Performance Analysis of a Generic System Model for IDMA using Serial and Parallel Interference Cancellation

Oliver Nagy, Mark C. Reed, Zhenning Shi

Abstract

In this paper the system model of a fully synchronised interleave division multiple access (IDMA) scheme without channel coding is transformed into a new but equivalent matrix model. This model allows for alternative insights into the principles of IDMA and we show that the arising matrices are highly structured and very sparse. Based on this model, the optimal detector for IDMA is derived and compared to the standard iterative receiver. We study the performance of serial and parallel interference cancellation schemes using BER analysis and EXIT charts, and show that serial interference cancellation achieve the best performance. In a last step we demonstrate that any bit interleaved DS-CDMA can be viewed as a special case of IDMA and hence an IDMA detector is applicable to multiuser DS-CDMA systems while achieving comparable performance.

Index Terms

IDMA, CDMA, soft interference cancellation, optimal detector, SIC, PIC.

I. Introduction

AMONG the various communication techniques that have been studied, multiuser direct sequence code division multiple access (DS-CDMA) systems have been deployed around the world and extensively studied. The basic idea of DS-CDMA is to assign each user a unique code sequence to separate the different communication links. The process of applying the code sequence is usually referred to as spreading a bit into several chips, and typically the data stream of each user is interleaved prior to spreading with the aim of distributing consecutive bits over a longer period of time. The order of spreading and interleaving can also be swapped resulting in a chip interleaved system rather than a bit interleaved system. This technique is referred to as chip-interleaving [1] or interleave division multiple access (IDMA) [2, 3].

In [2] the details of IDMA were described, and motivated by soft interference cancellation techniques and iterative turbo receiver schemes, an iterative receiver was proposed. Equipped with this low complexity yet very powerful receiver it was shown by simulation and density analysis [4–6] that the performance of such systems is close to the channel limits. In [7, 8] power allocation schemes for IDMA were investigated to optimise throughput and in [3] it was shown that the capacity of IDMA systems for Gaussian channels is only limited by noise. Work on how to efficiently generate and distribute interleaver configurations throughout the network was carried out in [9].

Clearly, numerous aspects of IDMA have been studied so far, but some fundamental problems remain. In this paper we set out to transform the traditional IDMA block diagram into a generic, yet simple, linear model whereby the optimal detector can be derived. We then present an IDMA issue that degrades system performance when the repetition encoder is not alternating the sign (termed "masking" in this paper), and with our new model we can demonstrate that such systems suffer from a low rank. After that we focus on the iterative receiver to study parallel and serial interference cancellation, and we demonstrate that the latter is capable of achieving the same performance irrespective of whether or not a mask is used. In a last step it is shown, by again utilising our model, that bit interleaved multiuser DS-CDMA systems are a special case of IDMA and that an IDMA receiver can be applied to detect such signals.

Section II summarises the IDMA basics and derives the new system model, in Section III the optimal detector is derived, Section IV covers serial and parallel interference (SIC and PIC respectively) cancellation fashions, Section V illustrates the low rank when neglecting the mask, and Section VI demonstrates the performance of SIC and PIC with and without masking, using BER and EXIT analysis. The IDMA and DS-CDMA relationship is studied in Section VII and conclusions are finally drawn in Section VIII.

For convenience, we will refer to multiuser DS-CDMA simply as CDMA throughout the rest of the paper.

Notation: In this paper vectors are $\mathbf{x} = (x_1, \ldots, x_n)$, whereas $\mathbf{x}_1$ and $\mathbf{x}_2$ denote two different vectors but the elements of such vectors are never addressed explicitly. Random variables/vectors are denoted by capital letters only. To address the k^{th}

[1]This work was conducted while O. Nagy was visiting NICTA. M.C. Reed, and Z. Shi are with National ICT Australia and affiliated with the Australian National University. National ICT Australia is funded through the Australian Government's Backing Australia's Ability initiative and in part through the Australian Research Council.

column of the matrix G we write $\mathsf{G}(\cdot, k)$, and to address just a single element we write $\mathsf{G}_{i,j}$. The bit energy E_b is always 1, *all* signals are real valued and the signal to noise ratio (SNR) in this paper is defined as

$$\text{SNR} \triangleq \frac{E_b}{\sigma^2} = \frac{1}{\sigma^2}$$

where σ^2 is the noise power.

II. Alternative Transmitter Model

At the core of any IDMA system is a chip interleaver and a repetition encoder[2]. The repetition encoder does not necessarily only repeat the bits but may as well alter their signs according to some predefined pattern. For the alternating sign pattern $(+1, -1, +1, -1, \ldots)$, which is often used with IDMA $[2, 4, 5]$, we will use the term "flip-mask" throughout this paper.

As in [2] the usual transmitter structure for an IDMA system is depicted in Figure 1. Each user k has a designated bit vector $\mathbf{B}_k$ to transmit, a rate r repetition encoder (spreader) that maps the bit vector $\mathbf{B}_k$ to the chip vector $\mathbf{X}_k$ and finally a chip interleaver φ_k specific to each user is applied. The transmitted data vector $\mathbf{Z}$ is the superposition of all interleaved chip-streams and the transmit signal is modelled as

$$Z_j = \sum_{k=1}^{n} Y_k. \tag{1}$$

This system model has two major shortcomings:

1) the actual mask- and interleaver choice is not part of the model and consequently there is no way to optimise it based on this model
2) it *assumes* that the received signal $\mathbf{Z}$ is decomposable into the data streams $\mathbf{B}_1$ to $\mathbf{B}_n$, but it is easy to construct cases where this is impossible.

To elaborate on point 2, simply consider the case where all interleavers are identical and all users use the same mask (e.g. flip-mask), then there is no parameter left that uniquely identifies the users.

The model we propose avoids these shortcomings and accurately models the IDMA signal $\mathbf{Z}$ based on the user data, interleaver and mask selection, and is written as

$$\mathbf{Z} = \mathsf{G}\mathbf{B}, \tag{2}$$

where $\mathbf{B}$ is the user data and G describes the chosen interleaver and mask. Therefore our system model allows for better mathematical insight into the fundamental IDMA structure.

We now show how G and $\mathbf{B}$ relate to the data streams $\mathbf{B}_k$, interleavers φ_k, and masks.

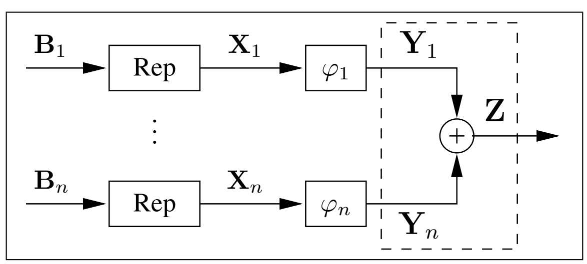


Fig. 1. Transmitter part of IDMA system for n users

The repetition encoding of the data stream $\mathbf{B}_k$ is described by

$$\mathbf{X}_k = \begin{bmatrix} \mathbf{1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{1} \end{bmatrix} \mathbf{B}_k = \mathsf{G}_k \mathbf{B}_k,$$

where $\mathbf{1}$ denotes the all-one column vector of length $1/r$ to achieve a rate r repetition encoding and G_k denotes the generator matrix. Interleaving the chip stream $\mathbf{X}_k$ is equivalent to permuting the corresponding rows in the generator matrix G_k (note that G_k is usually not square). Let G'_k denote the result of permuting the rows of G_k according to the interleaver φ_k, then the encoding and interleaving operation for user k is fully described by

$$\mathbf{Y}_k = \mathsf{G}'_k \mathbf{B}_k,$$

[2]The term *spreading* and *repetition encoding* are interchangeable, and hence the code rate r is the inverse of the spreading gain

and the output signal can be written as

$$\mathbf{Z} = \sum_{k=1}^{n} \mathbf{Y}_k = \sum_{k=1}^{n} \mathsf{G}'_k \mathbf{B}_k = \begin{bmatrix} \mathsf{G}'_1 & \dots & \mathsf{G}'_n \end{bmatrix} \begin{bmatrix} \mathbf{B}_k \\ \vdots \\ \mathbf{B}_n \end{bmatrix}.$$

Introducing the formal substitutions

$$\mathsf{G} \triangleq [\mathsf{G}'_1 \dots \mathsf{G}'_n] \tag{3}$$

$$\mathbf{B} \triangleq [\mathbf{B}_1 \dots \mathbf{B}_n]^T, \tag{4}$$

the fully synchronised IDMA transmitter can be described in the compact form (2).

A. Example for $n = 2$ Users, r=1/2, Interleaving over 4 Chips

Let the two users be denoted by $\mathscr{A}$ and $\mathscr{B}$ transmitting the bit vectors $\mathbf{A} = (A_1, A_2)$ and $\mathbf{B} = (B_1, B_2)$ and using the interleaver mappings φ_A and φ_B respectively, the process is illustrated as follows:

$$\mathscr{A} : (A_1, A_2) \xrightarrow{\text{Rep.}} (A_1, A_1, A_2, A_2) \xrightarrow{\varphi_A} (A_1, A_1, A_2, A_2)$$
$$\mathscr{B} : (B_1, B_2) \xrightarrow{\text{Rep.}} (B_1, B_1, B_2, B_2) \xrightarrow{\varphi_B} (B_1, B_2, B_2, B_1).$$

In this case the interleaver for user φ_A does not alter the chip stream at all, and φ_B exchanges the second and the fourth chip. The transmitted signal vector $\mathbf{Z}$ is the sum of the two interleaved chip vectors, i.e.

$$\mathbf{Z} = [A_1 + B_1, A_1 + B_2, A_2 + B_2, A_2 + B_1]^T. \tag{5}$$

Both users use the same spreading code so the respective generator matrices are

$$\mathsf{G}_A = \mathsf{G}_B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

Swapping the rows in G_A and G_B according to the interleavers φ_A and φ_B yields

$$\mathsf{G}'_A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \qquad \mathsf{G}'_B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The final system matrix is then given by

$$\mathsf{G} = [\mathsf{G}'_A \quad \mathsf{G}'_B] = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix},$$

and the interleaved chip stream is obtained by

$$\mathbf{Z} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_1 + B_1 \\ A_1 + B_2 \\ A_2 + B_2 \\ A_2 + B_1 \end{bmatrix}$$

which equals equation (5).

B. Matrix Properties

Let l_b denote the block length of the encoder, then IDMA systems are using an $(l_b/r, l_b)$ repetition code. Consequently the dimension of the generator matrix for every user k is

$$\dim(\mathsf{G}_k) = \frac{l_b}{r} \times l_b.$$

Because the chip interleaver only permutes the columns of G_k,

$$\dim(\mathsf{G}'_k) = \dim(\mathsf{G}_k).$$

With the definition of the system matrix in (3) its dimension is

$$\dim(\mathsf{G}) = \frac{l_b}{r} \times n l_b. \tag{6}$$

For fully loaded systems (i.e. $r = 1/n$), G becomes square.

For example, if the block length is $l_b = 2$bits and the code rate is $r = 1/3$, then $\dim(\mathsf{G}_k) = 6 \times 2$. From the traditional point of view this is equivalent to the transmitter taking blocks of 2 bits, spreading them into a block of 6 chips that is interleaved and output on the channel.

The system matrix G is very structured and contains only elements from the set $\{0, \pm 1\}$. With the repetition encoding it holds that each column contains exactly $1/r$ non-zero entries. Furthermore each row of G_k (and consequently G'_k) contains one non-zero entry, and as G is composed of n such matrices, the number of nonzero entries is exactly n, or more formally

$$\sum_{k=1}^{l_b/r} |\mathsf{G}_{k,j}| = 1/r, \qquad 1 \le j \le n l_b \tag{7}$$

$$\sum_{k=1}^{n l_b} |\mathsf{G}_{j,k}| = n, \qquad 1 \le j \le l_b/r. \tag{8}$$

For fully loaded systems each row and column sums to n, and using (8), the number of non-zero elements computes to $n l_b/r$. As a measure of sparseness, the ratio s between the number of nonzero elements and the total number of elements of G is introduced:

$$s = \frac{\frac{n l_b}{r}}{\frac{l_b}{r} n l_b} = \frac{1}{l_b}. \tag{9}$$

Note that (9) depends neither on the code rate r nor the number of users n anymore but is a function of the block length l_b only. In [5] for instance a block length of $l_b = 128$ is used which translates to $s = 1/128$ or less than 1% of all entires are non-zero. Consequently it can be concluded that IDMA matrices for reasonably large systems are very sparse.

III. Optimal Detector

Figure 2 shows the extension of the system model (2) with a (possibly stochastic) channel model H and additional noise $\mathbf{N}$. The received signal vector $\mathbf{R}$ is obtained by

$$\mathbf{R} = \mathsf{H}\mathsf{G}\mathbf{B} + \sigma\mathbf{N}. \tag{10}$$

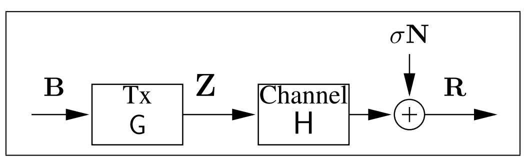


Fig. 2. Generalised IDMA model including a channel and noise

For the derivation of the optimal detector, the channel matrix H is assumed to be deterministic and the noise model is additive white Gaussian noise (AWGN). For simplicity of notation the auxiliary matrix

$$\mathsf{M} \triangleq \mathsf{H}\mathsf{G}$$

is introduced. The system model is then summarised as

$$\begin{aligned}
&\mathbf{R} = \mathsf{M}\mathbf{B} + \sigma\mathbf{N} \\
&\mathbf{B} \in \mathcal{B}, \qquad \mathsf{M} \in \mathbb{R}^{n \times n} \\
&N_k \sim \mathcal{N}(0,1) \quad \text{and i.i.d.} \quad \forall k \\
&B_k \sim \mathcal{U} \quad \text{and i.i.d.} \quad \forall k
\end{aligned}$$

where $\mathcal{B}$ denotes the set of source symbol vectors, $\mathcal{N}(0,1)$ a zero mean Gaussian random variable with variance 1, and $\mathcal{U}$ a uniform distribution. Consequently, for a given data vector $\mathbf{b}$, each element of $\mathbf{R} = (R_1, \ldots, R_k, \ldots, R_n)$ is a Gaussian random variable too with

$$\begin{aligned}
\mu_k &\triangleq \mathsf{M}(k,\cdot)\mathbf{b} \qquad (11) \\
r_k &= \mathsf{M}(k,\cdot)\mathbf{b} + \sigma n_k = \mu_k + \sigma n_k \\
R_k &\sim \mathcal{N}(\mu_k, \sigma^2),
\end{aligned}$$

where n_k is the current realisation of the noise for user k.

The optimal detector applies the MAP decoding rule, i.e.

$$\hat{\mathbf{b}} = \arg\max_{\mathbf{b}\in\mathcal{B}} \; p(\mathbf{b}|\mathbf{r}) = \arg\max_{\mathbf{b}\in\mathcal{B}} \; p(\mathbf{b},\mathbf{r}),$$

with the joint probability being

$$\begin{aligned} p(\mathbf{b},\mathbf{r}) &= p(\mathbf{b})p(\mathbf{r}|\mathbf{b}) \\ &= p(\mathbf{b})p(r_1|\mathbf{b})\prod_{k=2}^{n} p\big(r_k|r_1,\ldots,r_{k-1},\mathbf{b}\big) \\ &= p(\mathbf{b})p(r_1|\mathbf{b})\prod_{k=2}^{n} p\big(r_k|\mu_1+\sigma n_1,\ldots,\mathbf{b}\big) \\ &= p(\mathbf{b})p(r_1|\mathbf{b})\prod_{k=2}^{n} p\big(r_k|\sigma n_1,\ldots,\sigma n_{k-1},\mathbf{b}\big) && (12) \\ &= p(\mathbf{b})\prod_{k=1}^{n} p\big(r_k|\mathbf{b}\big), && (13) \end{aligned}$$

where (12) holds because $\mathbf{b}$ defines μ_k by (11) and equation (13) holds because the noise samples are independent by assumption. With the AWGN model, (13) expands to

$$\begin{aligned} p(\mathbf{b},\mathbf{r}) &= p(\mathbf{b})\prod_{k=1}^{n} p\big(r_k|\mathbf{b}\big) \\ &= p(\mathbf{b})\prod_{k=1}^{n}\big(2\pi\sigma^2\big)^{-1/2}\exp\left(-\frac{(r_k-\mathsf{M}(k,\cdot)\mathbf{b})^2}{2\sigma^2}\right) \\ &= p(\mathbf{b})\prod_{k=1}^{n}\big(2\pi\sigma^2\big)^{-1/2}\exp\left(-\frac{(r_k-\mu_k)^2}{2\sigma^2}\right) && (14) \\ &= p(\mathbf{b})\big(2\pi\sigma^2\big)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(r_k-\mu_k)^2\right), \end{aligned}$$

where in (14) the substitution (11) was used to obtain a more compact notation, but μ_k is still a function of $\mathbf{b}$.

Because the input symbols $\mathbf{b}\in\mathcal{B}$ are uniformly distributed by assumption, the MAP decoding rule finally becomes

$$\begin{aligned} \hat{\mathbf{b}} &= \arg\max_{\mathbf{b}\in\mathcal{B}} \; \frac{1}{|\mathcal{B}|}\big(2\pi\sigma^2\big)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(r_k-\mu_k)^2\right) \\ &= \arg\max_{\mathbf{b}\in\mathcal{B}} \; \exp\left(-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(r_k-\mu_k)^2\right) \\ &= \arg\min_{\mathbf{b}\in\mathcal{B}} \; +\frac{1}{2\sigma^2}\sum_{k=1}^{n}(r_k-\mu_k)^2 \\ &= \arg\min_{\mathbf{b}\in\mathcal{B}} \; \sum_{k=1}^{n}(r_k-\mu_k)^2 \\ &= \arg\min_{\mathbf{b}\in\mathcal{B}} \; \|\mathbf{r}-\mu_k\|. && (15) \end{aligned}$$

The practical relevance of this brute force method is very limited due to the computational cost for already very small matrices, but in the next section it will be a useful tool for comparative reasons.

IV. Interference Cancellation in the Iterative Receiver

The iterative IDMA receiver consists of an interference cancellation (IC) unit and a de-spreader. Two IC methods are investigated: serial interference cancellation (SIC) and parallel interference cancellation (PIC).

The basic idea of interference cancellation is to subtract all but the user of interest data from the incoming data signal to reduce/cancel the interference from other users. In general the bits of other users are unknown and only an estimate obtained from the soft output of the decoder.

The received signal for user k is his chip X_k plus the sum of all the other users, i.e.

$$\begin{aligned} R_k &= X_k + \sum_{l \neq k} X_l + \sigma N \\ &= X_k + W_k, \end{aligned}$$

where the interference and the noise were merged into a single random variable W_k, which by the central limit theorem is modelled as a Gaussian random variable, allowing the simple model definition

$$R_k \approx X_k + W_k \qquad\qquad W_k \sim \mathcal{N}(\mu, \varsigma), \tag{16}$$

where μ is the mean and ς^2 is the variance of W_k. The iterative receiver performs the following steps (for a detailed derivation see for instance [5]):

1) Set all Λ values to zero
2) $$\varsigma_{k,j}^2 = \sum_{l \neq k} \left(E_c - \mathsf{E}[\Lambda_{l,j}]^2 \right) h_l^2 + \sigma^2 \tag{17}$$
3) $$\Lambda_{k,j} = \frac{2 h_k \sqrt{E_c}}{\varsigma_{k,j}^2} \left(\mathsf{R}_j - \sum_{l \neq k} \mathsf{E}\left[\bar{X}_{l,j}\right] h_l \right) \tag{18}$$
4) $$\mathsf{E}\left[\bar{\mathsf{X}}_{k,j}\right] = \sqrt{E_c} \tanh\left(\frac{\sum_{l \neq j}^{r_0} \Lambda_{k,l}}{2} \right) \tag{19}$$
5) go to step 2

where $\sqrt{E_c}$ is the chip energy, $\Lambda_{k,j}$ is the LLR of the j^{th} chip of user k, and $\bar{\mathsf{X}}_{k,j}$ is the corresponding extrinsic soft chip information. The channel coefficient for user k is denoted by h_k.

Equations (17) and (18) result from strictly applying MAP hypothesis testing to the system model (16). In the first step, the prior information is set to neutral (i.e. the LLR=0). The second step computes the signal variance after subtracting the estimates of the other users (soft cancellation). With increasing reliability of the estimates $\mathsf{E}[\Lambda_{l,j}]$, the variance of the remaining signal will converge towards the channel noise power σ^2, which is hoped to be achieved by carrying out steps 2-4 iteratively.

The bracketed expression in equation (18) is essentially the interference cancellation step for chip j, where an estimate of all but the k^{th} user is subtracted from the input signal R_j. This is somewhat similar to the bracketed expression in (17), just that this time amplitudes rather than energies are considered.

In step 4 finally, the extrinsic soft bit information is computed. The sum in the denominator of (19) is the (extrinsic) de-spreading operation, and the $\tanh$ expression is just a convenient way to compute the expectation for a given LLR value.

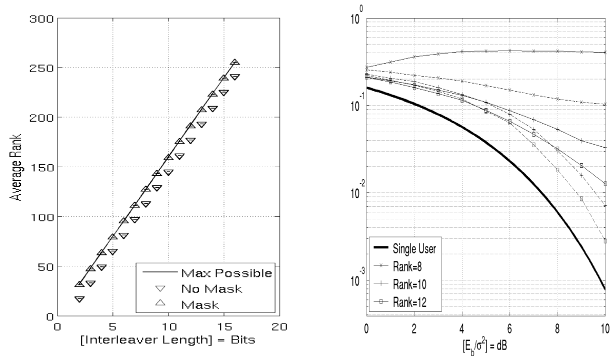
There are generally two ways how to schedule steps 2-4: either compute (17) for all users and chips, then (18) for all users and chips and finally (19) for all users and chips. Alternatively, fix the user k, compute (17) to (19) for all chips, and repeat this step until all users were processed. The first approach is referred to as parallel interference cancellation (PIC), the latter as serial interference cancellation (SIC). For both methods the first receiver iteration is outlined below:

Parallel IC	Serial IC
$\forall k, j : \Lambda_k, j = 0$	$\forall k, j : \Lambda_k, j = 0$
	For $k = 1 \ldots n$ do
$\forall k, j$: Compute ς_k, j	$\forall j$: Compute ς_k, j
$\forall k, j$: Compute Λ_k, j	$\forall j$: Compute Λ_k, j
$\forall k, j$: Compute $\mathsf{E}\left[\bar{\mathsf{X}}_{k,j}\right]$	$\forall j$: Compute $\mathsf{E}\left[\bar{\mathsf{X}}_{k,j}\right]$

V. Repetition Coding

In Section II it was shown how to represent any synchronous IDMA transmitter by a matrix model. Here we focus on the structure of the system matrix G induced by repetition coding and interleaving. The motivation for this investigation is to shed some light on a yet unsolved performance issue when using parallel soft interference cancellation without a mask, and to show a relationship to the matrix rank.

In terms of the system matrix G, the only notable difference between the two schemes is that $\mathsf{G} \in \{0, +1\}$ without a mask and $\mathsf{G} \in \{-1, 0, +1\}$ with a mask. For instance, the system matrix from the example in Section II-A, with and without flip



(a) Average Rank (b) BER

Fig. 3. Left: average rank of systems with and without a flip mask for $n = 16$ user and different interleaver lengths. Right: performance of the optimal (dashed lines) and iterative decoder (continuous lines) for $n = 4$ users and interleaving over $l_b = 3$ bits (12 chips).

mask is

$$\mathsf{G}_{\text{flip}} = \begin{bmatrix} +1 & 0 & +1 & 0 \\ -1 & 0 & 0 & -1 \\ 0 & +1 & 0 & +1 \\ 0 & -1 & -1 & 0 \end{bmatrix} \qquad \mathsf{G} = \begin{bmatrix} +1 & 0 & +1 & 0 \\ +1 & 0 & 0 & +1 \\ 0 & +1 & 0 & +1 \\ 0 & +1 & +1 & 0 \end{bmatrix}.$$

With the additional {-1} element the matrix is well balanced and as shown by simulation, the average matrix rank is significantly higher. Note that the equalities (7) to (9) are not affected by the additional element.

In Figure 3(a) the average rank of G for $n = 16$ users and code rate $r = 1/16$ is depicted. The height of the horizontal bar at each marker shows the standard deviation which is negligible small. The empirical results indicate that the flip mask almost surely produces matrices that are rank deficient by at most 1 (i.e. $\rho \geq n - 1$), whereas neglecting the flip-mask results in severely rank deficient matrices.

Figure 3(b) compares the performance of the optimal decoder (15) to that of the iterative receiver from [2]. Due to the exponential complexity of the optimal receiver with respect to the interleaver length, only a very small system with $n = 4$ users, code rate $r = 1/4$ and interleaving over $l_b = 3$ bits (12 chips) was considered, which results in a $\dim(\mathsf{G}) = 12 \times 12$ matrix. To achieve matrices with various rank numbers, randomly picked masks were used here. The general trend is that the performance degrades with the rank and that the optimal decoder always outperforms the iterative approach (irrespective of whether SIC or PIC is used).

Note that for a given rank ρ, the results in Figure 3(b) represent the performance of just a single matrix instantiation.

In this section we showed that for small matrices, rank and performance are linked. Such systems have only limited practical relevance however, and in Section VI it is shown that for large matrices the rank is a neglectable performance criterion.

VI. Simulation Results

The performance and convergence rates under different conditions are shown by conventional BER as well as EXIT charts. All the figures in this section were obtained for

- repetition coding (spreading) only, no additional outer channel code
- no channel distortions, only AWGN noise
- fully loaded systems only, i.e. code rate of repetition code $r = 1/n$
- fully synchronised system (i.e.perfect frequency and timing knowledge)

A. Conventional BER Performance

In Figure 4 all four combinations of IDMA systems with and without flip-mask using SIC and PIC are depicted. PIC systems without a flip mask do not converge and therefore fail, whereas the other three combinations perform more or less equally well. Figures 5 and 6 show the BER convergence over iterations for SNR=4dB and SNR=10dB respectively. Neglecting the PIC system without a mask, the conclusion is that the other three approaches achieve the same BER. The convergence rate however differs, and especially for higher SNR values, the SIC approaches perform noticeably better. The vertical bars attached to each data point show the standard deviation and suggest that for randomly chosen interleavers the convergence rate is very stable. The same holds for the BER, but no explicit data plot is shown here.

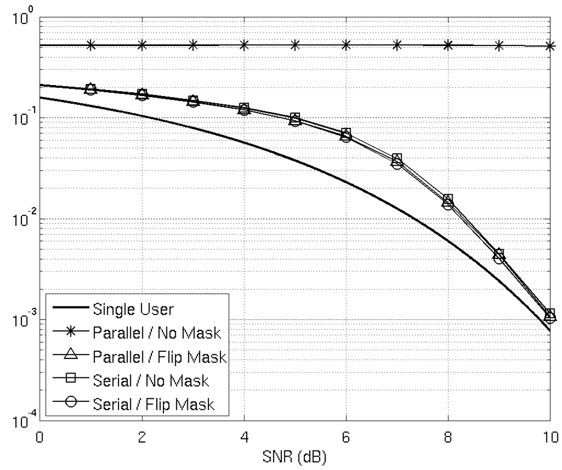


Fig. 4. BER for IDMA system after 10 receiver iterations for $n = 8$ users and interleaving over $l_b = 32$ bits

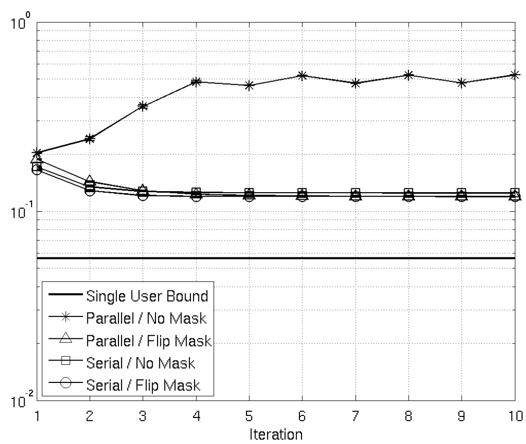


Fig. 5. Convergence for systems with $n = 8$ users, interleaver block length of $l_b = 32$ bits, SNR=4dB

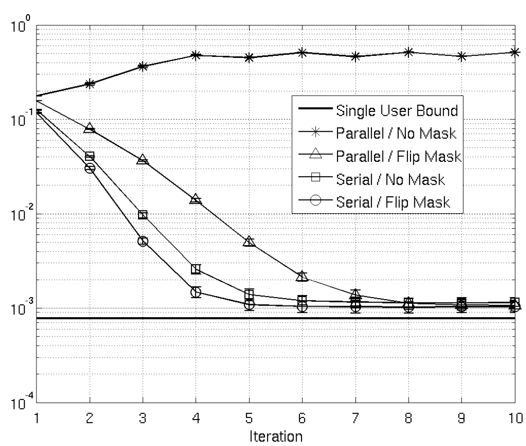


Fig. 6. Convergence for systems with $n = 8$ users, interleaver block length of $l_b = 32$ bits, SNR=10dB

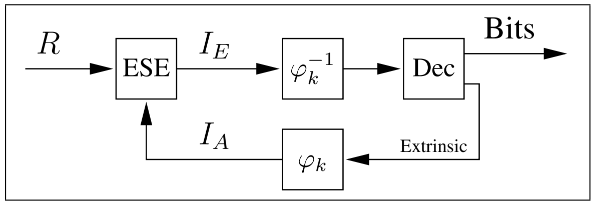


Fig. 7. Structure of iterative CBC receiver for user k. The ESE is also fed by the extrinsic information of all other users (not shown here).

B. EXIT Chart Analysis

Extrinsic information transfer (EXIT) charts as proposed in [10, 11] are another powerful tool to complement the BER analysis, as it allows for predicting and tracking the performance over iterations. Originally, EXIT charts were used to analyse the convergence behaviour of concatenated codes but its application can easily be extended to analyse generic turbo systems.

In IDMA, instead of two concatenated codes, the receiver consists of an elementary signal estimator (ESE) and a repetition decoder (de-spreader) block. The term ESE is used only to be consistent with the IDMA literature, as it is in fact the interference cancellation unit explained in Section IV.

EXIT charts relate mutual information between source bits and ESE output chips, to that between source bits and (extrinsic) decoder output bits. Figure 7 shows the iterative CBC model for user k, where R is the received data as in Figure 2, and I_E and I_A relate the mutual information of the data bits B to the ESE and decoder output respectively. Note that the position of the probes in Figure 7 is independent of whether the ESE operates in SIC or PIC mode. The predicted performance, though, changes insofar that for SIC the EXIT tunnel is broader than for PIC. Using PIC, all users are estimated in parallel first, and afterwards the prior information about all other users is updated simultaneously. In SIC this is different because the prior information is updated as soon as the first user was estimated, and consequently the second user possesses already more knowledge than the first user had and can do a more reliable interference cancellation, from which in turn the third user will benefit and so on. Hence the tunnel formed in the EXIT charts is broader for SIC than it is for PIC as will be shown shortly.

EXIT analysis of all four cases (PIC and SIC, with and without mask) is determined for a system with $n = 32$ users, code rate $r = 1/32$, interleaving over $l_b = 16$ bits (512 chips) and SNR=10dB. Figure 8(a) shows the convergence with PIC and flip-mask, and the actual performance stays tightly within the predicted bounds. Neglecting the mask as in Figure 8(b) causes the mutual information to spiral around the origin and shows no sign of convergence, hence the BER drops to 50% as was already shown in Figure 4 before.

Figures 8(c) and 8(d) show the same two cases, but this time using serial interference cancellation. As already mentioned, a notable difference to the PIC results is the bigger gap between predicted ESE and decoder performance, hence the receiver can gain more mutual information in a single iteration. This can be seen by comparing Figure 8(a) and 8(c), where the SIC converges significantly faster due to the larger steps. However, the crossover point is still approximately the same for both versions. These observations coincide with the BER performance shown in Figure 4 before, where SIC did not have any advantage over PIC when using a mask, except for faster convergence at high SNR.

Figure 8(d) finally shows the SIC case without a mask. Contrary to PIC, this system remains functional, has again the same predicted crossover point but shows slower convergence than with masking. The reason for this is the decoder performance lacking behind the prediction, and therefore more (smaller) steps are needed upon converge.

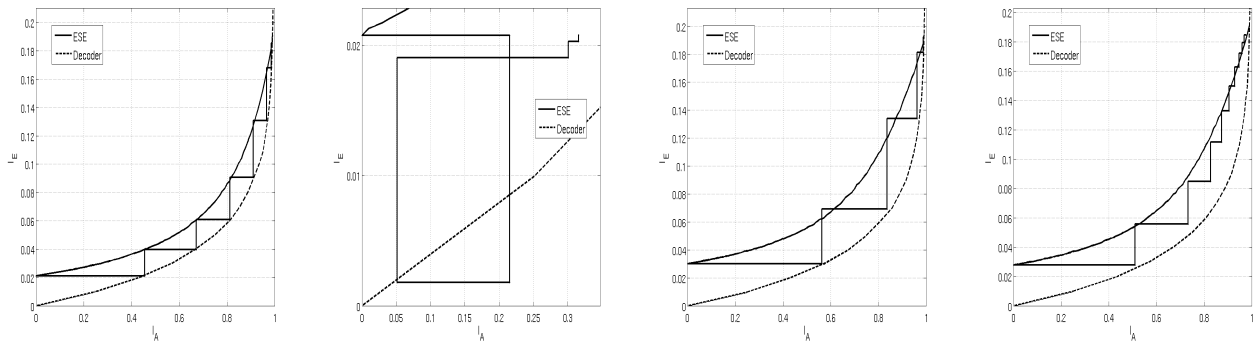
According to (6), the used system with $n = 32$ users and interleaving over $l_b = 16$ bits can be fully described by a $\dim(\mathsf{G}) = 512 \times 512$ matrix. One important difference in the matrix properties is the rank ρ. With a flip mask, the rank is almost always full (Figure 8(a) and 8(c)), but without masking the rank is significantly deficient (Figure 8(b) and 8(d)) and there is a chance that different input vectors map to the same output vector. Since the SIC performance does not suffer from the low rank, it can be concluded that such ambiguities occur rarely if at all. Consequently, the rank is hardly an IDMA performance criterion when the matrices are large, i.e. for any realistic system.

VII. Differences and Similarities between IDMA and Multiuser DS-CDMA

In this section we will show how IDMA and CDMA relate to each other, and that an IDMA receiver can be used to detect CDMA signals as well.

In Figure 9 the transmitter branch for a particular user k is depicted. The repetition encoder can be represented as a bit repetition and a masking unit, which in turn equals a conventional spreading operation in CDMA. The system model is then already translated to a conventional CDMA system with chip-level interleaving as in [1]. The only difference to the classical scheme are:

- each user has the same spreading code
- every user has a different chip level interleaver.



(a) PIC with mask (b) PIC without mask (c) SIC with mask (d) SIC without mask

Fig. 8. EXIT charts for both interference cancellation approaches with and without mask. The system supported $n = 32$ users, the interleaver block length was $l_b = 16$bit and the SNR=10dB.

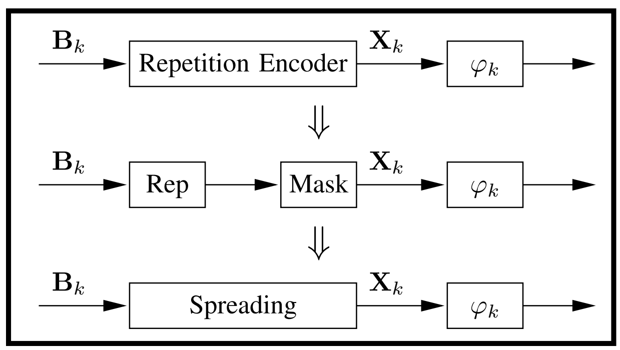


Fig. 9. Three equivalent representations of an uncoded IDMA transmitter branch for user k

Letting all users select the same spreading code makes the system simpler but can be neglected without contradicting the IDMA principle[3]. This allows us to assign each user an individual spreading code like in conventional CDMA, and it is always possible to find a chip interleaver configuration that resembles the traditionally used bit interleaving scheme. From this point of view, bit interleaved CDMA systems are a special case of IDMA.

The system model developed in Section II proves useful to illustrate the similarity. Equation 20a shows a CDMA system supporting two users $\mathcal{A}$ and $\mathcal{B}$ where each user transmits the three bits (A_1, A_2, A_3) and (B_1, B_2, B_3) respectively. The matrix columns contain the spreading codes associated with each bit, e.g. the first column belongs to bit A_1, the fifth column for bit B_2. The number of nonzero column entries determines the spreading factor and is 2 in this example. Interleaving the data bits is achieved by swapping the columns accordingly and a particular result is given in (20b), where the second and third bit were swapped.

$$\mathbf{Z}_\mathrm{C} = \begin{bmatrix} +1 & 0 & 0 & | & +1 & 0 & 0 \\ +1 & 0 & 0 & | & -1 & 0 & 0 \\ 0 & +1 & 0 & | & 0 & +1 & 0 \\ 0 & +1 & 0 & | & 0 & -1 & 0 \\ 0 & 0 & +1 & | & 0 & 0 & +1 \\ 0 & 0 & +1 & | & 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ B_1 \\ B_2 \\ B_3 \end{bmatrix} \tag{20a}$$

$$\Downarrow \text{Bit Interleaving} \Downarrow$$

$$\mathbf{Z}_\mathrm{C} = \begin{bmatrix} +1 & 0 & 0 & | & +1 & 0 & 0 \\ +1 & 0 & 0 & | & -1 & 0 & 0 \\ 0 & 0 & +1 & | & 0 & 0 & +1 \\ 0 & 0 & +1 & | & 0 & 0 & -1 \\ 0 & +1 & 0 & | & 0 & +1 & 0 \\ 0 & +1 & 0 & | & 0 & -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ B_1 \\ B_2 \\ B_3 \end{bmatrix} \tag{20b}$$

For IDMA the process is similar. Equation (21a) shows the un-interleaved system with a flip mask. After applying individual chip interleavers, the elements within the columns are permuted and a possible outcome is (21b), where the first user happened

[3]This represents a system flexible system possible – individual interleavers and individual spreading codes at the same time

to end up with a neutral interleaver not doing anything.

$$\mathbf{Z}_{\mathrm{I}} = \begin{bmatrix} +1 & 0 & 0 & | & +1 & 0 & 0 \\ -1 & 0 & 0 & | & -1 & 0 & 0 \\ 0 & +1 & 0 & | & 0 & +1 & 0 \\ 0 & -1 & 0 & | & 0 & -1 & 0 \\ 0 & 0 & +1 & | & 0 & 0 & +1 \\ 0 & 0 & -1 & | & 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ B_1 \\ B_2 \\ B_3 \end{bmatrix} \tag{21a}$$

$$\Downarrow \text{Chip Interleaving} \Downarrow$$

$$\mathbf{Z}_{\mathrm{I}} = \begin{bmatrix} +1 & 0 & 0 & | & +1 & 0 & 0 \\ -1 & 0 & 0 & | & 0 & 0 & -1 \\ 0 & +1 & 0 & | & 0 & -1 & 0 \\ 0 & -1 & 0 & | & 0 & 0 & +1 \\ 0 & 0 & +1 & | & -1 & 0 & 0 \\ 0 & 0 & -1 & | & 0 & +1 & 0 \end{bmatrix} \cdot \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ B_1 \\ B_2 \\ B_3 \end{bmatrix} \tag{21b}$$

To transform (21b) into (20b), the chip interleavers must be chosen such that the nonzero elements form the block structure again, and the mask sequences of user $\mathcal{A}$ in IDMA must be changed to the pattern $(+1, +1)$. It is therefore possible to represent any bit-interleaved CDMA system by an IDMA system by properly choosing the chip interleavers and masking sequences for each user.

The apparent difference between (20b) and (21b) is that CDMA blocks the nonzero elements together, whereas IDMA distributes them randomly. The implication is that in IDMA the individual bits are interconnected with much more bits than in CDMA. In example (20b) above, bit A_1 is only connected with bit B_1, A_2 only with B_2 and so on. Consequently, the receiver can only use knowledge of A_2 to detect B_2, and knowledge about other bits cannot aid in detection.

For bit interleaved CDMA systems it is also always possible to partition (20b) with the the vector $\mathbf{Z}_{\mathrm{C}} = (Z_1, Z_2, Z_3, Z_4)$ into the three smaller and independent equations

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \begin{bmatrix} +1 & +1 \\ +1 & -1 \end{bmatrix} \begin{bmatrix} A_1 \\ B_1 \end{bmatrix} \tag{22}$$

$$\begin{bmatrix} Z_3 \\ Z_4 \end{bmatrix} = \begin{bmatrix} +1 & +1 \\ +1 & -1 \end{bmatrix} \begin{bmatrix} A_3 \\ B_3 \end{bmatrix} \tag{23}$$

$$\begin{bmatrix} Z_5 \\ Z_6 \end{bmatrix} = \begin{bmatrix} +1 & +1 \\ +1 & -1 \end{bmatrix} \begin{bmatrix} A_2 \\ B_2 \end{bmatrix}. \tag{24}$$

These matrices do not have any nonzero components anymore and are typically associated with CDMA. The effective matrix size in this case is therefore only 2×2 anymore and not 6×6.

In the IDMA system (21b) this is different. Here bit A_1 is connected with B_1 and B_3, and those bits are in turn connected with yet other bits forming a network which can pass along extrinsic information. This is similar to linear coding, where the parity bits result from many different interconnection of the data bits to allow the decoder to pass on extrinsic information within a large and highly interconnected network. Also note that in general it is not possible anymore to partition the equation into smaller ones, and the effective matrix size here is indeed 6×6.

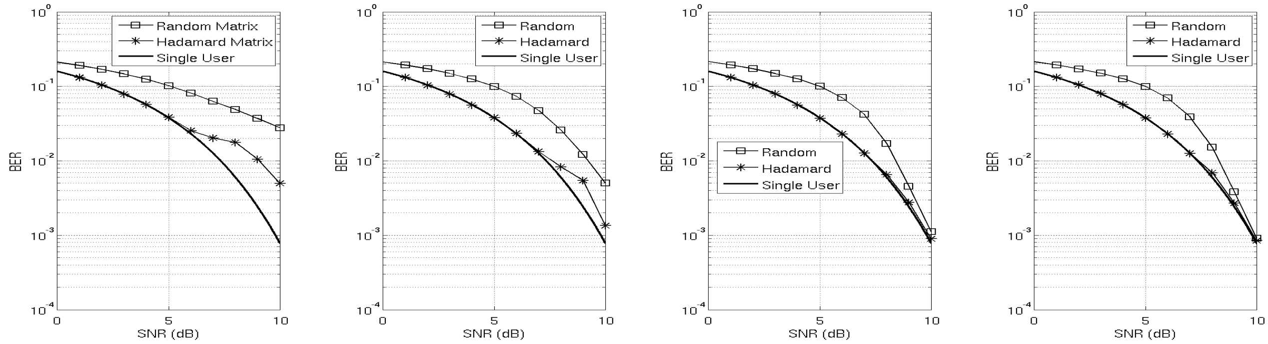
After showing that any bit interleaved CDMA system can be made up of an IDMA system with properly chosen masks and interleavers, Figure 10 depicts the performance of an iterative IDMA receiver when applied to a CDMA signal. The shown results are averages over multiple randomly created systems and show the performance for orthogonal and randomly created spreading codes. The orthogonal spreading sequences were included merely out of curiosity because there are simpler ways to deal with them and reach the single user bound. The number of users n specify the matrix size used for the simulation, i.e. $\mathsf{G} \in \{-1, +1\}^{n \times n}$. The general results are:

- CDMA signals can be detected with an IDMA receiver
- the performance improves with the matrix size.

Interestingly, for orthogonal matrices the receiver is unable to achieve single user performance at higher SNR values, but yet again, the effect diminishes with larger matrix sizes. The fact that the effective matrix size dominates the performance is the main reason why the IDMA receiver is inappropriate for practical CDMA systems, because spreading codes are usually rather short, leading to small matrices. In IDMA on the other side, the matrix size is mainly a function of the interleaver length and independent of the used spreading gain (see Section II-B).

A. Issues when Comparing IDMA and CDMA

In [5] the performance of an IDMA receiver for uncoded CDMA systems was briefly investigated and the authors showed that for a given number of users n and code rate r the performance of IDMA is far superior over CDMA. This however is a



(a) 8 Users (b) 64 Users (c) 256 Users (d) 512 Users

Fig. 10. BER results when detecting CDMA signals with the iterative IDMA receiver. All systems were fully loaded, i.e. the spreading gain matches the number of users

contradiction to the results shown above and is due to the effective matrix size used.

The IDMA system in [5] used a code rate of $r = 1/64$ and a 256 chips long random interleaver, and to make a reasonable comparison, the CDMA system used the same parameters. Considering the fully loaded case only where $n = 1/r = 64$, the CDMA matrix has dimension $\dim(\mathsf{G}_{\text{CDMA}}) = 64 \times 64$, while according to (6), IDMA uses a $\dim(\mathsf{G}_{\text{IDMA}}) = 256 \times 4n = 256 \times 256$ matrix. The IDMA matrix was therefore sixteen times larger than the CDMA matrix, and in the last section we showed how the performance improves with the matrix size. This comparison is equivalent to comparing Figures 10(b) and 10(c) with each other, where the latter performs significantly better in the high SNR region.

As an example for two equivalent systems, Figure 10(c) shows a CDMA system that has the same matrix size as the IDMA system used for Figure 4, and here the BER performance matches.

VIII. Conclusion

We showed how synchronised IDMA can be described by a system matrix G and translated the IDMA block diagram to the completely generic and very compact system model $\mathbf{Z} = \mathsf{G}\mathbf{B}$, with $\mathbf{B}$ being the data bits of all users. The new matrix model allows for additional analysis methods based on (sparse) matrix theory, helping to investigate and optimise IDMA further. It was shown that the resulting matrices are very sparse, and the percentage of non-zero elements declines with the inverse of the interleaver block length.

With the matrix model we derived the optimal decoding rule for IDMA which has exponential complexity unfortunately, and is of limited practical use only. For IDMA, however, a powerful low complexity iterative receiver exists and we outlined the differences between SIC and PIC with it. Using BER and EXIT chart analysis we showed that systems with alternating spreading codes using SIC converge the fastest. The point of convergence is the same for all approaches except PIC without masking, which does not converge at all. We therefore propose to always use a mask to assure convergence, irrespective of the used interference cancellation approach.

We briefly investigated the difference of matrices with and without mask and found that the latter suffer from a heavy loss in rank, whereas systems with a mask are almost surely rank deficient by at most 1. Using the optimal detector, the loss in rank was shown to be a problem for small matrices only. For a realistic system with reasonably large matrices the loss in rank, caused by neglecting the mask, does not lead to performance degradation when using SIC.

Finally, IDMA was compared to CDMA and we showed that an IDMA receiver can be used to detect CDMA signals, and the performance improves with the matrix dimension (for both, IDMA and CDMA).

References

[1] R. Mahadevappa and J. Proakis, "Mitigating multiple access interference and intersymbol interferece in uncoded CDMA systems with chip-level interleaving," *IEEE Trans. on Wireless Commun.*, vol. 1, no. 4, pp. 781–792, Oct. 2002.

[2] L. Ping, "Interleave-division multiple access and chip-by-chip iterative multi-user detection," *IEEE Commun. Magazine*, pp. S19–S23, Jun. 2005.

[3] P. A. Hoeher and H. Schoenreich, "Interleave-division multiple access from a multi user theory point of view," in *Turbo Coding 2006 Conference*, Munich, Germany, 3-7 April 2006.

[4] L. Ping, L. Liu, K. Wu, and W. Leung, "On interleave-division multiple-access," in *IEEE International Conference on Communications*, vol. 5, 20-24 June 2004, pp. 2869–2873.

[5] L. Ping, L. Liu, and W. Leung, "A simple approach to near-optimal multiuser detection: interleave-division multiple-access," *Wireless Communications and Networking Conference*, vol. 1, pp. S391–S396, 16-20 March 2003.

[6] L. Ping, L. Liu, K. Y. Wu, and W. K. Leung, "Interleave-division multiple-access," 2005, submitted to IEEE Trans. on Wireless Commun.

[7] L. Ping and L. Liu, "Analysis and design of IDMA systems based on SNR evolution and power allocation," in *IEEE Veh. Tech. Conference*, Sept. 2004, pp. 1068–1072.

[8] K. Li, X. Wang, and P. Li, “Analysis and optimization of interleave-division multiple-access communication systems,” *Acoustics, Speech, and Signal Processing, Proceedings, ICASSP ’05*, vol. 3, p. S917, 18-23 March 2005.

[9] H. Wu, L. Ping, and A. P. A, “User-specific chip-level interleaver design for IDMA systems,” *Electronics Letters*, vol. 42, pp. S233–S234, Februar 2006.

[10] S. ten Brink, “Convergence behavior of iteratively decoded parallel concatenated codes,” *IEEE Transactions on Communications*, vol. 49, no. 10, pp. 1727–1737, Oct 2001.

[11] F. Brannstrom, L. Rasmussen, and A. Grant, “Convergence analysis and optimal scheduling for multiple concatenated codes,” *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3354–3364, Sep 2005.