technique based on cyclic prefixing and fast Fourier transform (FFT). (See [10] for details.) This approach has complexity  $O(\log_2(N))$  and thus can reduce the ESE complexity to  $O(\log_2(N) + 2^K)$ . Later, we will show that the ESE complexity can be further reduced to  $O(\log_2(N) + K)$  using SCM.

### III. PERFORMANCE ANALYSIS

We now proceed to the analysis problem for the LMMSE scheme outlined above. We employ the SNR-variance evolution technique developed in [9] that is in spirit similar to the *extrinsic* information transfer (EXIT) chart method. We will extend the discussions in [9] to multi-ary cases. Our main finding is a design criterion that can minimize the uncertainty introduced at the “soft mapping” stage before the LMMSE operation. Such uncertainty stems from characterizing the distribution of the transmitted signal by its mean and variance only. We will show below that minimizing such uncertainty leads to maximizing the SNR at the output of LMMSE detection and hence performance improvement.

For illustration, we assume a single-user multipath channel but the discussions below can be extended to more general linear channels. We first discuss the analysis technique and then investigate the impact of signaling schemes on performance.

#### A. Transfer Functions

We employ the evolution model in Fig. 2 to characterize the receiver in Fig. 1. We assume fixed  $\mathbf{H}$ ,  $\sigma^2$  and infinite interleaving lengths (so, similar to the EXIT chart method [15], the results can only approximately characterize systems with finite codeword lengths). We now investigate the four modules in Fig. 2 one by one. We define one transfer function to characterize each module, similar to [9], [15].

**DEC:** The APP decoding algorithm is applied to the DEC. Following [15], we can use the EXIT function

$$I_\gamma = T_{\text{DEC}}(I_\lambda) \quad (14)$$

to fully characterize the DEC function. Here,  $I_\lambda$  is defined as the mutual information between the DEC inputs  $\{\lambda_k[n]\}$  and the coded bits  $\{b_k[n]\}$ , and  $I_\gamma$  is defined similarly for the

DEC outputs  $\{\gamma_k[n]\}$ . The bit-error-rate (BER) performance of the DEC can also be characterized by a monotone decreasing function  $g(\cdot)$  as

$$\text{BER} = g(I_\gamma). \quad (15)$$

**Soft Mapper:** The soft mapper takes the *extrinsic* LLRs  $\{\gamma_k[n]\}$  as inputs and produces the soft estimates of  $x[n]$ . Thus, it is reasonable to characterize the soft mapper by the variance  $v$  of its outputs, i.e.,

$$v = \mathbb{E}[\text{Var}[x[n]]] = T_{\text{SM}}(I_\gamma), \quad (16)$$

where the expectation  $\mathbb{E}[\cdot]$  is with respect to the distribution of  $\{\gamma_k[n]\}$  and  $\text{Var}[x[n]]$  is computed by (13). Assuming infinite interleaving lengths, we have  $v = 1/N \sum_{n=0}^{N-1} \text{Var}[x[n]]$ . At the beginning of the iterations, since there is no *a priori* information from the DEC (i.e.,  $I_\gamma = 0$ ),  $v = T_{\text{SM}}(0) = 1$ , where we have assumed that the average transmitted power  $\mathbb{E}[|x[n]|^2] = 1$ . When perfect *a priori* information is available, we have  $I_\gamma = 1$  and  $v = T_{\text{SM}}(1) = 0$ .

**LMMSE Estimator:** The LMMSE estimate  $\hat{x}[n]$  in (9) can be viewed as the output signal of an equivalent channel with multiplicative coefficient  $\phi[n]$  and additive noise  $\xi[n]$ . The SNR based on the modeling in (9) can be computed as

$$\Gamma[n] \equiv \frac{|\phi[n]|^2}{\text{Var}[\xi[n]]} = (\mathbf{h}[n])^H (\mathbf{R}[n])^{-1} \mathbf{h}[n] \quad (17)$$

where  $\mathbf{R}[n] = v \sum_{n' \neq n} \mathbf{h}[n'] (\mathbf{h}[n'])^H + \sigma^2 \mathbf{I}$ . Consider an  $L$ -tap multipath channel with coefficients  $\{h_0, h_1, \dots, h_{L-1}\}$ . It is shown in [9] that  $\Gamma[n]$  can be approximated as follows:

$$\Gamma[n] \approx \Gamma = \frac{u}{1 - vu}, \quad \forall n \quad (18)$$

where

$$u = \frac{1}{N} \sum_{n=0}^{N-1} \frac{|g[n]|^2}{v|g[n]|^2 + \sigma^2} \quad (19)$$

with  $g[n] = \sum_{l=0}^{L-1} h_l \exp(i \frac{2\pi n l}{N})$  being the frequency-domain channel gain. From the above discussions,  $\Gamma$  is a deterministic function of  $v$ ,  $\mathbf{H}$  and  $\sigma^2$ , denoted by

$$\Gamma = T_{\text{MMSE}}(v, \mathbf{H}, \sigma^2). \quad (20)$$

When the *a priori* information from the DEC is perfect,  $v = 0$  and  $\Gamma$  converges to the upper limit

$$\Gamma = u = \frac{1}{N} \sum_{n=0}^{N-1} \frac{|g[n]|^2}{\sigma^2} = \frac{\sum_{l=0}^{L-1} |h_l|^2}{\sigma^2} \quad (21)$$

where the last equality follows Parseval's theorem.

**Soft De-Mapper:** As illustrated in Fig. 2, the soft de-mapper performance is determined by  $\Gamma$  and  $I_\gamma$ :

$$I_\lambda = T_{\text{DEM}}(\Gamma, I_\gamma). \quad (22)$$

#### B. Evolution Analysis

Among the four transfer functions above, only  $T_{\text{MMSE}}(v, \mathbf{H}, \sigma^2)$  is a function of channel matrix. Fortunately,  $T_{\text{MMSE}}(v, \mathbf{H}, \sigma^2)$  has a closed-form expression in (18) and thus it can be quickly evaluated on-the-fly (rather than

pre-simulated). The other three transfer functions  $T_{\text{DEC}}(I_\lambda)$ ,  $T_{\text{SM}}(I_\gamma)$  and  $T_{\text{DEM}}(\Gamma, I_\gamma)$  are not functions of  $\mathbf{H}$ . They can be obtained by applying the Monte-Carlo method to an AWGN channel and characterized by look-up tables. Therefore, given  $\{\mathbf{H}, \sigma^2\}$ , the overall receiver performance can be predicted by the following evolution procedures, where  $T_{\text{DEC}}(I_\lambda)$ ,  $T_{\text{SM}}(I_\gamma)$  and  $T_{\text{DEM}}(\Gamma, I_\gamma)$  are assumed to be known and  $T_{\text{MMSE}}(v, \mathbf{H}, \sigma^2)$  is computed on line.

**Initialization:** Set  $I_\gamma = 0$ .

**Recursion:** Update  $I_\gamma$  as

$$I_\gamma = T_{\text{DEC}} \left( T_{\text{DEM}} \left( T_{\text{MMSE}} \left( T_{\text{SM}}(I_\gamma), \mathbf{H}, \sigma^2 \right), I_\gamma \right) \right).$$

**Termination:** After a preset number of recursions, estimate the BER by substituting the final value of  $I_\gamma$  into (15).

Later we will see that the above evolution method can provide quick and accurate performance prediction. We can also find the average performance (averaged over the distribution of  $\mathbf{H}$ ) by applying the above method to repeatedly generated samples of  $\mathbf{H}$ .

Note that an alternative to the above four-module approach is to characterize the whole ESE block using a single pre-simulated transfer function. This is feasible for a fixed channel matrix  $\mathbf{H}$ , but becomes intractable for random  $\mathbf{H}$ .

### C. Impact of Signaling Schemes

Given the component code, the system performance depends heavily on the signaling scheme employed. Some common examples of signaling methods can be found in [1]-[3]. Another example is SCM [12], [13] that generates the mapped symbol  $x[n]$  as

$$x[n] = \sum_{k=1}^K \beta_k (-1)^{b_k[n]}, \quad (23)$$

where  $\{\beta_k\}$  are complex constants.

Assume that the coding and decoding methods are fixed. Let us consider the impact of signaling schemes on system performance. Among the four modules in Fig. 2, only the soft mapper and soft de-mapper are dependent on the signaling method. The analysis of the de-mapper is a complicated issue. We will briefly return to it later in Section III-E.

We now focus on the soft-mapper performance. It can be shown based on (18) that the output SNR of the LMMSE estimator is a monotonously decreasing function of the variance  $v$  at the output of the soft mapper (see (16)). Thus,  $v$  should be minimized when the LMMSE performance is concerned. Interestingly, some explicit results can be obtained regarding the SNR maximization (or, equivalently, variance minimization) criterion.

We first make the following assumptions.

*Assumption 1:* The mapping  $\Phi: \mathcal{B} \rightarrow \mathcal{S}$  is unbiased and with unit average power:

$$\sum_{s \in \mathcal{S}} s = 0, \quad 2^{-K} \sum_{s \in \mathcal{S}} |s|^2 = 1. \quad (24)$$

*Assumption 2:* The elements of  $\{\gamma_k[n]\}$  are independent, identically distributed (i.i.d.). The statistics of the *a priori* probabilities are symmetric, which implies that

$$\mathbb{E}[\text{Pr}(b_k[n] = 0)] = \mathbb{E}[\text{Pr}(b_k[n] = 1)] = 1/2, \forall k, \quad (25)$$

there is a constant  $\eta$  such that

$$\eta = \mathbb{E}[\text{Pr}^2(b_k[n] = 0)] = \mathbb{E}[\text{Pr}^2(b_k[n] = 1)], \forall k, \quad (26)$$

and the elements in  $\mathcal{S}$  have equal occurrence probabilities:

$$\mathbb{E}[\text{Pr}[s]] = 2^{-K}, \forall s \in \mathcal{S}. \quad (27)$$

Based on these two assumptions, we have the following theorems. (Due to space limitations, we omit the proof.)

*Theorem 1:* Under Assumptions 1 and 2, the minimum variance

$$\min_{\mathcal{S}, \Phi} v = 2 - 4\eta \quad (28)$$

where the minimization is over all possible selections of signal constellation  $\mathcal{S}$  and mapping rule  $\Phi$ .

*Proof:* See [16]. ■

*Theorem 2:* Under Assumptions 1 and 2, for arbitrary  $K$  and arbitrary  $\{\beta_k\}$ , the SCM signaling defined by (23) achieves the minimum variance.

*Proof:* See [16]. ■

Theorems 1 and 2 show that SCM yields SNR optimization for the LMMSE estimator and potentially improved performance. In this respect, many commonly known signaling schemes (such as QAM with the MSP mapping [1]) are sub-optimal. It will be shown later that QAM with the Gray mapping yields performance close to that of SCM.

### D. Complexity of SCM

An additional advantage of SCM is its low complexity. To see this, let  $x[n]$  be constructed using (23). Then we can rewrite the LMMSE estimator output in (9) as

$$\hat{x}[n] = \phi[n] \sum_{k=1}^K \beta_k (-1)^{b_k[n]} + \xi[n] \quad (29a)$$

$$= \phi[n] \beta_k (-1)^{b_k[n]} + \zeta_k[n] + \xi[n] \quad (29b)$$

where  $\zeta_k[n] = \phi[n] \sum_{m \neq k} \beta_m (-1)^{b_m[n]}$ .

Instead of (11), we can adopt a fast technique in computing the LLR for  $b_k[n]$  by approximately treating (29) as a binary-input system (since  $(-1)^{b_k[n]} \in \{+1, -1\}$ ) with Gaussian noise  $\zeta_k[n] + \xi[n]$ . This basically follows the detection principles of interleave-division multiple-access systems [11] and the related de-mapping complexity is  $O(1)$  per bit and  $O(K)$  per symbol. From the linear nature of (23), the complexity of the soft mapper operations (outlined in Section II-D) is also  $O(K)$ . Thus, the overall ESE complexity with SCM is  $O(\log_2(N) + K)$  if the FFT-based fast technique [9] is applied. This is very low compared with the complexity related to (11). We observed by simulation that, for SCM, the performance difference is marginal between the detection techniques based on (11) and (29). Note that, in general, (11) has to be used for other signaling schemes involving non-linear mapping rules.

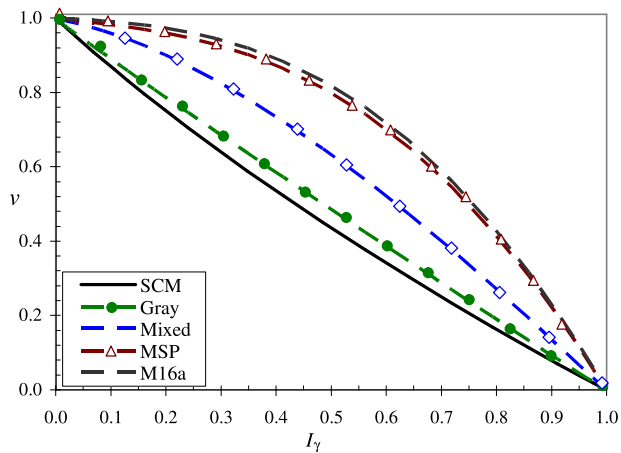


Fig. 3. Comparison of the variance achieved by SCM and four 16-QAM schemes with the Gray, Mixed, MSP [1] and  $M16^a$  mappings [2].

### E. Soft De-Mapper Again

As mentioned earlier, the signaling scheme also affects the soft de-mapper performance, but the related analysis is a complicated issue. There are different criteria for soft de-mapper design, e.g., those based on the distance or mutual information measurements. We are still studying this issue and, in this paper, we rely on simulation results.

An interesting property for SCM is that it can maximize mutual information (between the channel inputs and outputs) when  $K \rightarrow \infty$ , but we will omit the related discussions for space limitation. We would like to comment here that maximizing mutual information is not necessarily an optimal option for systems with limited code lengths.

## IV. NUMERICAL RESULTS

In this section, we present simulation results to verify the above analysis. We first show the impact of signaling schemes on the variance  $v$ . Following [15], we model  $\{\gamma_k[n]\}$  as the output LLRs from a binary-input AWGN channel and characterize their distribution by the mutual information  $I_\gamma$ . Fig. 3 compares the results of five different 16-ary signaling schemes. Clearly, SCM has the smallest  $v$  among all options. Note that the 16-QAM with Gray mapping yields  $v$  close to that of SCM, indicating its property in providing good SNR for the LMMSE estimator.

We next compare the overall system performance. In Fig. 4, we consider single-user BICM-ID systems over a multipath channel. The Proakis B channel is assumed. The simulated and predicted performance for SCM and the MSP signaling are compared. It is seen that SCM significantly outperforms the MSP signaling in the multipath channel with iterative LMMSE detection. (We observed that the performances of the Mixed and Gray signaling schemes considered in Fig. 3 are in between those of SCM and the MSP signaling.) This demonstrates that the advantage of SCM in maximizing the output SNR of the LMMSE detector can indeed lead to significant improvement in BER. It is also shown that the prediction results are quite close to the simulation results.

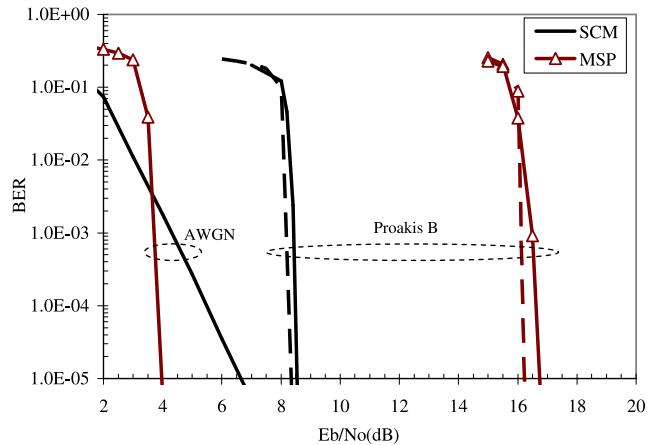


Fig. 4. Comparison of BICM-ID with different 16-ary signaling schemes over the Proakis B channel. A 4-state rate-1/2 convolutional code  $(5,7)_8$  is used and the information block length is 65536. The system throughput  $R = 2$  bits/channel use. The Proakis B channel [a 3-tap multipath channel with tap coefficients  $(0.407, 0.815, 0.407)$ ] is normalized in simulations. For the SCM,  $K = 4$ ,  $\beta_1 = i, \beta_2 = 1, \beta_3 = i, \beta_4 = 1.5$ , where  $i = \sqrt{-1}$ . The dashed and solid curves represent the prediction and simulation results, respectively.

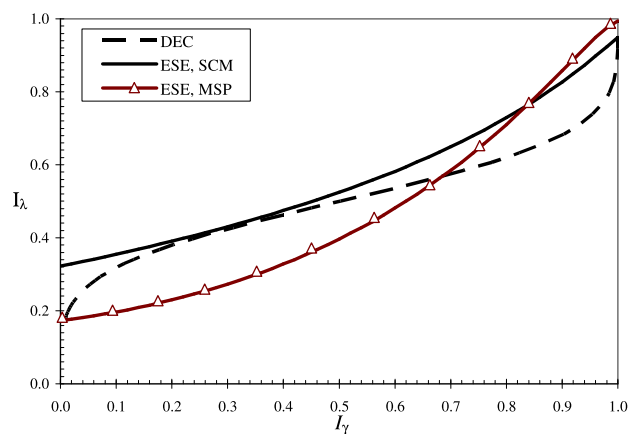


Fig. 5. EXIT chart for BICM-ID with different 16-ary signaling schemes over the Proakis B channel. The system parameters are the same as those in Fig. 4.  $E_b/N_0 = 8$  dB.

Note that for an AWGN channel, the ESE reduces to a symbol-by-symbol MAP estimator, since there is no inter-symbol-interference in this case. Then minimizing variance has no effect on performance and the MSP signaling can outperform SCM, as seen in Fig. 4.

The advantage of SCM can also be verified by examining the asymptotic convergence behavior of the iterative decoding. We apply the EXIT chart technique for this purpose and consider the fixed Proakis B channel used in Fig. 4. The ESE and DEC are characterized by the EXIT curves  $I_\gamma \rightarrow I_\lambda$  and  $I_\lambda \rightarrow I_\gamma$ , respectively. From Fig. 5, we can see that for the MSP signaling, the tunnel between the two curves closes at small  $I_\gamma$ , resulting in poor convergence and high BER.

In the above, we have assumed a fixed multipath channel. Next we investigate the average performance and consider a

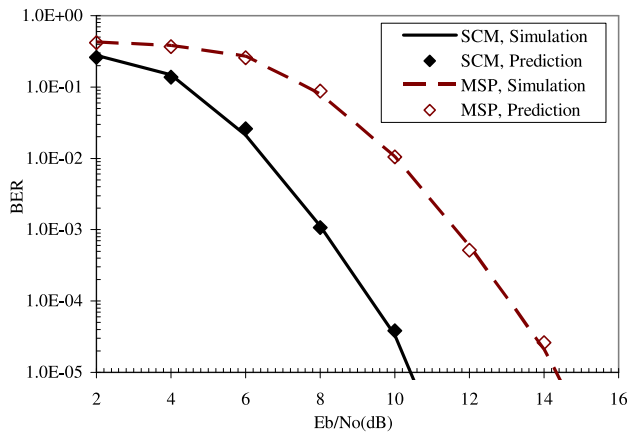


Fig. 6. Performance of BICM-ID with different 16-ary signaling schemes over random multipath channels. The system parameters are the same as those in Fig. 4. For the SCM,  $K = 4$ ,  $\beta_1 = i\beta_2 = 1$ ,  $\beta_3 = i\beta_4 = 1.5$ .

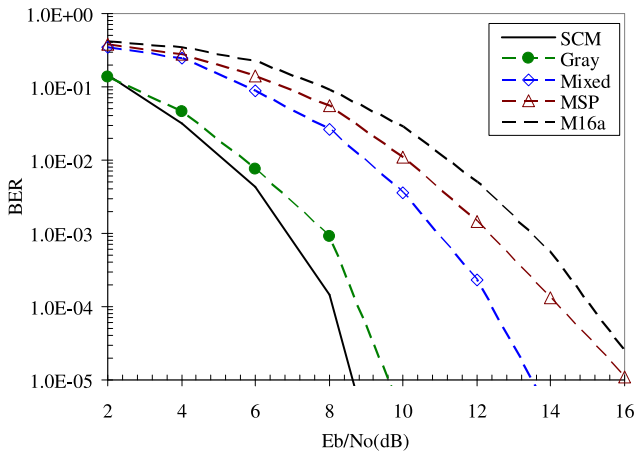


Fig. 7. Performance of BICM-ID with different 16-ary signaling schemes over random MIMO multipath channels. The rate-1/2 convolutional code  $(5, 7)_8$  is used. The system throughput is 4 bits/channel use. The  $2 \times 2$  channel has 4 taps with complex Gaussian coefficients. For the SCM,  $K = 4$ ,  $\beta_1 = i\beta_2 = 1$ ,  $\beta_3 = i\beta_4 = 2$ .

16-tap random multipath channel where each tap coefficient is a complex Gaussian random variable with variance  $1/16$ . The performance averaged over the distribution of the channel realizations is presented in Fig. 6. We observe that SCM is again advantageous in this case.

Finally, we show an example of MIMO multipath channels. At the transmitter, a BICM encoder is applied and the mapped symbols are multiplexed to multiple antennas in a BLAST manner. Random  $2 \times 2$  multipath channels with 4 taps are assumed. The performances with the five signaling schemes in Fig. 3 are compared in Fig. 7. It is seen that SCM yields the best performance. The gap between the SCM and MSP performance is large than 7 dB at  $\text{BER} = 10^{-5}$ . The 16-QAM Gray signaling leads to performance close to that of SCM, which agrees well with the variance results in Fig. 3.

## V. CONCLUSIONS

We have shown by variance analysis that the iterative LMMSE detection performance of multi-ary systems is highly related to the signaling schemes employed. We show that SCM can maximize the output SNR of the LMMSE detector, which is beneficial for overall system performance. Consequently, SCM can outperform other conventional signaling schemes over various channels such as multipath channels and MIMO channels. The simulation results agree well with the analysis.

## ACKNOWLEDGMENT

This work was fully supported by a grant from the Research Grant Council of the Hong Kong SAR, China [Project No. CityU 117305].

## REFERENCES

- [1] A. Chindapol and J. A. Ritcey, "Design, analysis, and performance evaluation for BICM-ID with square QAM constellations in Rayleigh fading channels," *IEEE J. Select. Areas Commun.*, vol. 19, no. 5, pp. 944–957, May 2001.
- [2] F. Schreckenbach, N. Gortz, J. Hagenauer, and G. Bauch, "Optimization of symbol mappings for bit-interleaved coded modulation with iterative decoding," *IEEE Commun. Letters*, vol. 7, pp. 593–595, Dec. 2003.
- [3] J. Tan and G. L. Stuber, "Analysis and design of symbol mappers for iteratively decoded BICM," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 662–672, Mar. 2005.
- [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] M. Tuchler, R. Kowtner, and A. C. Singer, "Turbo equalization: principles and new results," *IEEE Trans. Commun.*, vol. 50, pp. 754–767, May 2002.
- [6] A. Dejonghe and L. Vandendorpe, "Turbo-equalization for multilevel modulation: an efficient low-complexity scheme," in *Proc. IEEE ICC'02*, New York, USA, May 2002, vol. 3, pp.1863–1867.
- [7] V. Ramon, C. Herzet, and L. Vandendorpe, "A semi-analytical method for predicting the performance and convergence behavior of a multiuser Turbo-equalizer/demapper," *IEEE Trans. Signal. Process.*, vol. 55, no. 3, pp. 1104–1117, Mar. 2007.
- [8] A. Dejonghe and L. Vandendorpe, "Bit-interleaved turbo equalization over static frequency-selective channels: Constellation mapping impact," *IEEE Trans. Commun.*, vol. 52, no. 12, pp. 2061–2065, Dec. 2004.
- [9] X. Yuan, Q. Guo, X. Wang, and Li Ping, "Evolution analysis of low-cost iterative equalization in coded linear systems with cyclic prefixes," *IEEE J. Select. Areas Commun.*, vol. 26, no. 2, pp. 301–310, Feb. 2008.
- [10] D. Falconer, S. Ariyavitakul, A. Benyamin-Seeyar and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 40, no. 4, pp. 58–66, Apr. 2002.
- [11] 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.
- [12] L. Duan, B. Rimoldi, and R. Urbanke, "Approaching the AWGN channel capacity without active shaping," in *Proc. IEEE Int. Symp. Information Theory*, Ulm, Germany, June/July 1997, p. 374.
- [13] X. Ma and Li Ping, "Coded modulation using superimposed binary codes," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3331–3343, Dec. 2004.
- [14] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall PTR, 1993.
- [15] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.
- [16] J. Tong and Li Ping, "Iterative soft compensation for coded OFDM systems with peak-power limitation," in preparation.