# Interleave-Division Multiple-Access (IDMA) Communications [1]

Li Ping, Lihai Liu, K. Y. Wu, and W. K. Leung
Department of Electronic Engineering
City University of Hong Kong, Hong Kong
eeliping@cityu.edu.hk

**Abstract:** *In this paper, we study a multiple access scheme referred to as interleave-division multiple-access (IDMA), in which interleavers are used as the only means for user separation. IDMA inherits many advantages from CDMA, such as diversity against fading and mitigation of the worst-case other-cell user interference problem. A low-cost iterative chip-by-chip multi-user detection algorithm is described with complexity independent of the user number and increasing linearly with the path number. The advantages of low-rate coded systems are demonstrated using analytical and simulation results. Throughput of 3 bits/chip is observed with one receive antenna and 6 bits/chip with two receive antennas in multipath channels. Performance in a Gaussian multiple access channel at 1.4 dB away from the theoretical limit is demonstrated. These low-cost, high performance advantages can be maintained in asynchronous multipath environments.*

***Keywords**:* CDMA, channel capacity, iterative decoding, multi-user detection.

## 1. INTRODUCTION

The performance of code-division multiple-access (CDMA) systems is mainly limited by multiple access interference (MAI) and intersymbol interference (ISI). Encouraged by the success of turbo codes [1] in additive white Gaussian noise (AWGN) channels, turbo-type iterative multi-user detection (MUD) has been extensively studied [2]-[9] to mitigate such interference and significant progress has been made.

A conventional random waveform CDMA (RW-CDMA) system (such as IS-95) involves separate coding and spreading operations. Theoretical analysis [10][11] shows that the optimal multiple access channel (MAC) capacity is achievable only when the entire bandwidth expansion is devoted to coding. This suggests combining the coding and spreading operations using low-rate codes to maximize coding gain. However, how can we separate different users without spreading within the CDMA framework?

One possible solution to this problem is to employ chip-level interleavers for user separation. This principle has been considered previously and its potential advantages have been demonstrated [2][12]-[17]. Ref. [2] showed the possibility of employing interleaving for user separation in coded systems. Ref. [12] proposed narrow-band coded-modulation schemes using trellis code structures for user separation and interleaving was considered as an option. For wideband systems, the performance improvement by assigning different interleavers to different users in conventional CDMA has been demonstrated in [13] and [14]. Ref. [15] studied a chip interleaved CDMA scheme and a maximal-ratio-combining (MRC) technique for ISI MACs. It clearly demonstrated the advantages of introducing chip-level interleavers. An interleaver-based multiple access scheme has also been studied in [16][17] for high spectral efficiency, improved error performance and low receiver complexity.

This paper concerns transmission and detection principles using interleavers as the only means for user separation, incorporating the principles developed in [2][12]-[17]. The scheme considered is a special case of CDMA in which bandwidth expansion is entirely performed by low-rate coding. For convenience, we will refer to it as interleave-division multiple-access (IDMA) [16][17]. It can also be regarded as a special form of CDMA by treating interleaving index sequences as multiple access codes. IDMA inherits many advantages from CDMA, in particular, diversity against fading and mitigation of the worst-case other-cell user interference problem. Furthermore, IDMA allows a very simple chip-by-chip iterative MUD strategy [16]. The normalized MUD cost (per user) can be made independent of the user number. A low-cost rake MUD structure is described for multipath environments and joint channel estimation and MUD operations are investigated. Simulation results demonstrate the advantages of the IDMA scheme in terms of both bandwidth and power efficiencies for systems with large numbers of users. For example, with simple convolutional/repetition codes, an overall throughput of 6 bits/chip with two receive antennas is observed in multipath environments. Also, with a turbo-Hadamard code [18], MUD performance at 1.4 dB away from the theoretical limit is demonstrated in a Gaussian MAC.

## 2. MOTIVATIONS

### 2.1. Practical Considerations

Inspired by the success of turbo codes, iterative MUD has been extensively studied in recent years. An illustrative iterative MUD scheme is shown in Fig. 1. At the transmitter side, the data from user-*k* is first encoded

[1] This work was fully supported by a grant from the Research Grant Council of the Hong Kong Special Administrative Region, China [Project No. CityU 1110/00E].

by a forward error correction (FEC) code $C$ followed by an interleaver $\pi_k$. A spreading operation is then applied to produce the transmitted signal.

At the receiver side, the received signal is passed through a bank of correlators. A turbo process is then applied involving two functions: an elementary multi-user detector (EMUD) and a bank of $K$ soft-in-soft-output decoders (DECs) based on $C$.

In the receiver in Fig. 1, two constraints must be considered: (i) the FEC code $C$ and (ii) the correlation among signature sequences. Finding a joint optimal solution is usually computationally prohibitive. The turbo processor takes a sub-optimal approach by decomposing the task into two parts. The EMUD processes constraint (ii) and ignores (i). The DECs process (i) and ignore (ii). A global iterative process is then applied to refine the results.

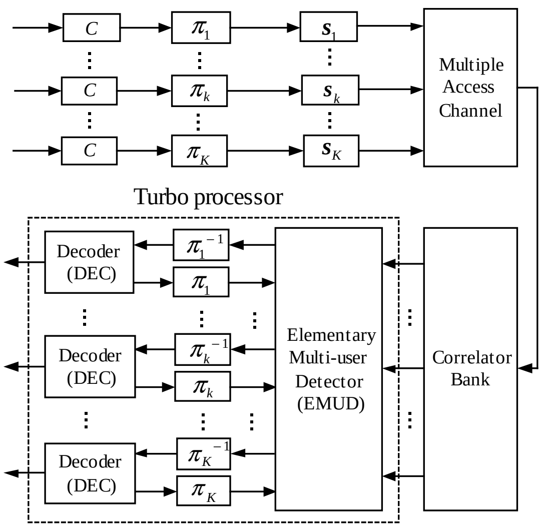


Figure 1. Conventional CDMA transmitter and iterative MUD receiver.

Generally speaking, in Fig. 1, the DEC complexity per user is independent of the simultaneous user number $K$. The EMUD complexity per user, on the other hand, increases rapidly with $K$. For example, the complexity of the well-known MMSE technique [4] is $O(K^2)$ per coded bit per user. This can be a concern when $K$ is large.

Hence we may want to minimize the EMUD cost without seriously affecting the DEC cost. A potential method is to move the spreading operation into FEC coding so as to reduce the EMUD cost related to spreading. This is the strategy to be detailed later.

## 2.2. Theoretical Considerations

Besides cost issues, there are also theoretical ones. Although the existing RW-CDMA systems (such as IS-95) have advantages in multi-cell cellular environments, they have relatively low throughput in single-cell environments compared with FDMA and TDMA. This is mainly due to the effect of the same-cell-user MAI. However, the fundamental work in [10][11] shows that, at least theoretically, there is no penalty on throughput with RW-CDMA even in single-cell environments, provided that the signaling method is designed properly. We now explain this issue below.

Some definitions are in order. Let $R$ be the rate of the FEC code $C$ and $N$ the length of the signature sequence used by the spreaders. Assume the same $R$ and $N$ for all the users. After multiple access by $K$ users, on average $R{\times}K/N$ information bits are transmitted across the channel during a chip duration, so we refer to $R{\times}K/N$ as the system throughput. We denote by $\eta$ the maximum achievable value of $R{\times}K/N$ for a given $R$ for reliable communication ($\eta$ is called the spectral efficiency in [11]).

Fig. 2 shows the $\eta$ versus $R$ curves according to equation (9) in [11] for an RW-CDMA system with AWGN and $K\rightarrow\infty$. (We expect that they also shed light on cases with finite $K$.) It is seen from Fig. 2 that $\eta$ is maximized when $R$ is minimized for any fixed $E_b/N_0$. Some consequences of this observation are as follows.

- Assume that we fix the bandwidth expansion factor ($=N/R$) for each user. Then minimizing $R$ implies minimizing $N$. The minimum value of $N$ is 1, meaning no spreading.
- RW-CDMA incurs no penalty on throughput when $R\rightarrow 0$, since it can be verified that $\eta$ approaches the unconstrained AWGN channel capacity when $R\rightarrow 0$.
- There will be a non-negligible penalty on system throughput if $R$ is not sufficiently small.

We can also justify the above statements using intuition. Lower rate codes can provide higher coding gains, so they should be used in systems with bandwidth expansion. If we do not fully exploit redundancy related to bandwidth expansion using FEC coding, penalties on spectral or power efficiency result.

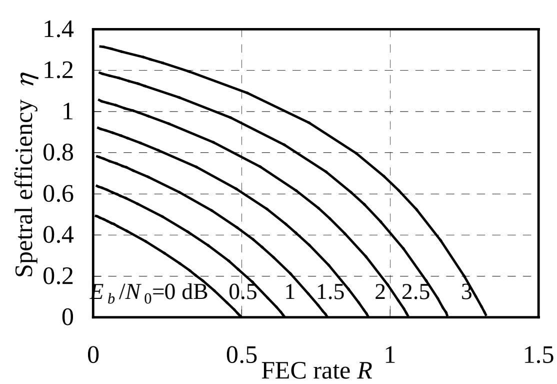


Figure 2. Spectral efficiency versus FEC rate in a real valued Gaussian multiple access channel for $K\rightarrow\infty$.

A conventional CDMA scheme, such as IS-95, employs FEC coding rate at around 1/2 ~ 1/3. Quite long signature sequences are required to support a large number of users. This is not an optimized approach (although perhaps a convenient one) according to the above discussion.

Clearly, some structural change is necessary if we want to make progress towards fully exploiting the available capacity. A key issue is how to achieve multiple access in low-rate coded systems.

# 3. IDMA TRANSMITTER AND RECEIVER PRINCIPLES

The discussion in Section 2 points to simplifying the EMUD and minimizing the FEC coding rate. Fig. 3 shows transmitter and (iterative) receiver structures based on these principles, incorporating the work developed in [2][12]-[17]. Since interleaving is the only mechanism for user separation here, it is referred to as interleave-division multiple-access (IDMA) [16][17].

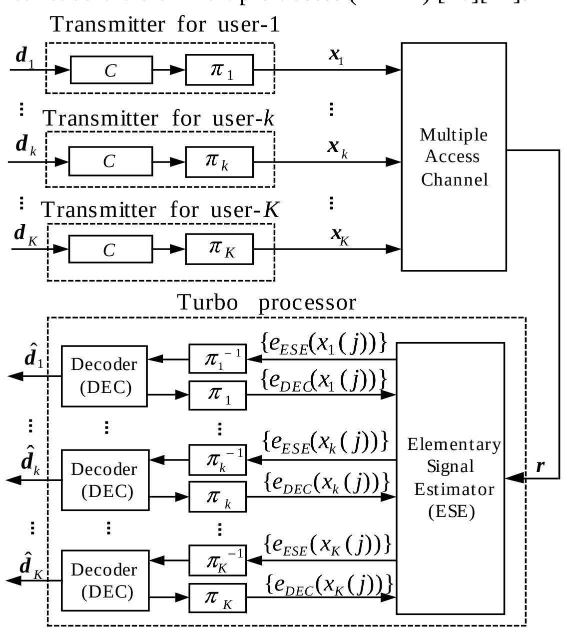


Figure 3. Transmitter and (iterative) receiver structures of an IDMA scheme with $K$ simultaneous users.

## 3.1. Transmitter Structure

The upper part of Fig. 3 shows the transmitter structure of an IDMA system with $K$ simultaneous users. The input data sequence $\boldsymbol{d}_k$ of user-$k$ is encoded based on a low-rate code $C$. The coded sequence is then interleaved by a chip-level interleaver $\pi_k$, producing $\boldsymbol{x}_k \equiv [x_k(1), \ldots, x_k(j), \ldots, x_k(J)]^T$. Following the convention, we call the elements in $\boldsymbol{x}_k$ "chips".

The key principle of IDMA is that the interleavers $\{\pi_k\}$ should be different for different users. We assume that the interleavers are generated independently and randomly. These interleavers disperse the coded sequences so that the adjacent chips are approximately uncorrelated, which facilitates the simple chip-by-chip detection scheme discussed below.

Assume quasi-static single-path channels. After chip-matched filtering, the received signal from $K$ users can be written as

$$r(j) = \sum_{k=1}^{K} h_k x_k(j) + n(j), \qquad j = 1, 2, \ldots, J \quad (1)$$

where $h_k$ is the channel coefficient for user-$k$ and $\{n(j)\}$ are samples of an AWGN with variance $\sigma^2 = N_0/2$. We assume that the channel coefficients $\{h_k\}$ are known *a priori* at the receiver side.

## 3.2. Receiver Structure

The receiver operation is still based on two constraints: (i) the constraint of the FEC code $C$ and (ii) the constraint due to the superposition of the transmitted chips. Again we adopt a sub-optimal receiver structure, as illustrated in Fig. 3. Without signature-sequences, the processing related to the second constraint becomes very simple. The chip-level interleavers $\{\pi_k\}$ allow us to adopt a chip-by-chip estimation technique [16], as detailed below.

The receiver in Fig. 3 consists of an elementary signal estimator (ESE) and $K$ single-user *a posteriori* probability (APP) decoders (DECs).

For simplicity, we first consider BPSK signaling, i.e., $x_k(j) \in \{+1, -1\}$, and real channel coefficients. The generalization to situations without these restrictions will be discussed in Section 3.5.

Denote by $e(x_k(j))$ the extrinsic information about $x_k(j)$ [16][17]. It is further distinguished by subscripts, i.e., $e_{ESE}(x_k(j))$ and $e_{DEC}(x_k(j))$, depending on whether it is generated by the ESE or DECs. During the turbo-type iterative process, the extrinsic information generated by the ESE is used (after de-interleaving) as the *a priori* information in the DECs, and vice versa.

The ESE uses $\{r(j)\}$ and $\{e_{DEC}(x_k(j))\}$ as its inputs and only constraint (ii) is considered here. The output of the ESE is the extrinsic information $\{e_{ESE}(x_k(j))\}$ about $\{x_k(j)\}$. The de-interleaved version of $\{e_{ESE}(x_k(j)), \forall j\}$ is used as the *a priori* information in the $k$th DEC for user-$k$. The output of the $k$th DEC (after interleaving) is the updated extrinsic information $\{e_{DEC}(x_k(j)), \forall j\}$ based on the code constraint of $C$.

The overall procedure is repeated a pre-set number of times. During the final iteration, the DECs produce hard decisions $\{\hat{\boldsymbol{d}}_k\}$ on information bits $\{\boldsymbol{d}_k\}$.

The APP decoding in the DECs is a standard function [1][18], so we will not discuss it in detail. In the following, we will only consider the ESE function.

## 3.3. The ESE Function

Recall from Section 2.1 that we wish to minimize the EMUD cost in Fig. 1. This is the strategy used in Fig. 3 where the ESE is an extremely simplified version of the

EMUD. The constraint of $C$ is ignored in the ESE. We first assume that the channel has no memory. Then the ESE operation can be carried out in a chip-by-chip manner, i.e., only one channel observation value $r(j)$ is used at a time. Clearly, not much can be done in the ESE, which implies that it must be very simple.

From (1), we write

$$r(j) = h_k x_k(j) + \zeta_k(j) \tag{2a}$$

where

$$\zeta_k(j) \equiv \sum_{\substack{k'=1 \\ k' \neq k}}^{K} h_{k'} x_{k'}(j) + n(j) \tag{2b}$$

is the distortion (including interference-plus-noise) contained in $r(j)$ with respect to user-$k$. From the central limit theorem, $\zeta_k(j)$ is approximately Gaussian, so $x_k(j)$ can be estimated from (2a) provided that the mean and variance of $\zeta_k(j)$ are available. A detailed derivation of the detection algorithm is given in [16]. Here we only list the algorithm, where E(·) and Var(·) denote the mean and variance respectively.

**(i) Initialization:** Set $e_{DEC}(x_k(j)) = 0, \quad \forall k, j$.

**(ii) Main operations:**

$$\mathrm{E}(x_k(j)) \Leftarrow \tanh(e_{DEC}(x_k(j))/2), \quad \forall k, j. \tag{3a}$$

$$\mathrm{Var}(x_k(j)) \Leftarrow 1 - (\mathrm{E}(x_k(j)))^2, \quad \forall k, j. \tag{3b}$$

$$\mathrm{E}(r(j)) \Leftarrow \sum_{k'=1}^{K} h_{k'} \mathrm{E}(x_{k'}(j)), \quad \forall j. \tag{4a}$$

$$\mathrm{Var}(r(j)) \Leftarrow \sum_{k'=1}^{K} |h_{k'}|^2 \mathrm{Var}(x_{k'}(j)) + \sigma^2, \quad \forall j. \tag{4b}$$

$$e_{ESE}(x_k(j)) \Leftarrow 2h_k \cdot \frac{r(j) - \mathrm{E}(r(j)) + h_k \mathrm{E}(x_k(j))}{\mathrm{Var}(r(j)) - |h_k|^2 \mathrm{Var}(x_k(j))}, \quad \forall k, j. \tag{5}$$

It can be verified that the above algorithm is an extremely simplified form of that derived in [4] (i.e., when the spreading sequences are all of length-1).

The normalized computational cost in (3)-(5) (excluding the APP decoding of $C$) is only a few additions and multiplications per chip per user. Note that the results of (4), i.e., E($r(j)$) and Var($r(j)$), can be shared by all the users. The cost per information bit per user increases linearly with the spreading length but is independent of the user number $K$.

## 3.4. The Rake ESE Function

We now consider the detection of IDMA systems in more complicated asynchronous channels with memory. The overall decoder structure is the same as that shown in Fig. 3 and so the iterative principle discussed above is still applicable. Therefore, we will only focus on the calculation of $\{e_{ESE}(x_k(j))\}$ in the ESE.

Consider a quasi-static multipath fading channel with $L$ tap-coefficients. Let $\{h_{k,0}, \cdots, h_{k,L-1}\}$ be the fading coefficients related to user-$k$. After chip-matched filtering, the received signal can be represented by

$$r(j) = \sum_{k=1}^{K} \sum_{l=0}^{L-1} h_{k,l} x_k(j-l) + n(j), \quad j = 1, \cdots, J+L-1. \tag{6}$$

In this case, each $x_k(j)$ is observed on $L$ successive samples $\{r(j), r(j+1), \ldots, r(j+L-1)\}$ in the received signal. A rake-type operation can be applied to combine the information about $x_k(j)$ from these samples as

$$e_{ESE}(x_k(j)) = \sum_{l=0}^{L-1} e_{ESE}(x_k(j))_l \tag{7a}$$

where

$$e_{ESE}(x_k(j))_l = 2h_{k,l} \cdot \frac{r(j+l) - \mathrm{E}(r(j+l)) + h_{k,l}\mathrm{E}(x_k(j))}{\mathrm{Var}(r(j+l)) - |h_{k,l}|^2 \mathrm{Var}(x_k(j))} \tag{7b}$$

is the soft estimate of $x_k(j)$ based on $r(j+l)$ using (5). The complexity (per chip per user) of this algorithm is $O(L)$.

There are two alternative treatments for channels with memory. The first is the maximal-ratio-combing (MRC) technique [15], in which $\boldsymbol{r}$ is passed through an MRC filter matched to the $L$ tap-coefficients for a particular user. An MMSE based chip detection is then applied to generate $\{e_{ESE}(x_k(j))\}$. Since $K$ different $L$-tap matched filters are used, the complexity is $O(KL)$ [15] per chip per user. The second is the joint Gaussian (JG) technique discussed in [17] with complexity $O(L^2)$. Generally speaking, the rake, MRC and JG techniques have similar performance in systems with sufficient bandwidth expansion. However, when throughput is very high or when the rate of $C$ is relatively high, the JG technique demonstrates better performance [17].

## 3.5. Generalizations

The detection principles discussed above are for situations with BPSK signaling, real channel coefficients and one receive antenna. We now show that the basic detection principle remains the same when these restrictions are removed.

- With complex channel coefficients, the imaginary part of a received signal sample simply provides an extra observation. Thus, a rake-type operation can be applied. The computational cost is doubled compared with that of the case with real channel coefficients.
- When $N_r$ receive antennas are used, there are $N_r$ times observations with respect to each $x_k(j)$. The rake-type technique in (7a) can again be extended to combine the estimates from all of these observations.
- When QPSK signaling is used, a chip can be transmitted using either in-phase or quadrature signaling. From the receiver point of view, the two situations are nearly the same apart from a 90° phase shift. The detection principle remains the same.

## 4. TURBO CODED SYSTEMS

When a turbo-type code is used as $C$, the DECs contain internal iterations. These internal iterations can be incorporated into the global iteration described above (e.g., one internal iteration per DEC per global iteration). In this case, the overall iterative detection process for each user involves more than two component processors. As an illustration of the basic principle, we assume that user-$k$ employs a standard turbo code $C$ with two constituent codes: $C1$ and $C2$. In this case, the $k$th DEC consists of two local decoders $\mathrm{DEC}_{k,1}$ and $\mathrm{DEC}_{k,2}$ for $C1$ and $C2$ respectively. The overall iteration for user-$k$ involves three component processors: ESE, $\mathrm{DEC}_{k,1}$ and $\mathrm{DEC}_{k,2}$. These processors are executed in a serial manner based on different constraints, i.e., ESE based on constraint (ii) in Section 3.2, $\mathrm{DEC}_{k,1}$ on $C1$ and $\mathrm{DEC}_{k,2}$ on $C2$. The extrinsic information from the two component processors is combined to form the *a priori* information input to the remaining component processor, as illustrated in Fig. 4. This is similar to the treatment employed in [6]. (The combining operations involve interleaving, de-interleaving and addition operations.)

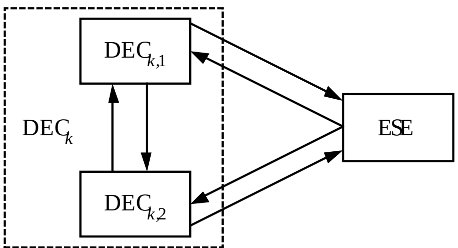


Figure 4. The exchange of extrinsic information among ESE, $\mathrm{DEC}_{k,1}$ and $\mathrm{DEC}_{k,2}$.

## 5. CHANNEL ESTIMATION

The discussions so far are based on the assumption that the receiver has ideal channel state information (CSI). We now proceed to consider the channel estimation issue.

We employ a pilot embedding technique together with iterative joint channel estimation and multi-user detection [19]. We first consider a synchronous channel without memory. For simplicity, write (1) in a vector form as

$$\boldsymbol{r} = \sum_{k=1}^{K} h_k \boldsymbol{x}_k + \boldsymbol{n} \tag{8}$$

where $\boldsymbol{r} \equiv [r(1),\cdots,r(J)]^T$ and $\boldsymbol{n} \equiv [n(1),\cdots,n(J)]^T$. Let $\boldsymbol{p}_k \equiv [p_k(1), \ldots, p_k(J)]^T$ be the pilot for user-$k$. Assume that $\{\boldsymbol{p}_k\}$ are generated randomly and independently. The transmitted signal from user-$k$ is the superposition of the data signal $\boldsymbol{x}_k$ and the pilot $\boldsymbol{p}_k$. The received signal (refer to (8)) can be written as

$$\boldsymbol{r} = \sum_{k=1}^{K} h_k (\boldsymbol{x}_k + \boldsymbol{p}_k) + \boldsymbol{n} \,. \tag{9}$$

Denote by $\hat{\boldsymbol{x}}_k \equiv [\hat{x}_k(1),\cdots,\hat{x}_k(j),\cdots,\hat{x}_k(J)]^T$ the soft estimate of $\boldsymbol{x}_k$. Suppose that we know $\hat{\boldsymbol{x}}_k$ for every $k$. The estimate $\hat{h}_k$ of $h_k$ is generated as

$$\hat{h}_k = \frac{(\hat{\boldsymbol{x}}_k + \boldsymbol{p}_k)^H \left( \boldsymbol{r} - \sum_{k' \neq k} \hat{h}_{k'} \cdot (\hat{\boldsymbol{x}}_{k'} + \boldsymbol{p}_{k'}) \right)}{(\hat{\boldsymbol{x}}_k + \boldsymbol{p}_k)^H (\hat{\boldsymbol{x}}_k + \boldsymbol{p}_k)} \,. \tag{10}$$

The subtraction operation in the numerator is basically a soft-cancellation. The denominator represents a normalization operation. Eqn. (10) is a very simplified version of that used in [19]. The simplification is justified since the pilot sequences are quite long here (as we are considering spread-spectrum systems) and so the simple correlation operation in (10) appears sufficient.

To improve the performance, we adopt an iterative process as shown in Fig. 5. We initialize $\hat{\boldsymbol{x}}_k = \boldsymbol{0}$ and $\hat{h}_k = 0$ for every $k$. The receiver begins by evaluating $\{\hat{h}_k\}$ using (10) in the channel estimator (CE). The ESE and DEC functions are then executed, using $\{\hat{h}_k\}$ as the channel information. The outputs of the ESE and DECs are used to update $\{\hat{\boldsymbol{x}}_k\}$ using the mean values calculated based on the DEC outputs. For BPSK signaling

$$\hat{x}_k(j) = \tanh\{\tfrac{1}{2} l_{DEC}(x_k(j))\}\,, \qquad \forall k, j \tag{11}$$

where

$$l_{DEC}(x_k(j)) \equiv e_{ESE}(x_k(j)) + e_{DEC}(x_k(j))\,, \quad \forall k, j \tag{12}$$

is the *a posteriori* information about $x_k(j)$ generated by the DECs. For QPSK signaling, $\hat{x}_k(j)$ is obtained from the soft estimates of its real and imaginary parts. The CE then refines its estimates on channel coefficients again and the iteration continues.

It is straightforward to generalize the above estimation technique to multipath channels with multiple tap-coefficients. We simply carry out (10) using shifted versions of $\boldsymbol{r}$ to estimate the tap-coefficients corresponding to different delays.

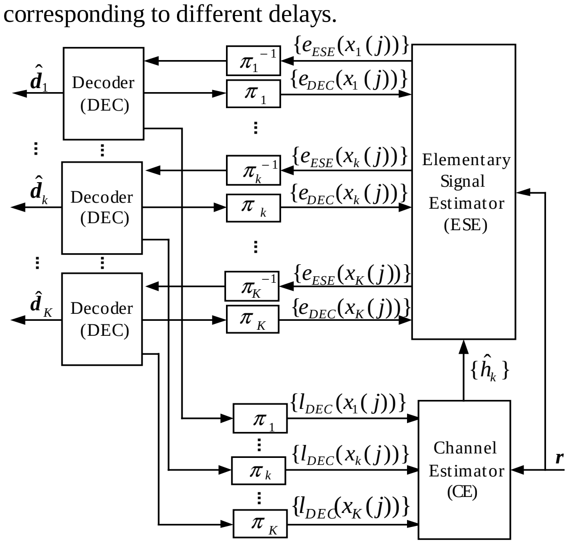


Figure 5. The iterative receiver structure of the joint channel estimation and multi-user detection.

## 6. NUMERICAL RESULTS

In this section, simulation results are provided to illustrate the performance of IDMA systems. Let $N_{\text{info}}$ be the number of information bits in a frame, $K$ the number of simultaneous users in the system, $L$ the tap number of an ISI channel, $N_r$ the number of receive antennas, *It* the iteration number, $R_C$ the rate of each user, and $K \times R_C$ the system throughput that is a measurement of the overall bandwidth efficiency. QPSK signaling is always assumed.

We first examine the bandwidth efficiency of the IDMA scheme based on a simple experimental system. We construct $C$ using a common rate-1/2 $(23, 35)_8$ convolutional code followed by (i.e., in serial concatenation with) a length-8 repetition code ($R_C$ =1/2×1/8=1/16). The repetition coding can also be viewed as a kind of spreading, except that all of the users use the same sequence. The resultant codeword is then multiplied by a mask sequence with alternative signs, i.e., [+1, −1, +1, −1, …], to balance the numbers of +1 and −1. Two independent chip interleavers are employed by each user to produce the in-phase and quadrature parts of the transmitted sequence. The purpose of the masking operation is to maximize the randomness among the transmitted sequences of different users.

Fig. 6 shows the performance of the above system in AWGN channels with different numbers of simultaneous users. As we can see, near single-user performance is achievable for very large $K$ values. Notice that for $K$ = 32, the corresponding system throughput is $K \times R_C$ = 2 bits/chip.

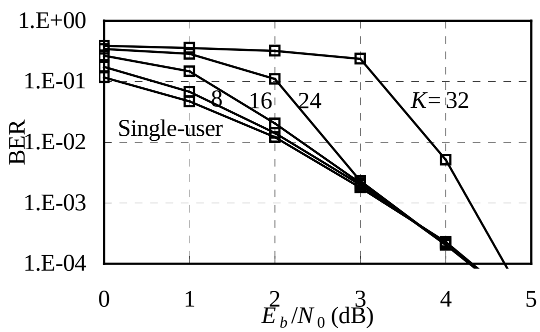


Figure 6. Performance of a convolutionally coded IDMA system in AWGN channels. *It* = 5 for the results with $K \leq 16$ and *It* = 10 for $K$ = 24 and 32. $N_r$ = 1 and $N_{\text{info}}$ = 128.

Fig. 7 shows the performance of the above system in quasi-static Rayleigh fading multipath channels with different tap numbers. $K$ = 32 and $K \times R_C$ = 2 bits/chip. The rake algorithm is used. From this figure, we can see that performance improves uniformly with increasing tap number due to improved diversity.

Fig. 8 shows the performance of the above system in quasi-static Rayleigh fading multipath channels with different tap numbers and receive antenna numbers. The corresponding single-user performance is also included for reference. It is observed that the system can achieve $K \times R_C$ = 3 bits/chip for $K$ = 48 using one receive antenna and $K \times R_C$ = 6 bits/chip for $K$ = 96 using two receive antennas with performance close to the single-user performance at BER = $10^{-4}$. Such throughputs are rather high, recalling that with TDMA we may require a 128-QAM trellis coded modulation scheme to achieve similar throughput and performance.

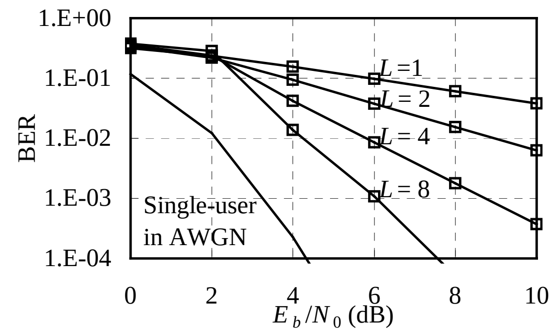


Figure 7. Performance of a convolutionally coded IDMA system in quasi-static Rayleigh fading multipath channels. $K$ = 32, $N_r$ = 1, $K \times R_C$ = 2 bits/chip, *It* = 10 and $N_{\text{info}}$ = 128.

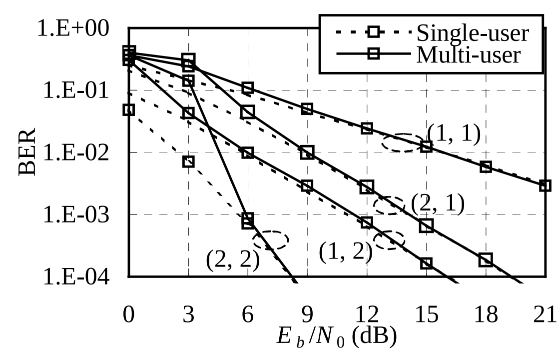


Figure 8. Performance of a convolutionally coded IDMA system in quasi-static Rayleigh fading channels. The ($L$, $N_r$) pair is marked in the figure. $K$ = 48 for one receive antenna and $K$ = 96 for two receive antennas. *It* = 10, $N_{\text{info}}$ = 128.

It is interesting to examine the diversity degree defined as $L \times N_r$, i.e., the total number of paths considering all the antennas. The asymptotic slopes of the curves in Fig. 8 show their corresponding diversity degrees. Clearly, a higher diversity degree results in better performance. For ($L$, $N_r$) = (2, 1) and ($L$, $N_r$) = (1, 2), the diversity degrees are both 2. The asymptotic slopes for the corresponding curves are the same. The performance for ($L$, $N_r$) = (1, 2) has a 3 dB gain since two receive antennas are used. We can see that the IDMA scheme together with the chip-by-chip detection algorithm can efficiently explore the space diversity provided by the channel.

Fig. 9 illustrates the convergence property of the above IDMA system in quasi-static Rayleigh fading multipath channels with $K$ = 32, $L$ = 4 and $N_r$ = 1. It can be seen that convergence is generally achieved within four iterations.

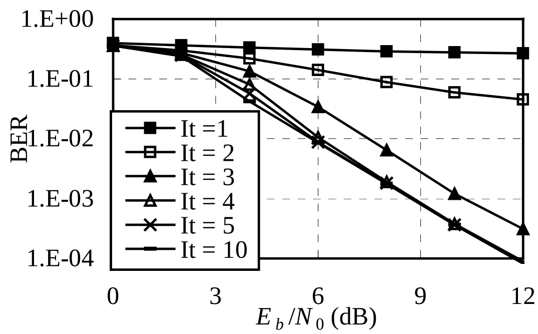


Figure 9. Convergence property of the convolutionally coded IDMA system in quasi-static Rayleigh fading multipath channels with $K$ = 32, $L$ = 4, $N_r$ = 1 and $N_{\text{info}}$ = 128.

Fig. 10 illustrates the performance of the above IDMA system in a 4-tap multipath fading channel with channel estimation. The pilot embedding method discussed in Section 5 is used. We define the energy overhead $\rho$ due to pilots as the ratio between the pilot energy and the total energy

$$\rho \equiv \frac{\text{pilot energy}}{\text{pilot energy} + \text{data energy}}. \tag{13}$$

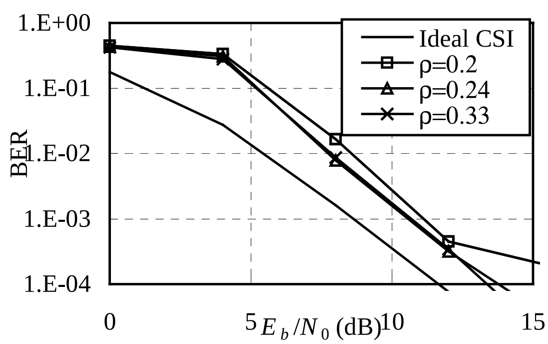


Figure 10. Performance of convolutionally coded IDMA systems with different pilot overheads in quasi-static Rayleigh fading multipath channels with $K$ = 16, $L$ = 4, $N_r$ = 1, $It$ = 10 and $N_{\text{info}}$ = 128. $E_b$ includes the overhead due to pilots.

From Fig. 10, for $\rho$ = 0.24, we observe that there is about 2dB performance degradation for the system without CSI compared to that with ideal CSI. This 2dB degradation is mainly caused by the energy overhead due to pilots. This is in line with the current overhead for CDMA channel sounding.

Next we consider using a more sophisticated low-rate code to improve power efficiency that is actually closely related to spectral efficiency. With higher power efficiency, the transmission power of each user can be reduced. Since the performance of cellular systems is mainly limited by the interference among users, lower transmission power from each user leads to reduced interference, and consequently a larger number of simultaneous users supported [20].

A good choice of low-rate codes is the turbo-Hadamard code studied in [18] that can achieve performance close to the ultimate Shannon limit in AWGN channels. The turbo-Hadamard code used in this paper is constructed by concatenating 3 convolutional-Hadamard codes in parallel, each generated from a length-32 Hadamard code and a convolutional code with polynomial $G(x) = 1/(1+x)$. The information bits in all component codes except one are punctured. A random puncturing operation on parity bits is also adopted to make $R_C$ = 1/16.

Fig. 11 illustrates the performance of IDMA systems based on the turbo-Hadamard code (Scheme I) in AWGN channels. From Fig. 11, performance of BER = $10^{-5}$ is observed at $E_b/N_0 \approx 1.4$ dB with $K$ = 16, which corresponds to $K\times R_C$ = 1 bit/chip. This is only about 1.4 dB away from the corresponding Shannon limit, which is $E_b/N_0$ = 0 dB for a throughput of 1 bit/chip, the same as that for a single-user AWGN channel [11].

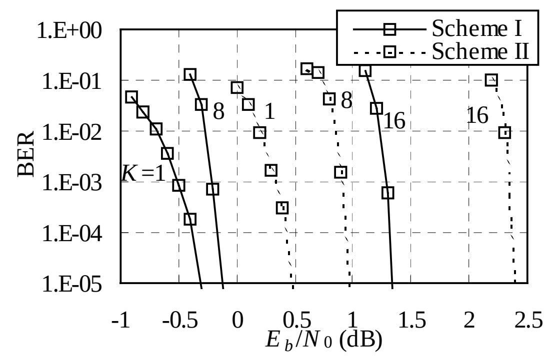


Figure 11. Performance of IDMA systems based on the turbo-Hadamard code and turbo code over AWGN channels with quadrature modulation. $N_r$ = 1, $It$ = 30, $N_{\text{info}}$ = 4095 for Scheme I and $N_{\text{info}}$ = 4096 for Scheme II.

For comparison, in Fig. 11, we have also included the performance of an IDMA system based on a standard turbo code (Scheme II), in which $C$ is constructed using a rate-1/3 $(1, 35/23)_8$ turbo code followed by a length-6 repetition code. Puncturing is then applied to make $R_C$ = 1/16. The advantage of using a low-rate code is clearly seen in Fig. 11. With $K$ = 16, Scheme I demonstrates about 1dB performance advantage over Scheme II, due to the higher coding gain offered by the turbo-Hadamard code. The decoding costs of Schemes I and II are quite similar.

## 7. DISCUSSIONS AND CONCLUSIONS

Consider the first motivation outlined in Section 2 to minimize the cost of the EMUD. Fig. 3 is based on this principle, where spreaders are removed. Bandwidth expansion in IDMA is accomplished as part of the FEC coding. At the receiver side, the ESE in Fig. 3 is much simpler than the EMUD in Fig. 1. This is achieved without the expense of significantly increased cost in the DECs. If a repetition code is used as part of *C*, it only requires two additions and two multiplications per chip in the DECs.

When the rate of *C* in Fig. 1 is lowered, extra coding gain comes as a bonus if *C* is properly designed. This is the second motivation from Section 2. Some low-rate codes such as the super-orthogonal convolutional codes [10] and turbo-Hadamard codes [18] have comparable or even lower complexity than their medium rate counterparts such as convolutional codes and turbo codes. Thus the cost increase in the DECs is usually moderate (unless, of course, one really wants to push towards the capacity).

The benefits of the IDMA scheme are substantial as seen from Figs. 6-11. These include low-cost MUD for systems with large numbers of users, robustness and diversity in multipath environments, very high spectral efficiency and near limit performance.

In conclusion, we have explained the feasibility of interleaver-based multiple access and demonstrated the efficiency of a very low-cost chip-by-chip MUD technique for systems with a large number of users and channels with memory. We expect the basic principles can be extended to other applications, such space-time codes [16] and ultra wideband (UWB) systems [21].

## REFERENCES

[1] C. Berrou and A. Glavieux, "Near Shannon limit error correcting coding and decoding: Turbo-codes," *IEEE Trans. Commun.*, vol. 44, pp. 1261–1271, Oct. 1996.

[2] M. Moher and P. Guinand, "An iterative algorithm for asynchronous coded multi-user detection," *IEEE Commun. Lett.*, vol. 2, pp. 229–231, Aug. 1998.

[3] M. C. Reed, C. B. Schlegel, P. D. Alexander, and J. A. Asenstorfer, "Iterative multi-user 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] Z. Shi and C. Schlegel, "Joint iterative decoding of serially concatenated error control coded CDMA," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1646–1653, Aug. 2001.

[6] A. AlRustamani, A. D. Damnjanovic, and B. R. Vojcic, "Turbo greedy multi-user detection," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1638–1645, Aug. 2001.

[7] A. AlRustamani and B. R. Vojcic, "A new approach to greedy multi-user detection," *IEEE Trans. Commun.*, vol. 50, pp. 1326–1336, Aug. 2002.

[8] M. C. Reed and P. D. Alexander, "Iterative multi-user detection using antenna arrays and FEC on multipath channels," *IEEE J. Select. Areas commun.*, vol. 17, pp. 2082–2089, Dec. 1999.

[9] J. Boutros and G. Carie, "Iterative multi-user joint decoding: Unified framework and asymptotic analysis," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1772–1793, July 2002.

[10] A. J. Viterbi, "Very low rate convolutional codes for maximum theoretical performance of spread spectrum multiple-access channels," *IEEE J. Select. Areas Commun.*, vol. 8, pp. 641–649, Aug. 1990.

[11] S. Verdú and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Thoery*, vol. 45, pp. 622–640, Mar. 1999.

[12] F. N. Brannstrom, T. M. Aulin, and L. K. Rasmussen, "Iterative decoders for trellis code multiple-access," *IEEE Trans. on Communs.*, vol. 50, pp. 1478–1485, Sept. 2002.

[13] S. Brück, U. Sorger, S. Gligorevic, and N. Stolte, "Interleaving for outer convolutional codes in DS-CDMA Systems," *IEEE Trans. Commun.*, vol. 48, pp. 1100–1107, July 2000.

[14] A. Tarable, G. Montorsi, and S. Benedetto, "Analysis and design of interleavers for CDMA systems," *IEEE Commun. Lett.*, vol. 5, pp. 420–422, Oct. 2001.

[15] R. H. Mahadevappa and J. G. Proakis, "Mitigating multiple access interference and intersymbol Interference in uncoded CDMA Systems with chip-level interleaving," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 781–792, Oct. 2002.

[16] Li Ping, L. Liu, K. Y. Wu, and W. K. Leung, "A unified approach to multi-user detection and space-time coding with low complexity and nearly optimal performance," in *Proc. 40$^{th}$ Allerton Conference*, Allerton House, USA, Oct. 2002, pp. 170–179. (www.ee.cityu.edu.hk/~liping/research)

[17] L. Liu, W. K. Leung, and Li Ping, "Simple chip-by-chip multi-user detection for CDMA systems," in *Proc. IEEE VTC*, Korea, Apr. 2003.

[18] Li Ping, W. K. Leung, and K. Y. Wu, "Low-rate turbo-Hadamard codes," submitted for publication, available: www.ee.cityu.edu.hk /~liping/research.

[19] H. Zhu, B. Boroujeny, and C. Schlegel, "Pilot embedding for joint channel estimation and data detection in MIMO communication systems," *IEEE Commun. Lett.*, vol. 7, pp. 30–32, Jan. 2003.

[20] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver, and C. E. Wheatly, "On the capacity of a cellular CDMA system," *IEEE Trans. Vehicular Technology*, vol. 40, pp.303–312, May 1991.

[21] E. Fishler and H. V. Poor, "On the tradeoff between two types of processing gain," in *Proc. 40$^{th}$ Allerton Conference*, Allerton House, USA, Oct. 2002, pp. 1178–1187.