Efficient Generation of Interleavers for IDMA

Ioachim Pupeza*, Aleksandar Kavčić† and Li Ping††
*Technical University Braunschweig, Germany (Visiting Fellow at Harvard University)
†Department of Electronic Engineering, City University of Hong Kong (on leave from Harvard University)
††Department of Electronic Engineering, City University of Hong Kong

Abstract— **We consider the design of practical interleavers for interleaver division multiple access (IDMA) systems. A set of interleavers is considered to be practical if it satisfies two criteria: 1) It is easy to generate (i.e., the transmitter and receiver need not store or communicate many bits in order to agree upon an interleaver), and 2) no two interleavers in the set "collide". We show that a properly defined correlation between interleavers can be used to formulate a collision criterion, where zero-correlation (i.e., orthogonality) implies no collision. Computing the correlation among non-orthogonal interleavers is generally computationally very expensive, so we also design an upper-bounding technique to efficiently check whether two interleavers have low correlation. We then go on to propose several methods to design practical interleavers for IDMA: one method to design orthogonal interleavers, and two methods to design non-orthogonal interleavers (where the upper-bounding technique is used to verify their cross-correlation is low). Simulation results are presented to show that the designed practical interleavers perform as well as random interleavers in an IDMA system.**

Index Terms— **IDMA, orthogonal interleavers, correlation between interleavers**

I. Introduction

Interleaver division multiple access (IDMA) is a technique that relies on different interleavers to separate signals from different users in a multiuser spread-spectrum communication system. In [1], an IDMA system that uses randomly and independently generated interleavers is presented. With these interleavers, the IDMA system in [1] performs similarly and even better than a comparable CDMA system.

The condition for IDMA to be successfully implemented is that the transmitter and receiver agree upon the same interleaver. For random interleavers, the entire interleaver matrix has to be transmitted to the receiver, which can be very costly. Our goal is to construct non-random interleavers for IDMA that perform as well as random interleavers and satisfy two design criteria:

- They are easy to specify and generate, i.e., the transmitter and receiver can send a small number of bits between each other in order to agree upon an interleaver, and then generate it.
- The interleavers do not "collide".

Organization: Section II contains an introduction of the IDMA communication system. In Section III, we explain what it means that interleavers do not collide. We define the correlation between two interleavers as a measure of "how strong two interleavers collide". We define orthogonal interleavers as interleavers with zero correlation and we bound the number of orthogonal interleavers. In Section IV, we present three families of interleavers that match the design criteria. Section V presents computer simulations of the IDMA system with the constructed interleavers. Section VI concludes the paper.

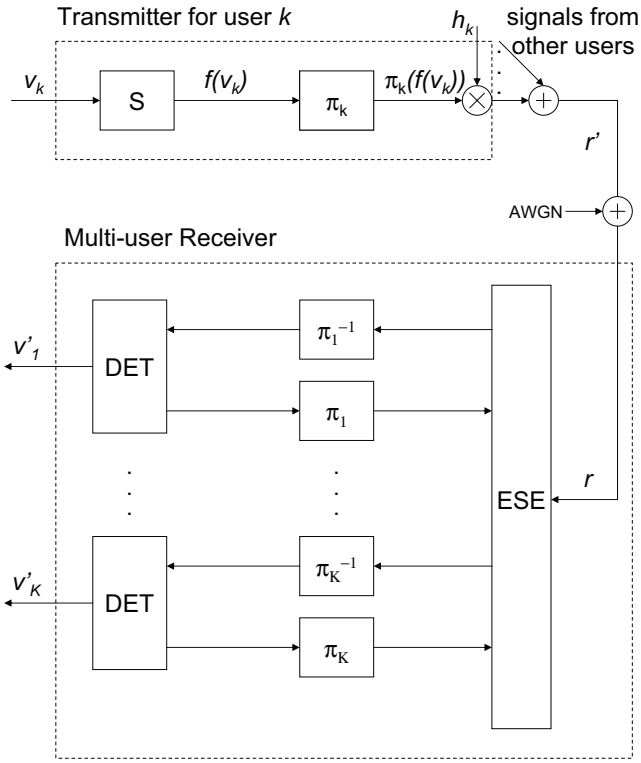

Fig. 1. IDMA system scheme over an AWGN channel.

Notation: In this paper, vectors are denoted as $v = (v[1], v[2], \ldots, v[n])^T$, where $v[j]$ denotes the j-th component of the vector v for $j \in \{1, 2, \ldots, n\}$. By an interleaver, we mean a bijective map, that maps every vector to a permuted version of itself. By K_{max} we denote the maximal number of users allowed to communicate simultaneously in a multi-user communication system.

II. IDMA

We use the uncoded IDMA system described in [1], see Figure 1. There are K users in the system ($K \leq K_{max}$). User k transmits a bipolar input vector $v_k \in \{-1, 1\}^{\ell}$,

1-4244-0355-3/06/$20.00 (c) 2006 IEEE

where ℓ is the *block length* and $1 \leq k \leq K$. We call every vector $w \in \{-1,1\}^{\ell}$ a *word*, and let W be the *set of all words*. The spreading operation means that each bit (symbol) of user k's vector v_k multiplies a certain spreading sequence of length S (usually $S \geq 64$). The spreading sequence in IDMA is the same for all users. Usually it is $(1,-1,+1,-1,...,+1,-1)$, i.e., an alternating sequence of $+1$ and -1, of length S. Let the function $f(\cdot)$ be the mapping describing the spreading process. Thus, the obtained vector $f(v_k)$ has length ℓS. The interleaver π_k permutes the bits of $f(v_k)$, which yields a vector $\pi_k(f(v_k))$. The channel linearly combines the signals form all K users as $r = \sum_{k=1}^{K} h_k \pi_k(f(v_k)) + n$, where h_k is the channel coefficient, and n is a vector of additive white Gaussian (AWGN) noise samples, see Figure 1.

In [1], a multi-user receiver for IDMA is proposed that uses the turbo decoding principle (see [3]). The receiver consists of an elementary signal estimator (ESE), and K branches for the K users (see Figure 1). The ESE is responsible for estimating the signals in every iteration of the turbo decoding process. Each branch k uses the same interleaver π_k as the corresponding branch of the transmitter, and an a posteriori probability (APP) detector (DET), which is identical for every user.

III. Correlation Between Interleavers

A. Motivation

Since the separation of users is achieved by interleavers, an obvious interleaver design criterion is that every two interleavers out of a set of interleavers "collide" as little as possible. The goal in this section is to define correlation among interleavers for IDMA in order to measure the level of "collision" among interleavers.

Unlike in classical turbo coding/decoding (see [3]), where the task of a single interleaver is to decorrelate different sequences of bits, here we have a set of interleavers, that not only need to decorrelate different bit sequences, but also different users. The correlation between interleavers should measure how strongly signals from other users affect the decoding process of a specific user. Hence, the additive noise should not play a role in the correlation of interleavers, and throughout this section, we consider the noiseless IDMA system. In that case, a non-turbo decoder depicted in Figure 2 suffices, where the decoder for user j consists of the user-specific deinterleaver π_j^{-1} and a despreader (DES).

B. Definition of Correlation and Orthogonal Interleavers

Definition 1: Let π_i and π_j be two interleavers and let w and v be two words. We define the *correlation* $\mathrm{C}(\pi_i, w, \pi_j, v)$ between π_i and π_j with respect to the words w and v as the scalar product between $\pi_i(f(w))$ and $\pi_j(f(v))$:

$$\mathrm{C}(\pi_i, w, \pi_j, v) = \langle \pi_i(f(w)), \pi_j(f(v)) \rangle . \qquad (1)$$

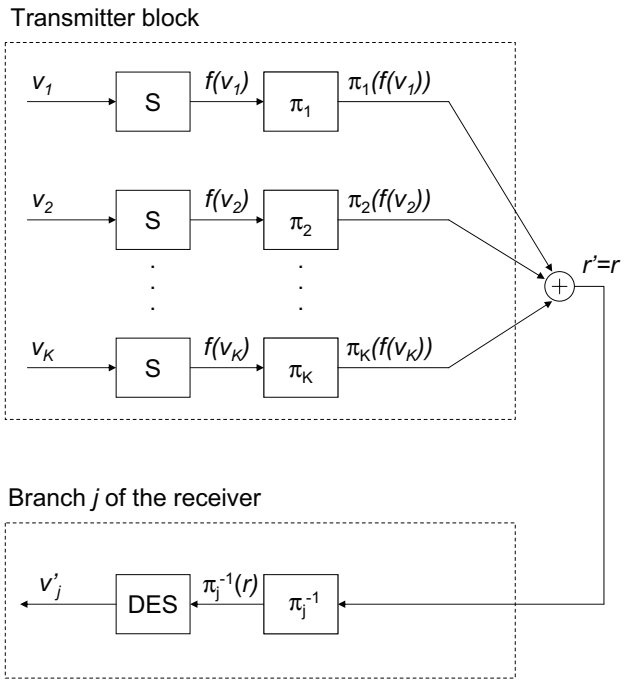

Fig. 2. Noiseless IDMA system.

Definition 2: Two interleavers π_i and π_j (where $\pi_i \neq \pi_j$) are called *orthogonal*, if for any two words w and v, we have

$$\mathrm{C}(\pi_i, w, \pi_j, v) = \langle \pi_i(f(w)), \pi_j(f(v)) \rangle = 0. \qquad (2)$$

It is easy to verify that if a set of mutually othogonal interleavers is used in the IDMA system, then the decoder in Figure 2 perfectly dedodes user j, i.e., $v'_j = v_j$. In this sense, zero-correlation (or orthogonality) implies no "collision" among interleavers.

C. Bound on the Number of Orthogonal Interleavers

Theorem 1: Let S be the spreading length. For any block length ℓ, a set of orthogonal interleavers has at most S elements, i.e., the number of orthogonal interleavers is at most S.

The proof is given in Appendix A.

D. Bounding the Correlation between Interleavers

We have shown that it is impossible to find a set of more than S orthogonal interleavers. If we want to build an IDMA system that allows more than S simultaneous users, we need to use interleavers with non-zero correlation. However, evaluating the correlation between two interleavers with respect to every possible pair of two words is very computationally complex. This is because there are 2^{ℓ} possibilities to choose the first word and other 2^{ℓ} possibilities to choose the second word. In this section we suggest a method for upper bounding the correlation between interleavers.

For two "good" interleavers, the correlation term in (1) should be close to 0. For $i \neq j$ or $w \neq v$, this is equivalent to minimizing the magnitude

$$|\mathrm{C}(\pi_i, w, \pi_j, v)| = |\langle \pi_i(f(w)), \pi_j(f(v))\rangle| \,. \quad (3)$$

In order to find upper bounds for (3), some definitions are helpful. From now on, we assume that $i \neq j$ or $w \neq v$ and $\ell \geq 3$.

Definition 3: Let $\ell \in \mathbb{N}$. The *canonical basis* of $\mathbb{R}^\ell$ is the set of basis vectors e_i

$$\{e_i : e_i[i] = 1, e_i[j] = 0, i \in \{1, 2, \ldots, \ell\}, j \neq i\} \,.$$

Definition 4: Let $\ell \in \mathbb{N}$. A *generating set* of $W \subset \mathbb{R}^\ell$ is a set of ℓ vectors, such that every word in W can be written as a linear combination of the elements in the generating set.

Definition 5: Let the set $W_g = \{w_1, w_2, \ldots, w_\ell\}$ be defined as follows

$$W_g = \{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix}, \ldots, \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \\ \vdots \\ -1 \\ 1 \end{pmatrix} \}. \quad (4)$$

In words, w_1 is the all-ones vector; for $i \geq 2$, the first $i-1$ components of w_i are *-1* and all other components of w_i are *1*.

Proposition 1: W_g is a generating set of W.

Proposition 1 can be proved by showing that a matrix whose columns are the vectors w_i, has an inverse.

We now turn our attention to the following correlation value for $w_n \in W_g$

$$|\mathrm{C}(\pi_i, w, \pi_j, w_n)| = |\langle \pi_i(f(w)), \pi_j(f(w_n))\rangle| \,. \quad (5)$$

(Note that the second word in (5) is an element of W_g.)

Definition 6: Let π_i and π_j be two interleavers, $w \in W$ and $w_n \in W_g$. We call $\mathrm{C}(\pi_i, w, \pi_j, w_n)$ the *basis correlation* between π_i and π_j with respect to w and the basis word w_n.

The word w can be written as $w = \sum_{m=1}^{\ell} \alpha_m e_m$, where $e_1, e_2, \ldots, e_\ell$ build the canonical basis of $\mathbb{R}^\ell$ and $\alpha_m \in \{-1, +1\}$ for all $m \in \{1, 2, \ldots, \ell\}$. Then (5) becomes

$$|\mathrm{C}(\pi_i, w, \pi_j, w_n)| = \left| \sum_{m=1}^{\ell} \alpha_m \langle \pi_i(f(e_m)), \pi_j(f(w_n))\rangle \right| . \quad (6)$$

Using the triangle inequality, we bound (6) as

$$\begin{aligned} |\mathrm{C}(\pi_i, w, \pi_j, w_n)| &\leq \sum_{m=1}^{\ell} \underbrace{|\alpha_m|}_{1} \cdot |\langle \pi_i(f(e_m)), \pi_j(f(w_n))\rangle| \\ &= \sum_{m=1}^{\ell} |\langle \pi_i(f(e_m)), \pi_j(f(w_n))\rangle| \,. \quad (7) \end{aligned}$$

Proposition 2: The inequality *(7)* represents a tight upper bound, i.e. for any two interleavers π_i and π_j and for every $w_n \in W_g$, there exists a word $w \in W$, such that the bound *(7)* is met with an equality.

Proof of Proposition 2: Setting either $\alpha_m = \mathrm{sgn}(\langle \pi_i(f(e_m)), \pi_j(f(w_n))\rangle)$, for all $m \in \{1, 2, \ldots, \ell\}$, or $\alpha_m = -\mathrm{sgn}(\langle \pi_i(f(e_m)), \pi_j(f(w_n))\rangle)$, for all $m \in \{1, 2, \ldots, \ell\}$ will yield equality in (7).

Definition 7: Let π_i and π_j be two interleavers. The *peak basis correlation* between π_i and π_j is denoted by $P(\pi_i, \pi_j)$ and defined as

$$P(\pi_i, \pi_j) = \max_{w_n \in W_g} \sum_{m=1}^{\ell} |\langle \pi_i(f(e_m)), \pi_j(f(w_n))\rangle| \,. \quad (8)$$

With this definition, Proposition 2 immediately yields:

Proposition 3: For every $w \in W$ and every $w_n \in W_g$, the upper bound

$$|\mathrm{C}(\pi_i, w, \pi_j, w_n)| \leq \mathrm{P}(\pi_i, \pi_j) \quad (9)$$

is tight, i.e. for any two interleavers π_i and π_j and for every $w_n \in W_g$, there exists a word $w \in W$, such that $|\mathrm{C}(\pi_i, w, \pi_j, w_n)| = \mathrm{P}(\pi_i, \pi_j)$.

Definition 8: Let π_i and π_j be two interleavers. The *worst case correlation* between π_i and π_j is denoted by $W(\pi_i, \pi_j)$ and defined as

$$W(\pi_i, \pi_j) = \sum_{m,n=1}^{\ell} |\,\langle \pi_i(f(e_m)), \pi_j(f(w_n))\rangle\,| . \quad (10)$$

Proposition 4: For every $w \in W$ and every $v \in W$

$$|\mathrm{C}(\pi_i, w, \pi_j, v)| \leq \mathrm{W}(\pi_i, \pi_j). \quad (11)$$

We omit the proof of Proposition 4 due to spatial constraints.

Note that the computational complexities of computing the bounds in (9) and (11) are the same since the outer sum in (11) is replaced by the $\max$ operation in (9). Further, by combining Propositions 3 and 4, we get

$$|\mathrm{C}(\pi_i, w, \pi_j, v)| \leq \mathrm{W}(\pi_i, \pi_j) \leq \ell \cdot \mathrm{P}(\pi_i, \pi_j).$$

Hence, either the worst case correlation $W(\cdot, \cdot)$ or the peak basis correlation $\mathrm{P}(\cdot, \cdot)$ can be used to bound the correlation between two interleavers.

IV. Interleaver Design

The following definition is necessary in this section:

Definition 9: Let a and b be two arbitrarily chosen vectors of length ℓS and π an interleaver, such that $\pi(a) = b$. We define the bijective permutation map $\Pi : \{1, 2, \ldots, \ell S\} \to \{1, 2, \ldots, \ell S\}$, such that $a[i] = b[\Pi(i)]$ for all $i \in \{1, 2, \ldots, \ell S\}$.
Note that π and Π describe the same permutation.

A. Orthogonal Interleavers

According to Theorem 1, it is impossible to construct more than S orthogonal interleavers. In this section we will propose a method to construct a set of $S - 1$ orthogonal interleavers based on orthogonal binary sequences. A way to construct orthogonal sequences is

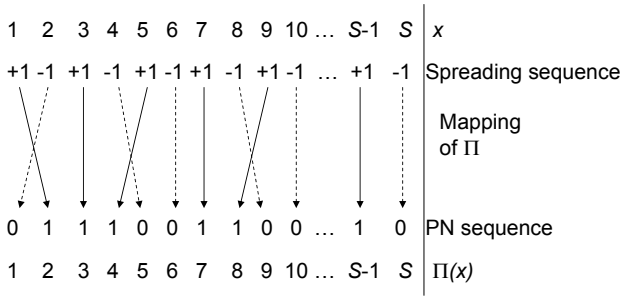

Fig. 3. Construction of an interleaver based on a PN sequence. In this example, $\Pi(5) = 4$ and $\Pi(8) = 9$.

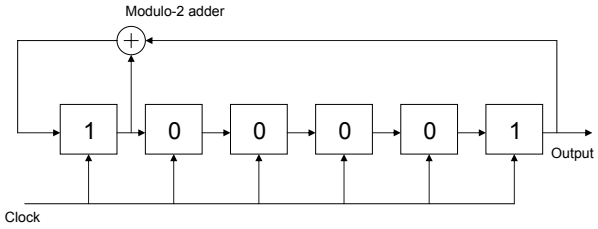

Fig. 4. Linear feedback shift register for $x^6 + x + 1$

by using *pseudonoise (PN) sequences* generated by a linear feedback shift register. For a comprehensive treatment of PN sequences, see [2].

Algorithm:

- Generate a PN sequence of length $S - 1 = 2^m - 1$ for some integer m using a linear feedback shift register with the connections defined by a primitive polynomial of degree m over the Galois field GF(2). Figure 4 shows the initial state of the linear feedback shift register that we used for our simulations. Both the feedback connections and the initial loading correspond to the coefficients of the primitive polynomial $x^6 + x + 1$.
- Store all $S - 1$ shifts of this sequence.
- Append a 0 at the end of each shift of the initial sequence such that the obtained binary sequences are orthogonal. This property is given by the balance of 1s and 0s within PN sequences (see [2]).
- For each sequence $i \in \{1, 2, \ldots, S - 1\}$ do:
  - For x from 1 to S do:
    if the x-th element of the spreading sequence is +1, let Π_i map x to the next free place in the PN sequence that contains a 1. Else, let Π_i map x to the next free place in the PN sequence that contains a 0 (see Figure 3).
  - For n from 1 to ℓ and x from 1 to S do:

$$\Pi_i(n \cdot S + x) = n \cdot S + \Pi_i(x). \tag{12}$$

The striking advantage of orthogonal interleavers is the compressibility. For example, by specifying only 6 bits, namely the coefficients of the polynomial generating the PN sequence, a mobile station can generate one of 63 ($63 = 2^6 - 1$) orthogonal interleavers as soon as it receives a number between 1 and 63 from the base-station. This works for any block length ℓ. Furthermore, after defining the mapping within the period of the interleaver, we could scramble the periodic segments (12) among themselves. In this case, the property of orthogonality remains preserved, because every two spreading blocks are mapped by two different interleavers to two orthogonal sequences, respectively. This may be useful if other constraints (such as multi-path propagation) are given.

B. Pseudo Random Interleavers

According to Theorem 1, for the implementation of more than S simultaneous users, a family of non-orthogonal interleavers is needed. A different approach than in Section IV-A to finding "good" interleavers for the uncoded IDMA system is to use the low-correlation property of long PN sequences. The correlation between these interleavers can be efficiently bounded using the peak basis correlation $\mathrm{P}(\cdot, \cdot)$ introduced in Section III-D.

Algorithm:

- We chose K primitive polynomials of degree m over the Galois field GF(2), such that $\ell S = 2^m$ for some integer m. The maximal number of users, K_{max}, may be larger than the spreading length S. All such polynomials are listed in [5] up to a large degree.
- Each of the K_{max} interleavers of length ℓS is generated by its corresponding polynomial, using the following algorithm:
  - A linear feedback shift register is implemented according to the coefficients of the generating polynomial, similar as the one depicted in Figure 4, but of degree m with $\ell S = 2^m$.
  - Let $t \in \{1, 2, \ldots, \ell S - 1\}$ denote the discrete time, where $\ell S - 1$ is the period of the PN sequence. At the initialization of the shift register, set $t = 1$. Let $q_b(t)$ be the vector representing the content of the shift register at the time t and let $q(t)$ be the decimal representation of the binary vector $q_b(t)$.
  - Every binary PN sequence of period $\ell S - 1$ will have a longest run of consecutive zeros at a unique time index x (this is a property of the PN sequences, see [2]). Then set:

$$\Pi(t) = \begin{cases} q(t) & \text{for } 1 \le t \le x - 1, \\ \ell S & \text{for } t = x, \\ q(t-1) & \text{for } x + 1 \le t \le \ell S. \end{cases} \tag{13}$$

An advantage of the set of pseudo random interleavers is the fact that every interleaver of length $\ell S = 2^m$ can be generated using only m bits that represent the coefficients of the primitive polynomials. The memory necessary to store the "seed" of these interleavers in the mobile stations is then $K_{max} \cdot m$. For example, in a system with $K_{max} = 120$ and $\ell S = 2^{14}$, only $120 \cdot 14 = 1680$ bits need to be stored in every mobile station.

C. Nested Interleavers

Another set of non-orthogonal interleavers can be constructed by using composition maps of a single interleaver. Let the symbol $\circ$ denote the composition of maps. That is, $\pi_j \circ \pi_i(x) = \pi_j(\pi_i(x))$. The correlation of these interleavers can be measured using the peak basis correlation $\mathrm{P}(\cdot,\cdot)$.

Algorithm:

- Choose a primitive polynomial and build one pseudo random interleaver π_1 using the same procedure as described in Section IV-B.
- Permute the images of the first interleaver π_1 by itself. This yields the second interleaver: $\pi_2 = \pi_1 \circ \pi_1$.
- Permute the images of the second interleaver π_2 by π_1 and get the third interleaver: $\pi_3 = \pi_1 \circ \pi_2$.
- Repeat the same procedure to obtain the interleavers π_i for $i \in \{4, 5, \ldots, K_{max}\}$.

The advantage of the nested over the pseudo random interleavers (Section IV-B) is that the memory required in the mobile station to generate the interleavers is K_{max} times lower (since only one master polynomial need be stored). A disadvantage is that their generation is slower, since after the generation of the initial interleaver, $O\left(\lceil \log_2 K \rceil\right)$ compositions of maps need to be performed in order to compute the interleaver for user K.

V. Computer Simulations Results

A. Performance of Uncoded IDMA

For all the simulations in this paper, the IDMA decoding algorithm described in [1] was used. The simulated curves in Figures 5 and 6 represent the average bit error rate of all users as a function of Eb/No[dB]. We have used the parameters $S = 64$ and $\ell = 256$. For every curve, the transmission of more than 1000 blocks per user was simulated. For 1, 32 and 63 users, the number of iterations performed in the decoding algorithm is 10. For 96, 110 and 120 users, the number of iterations is 30. Since the measured curves for the different families of interleavers are very similar to each other, in Figure 6 we only have depicted the results for pseudo random interleavers. For comparison, every figure also contains the results of simulations with random interleavers and the single user bound. The used decoder is sub-optimal in the sense that the channel we use is not noiseless. This explains the fact that the non-orthogonal interleaver families perform as good as the orthogonal interleavers (for up to 63 users).

B. Correlation of Interleavers

In this section results of the evaluation of the peak basis correlations $\mathrm{P}(\pi_i, \pi_j)$, as described in Section III-D, are given. Since the sets of interleavers contain on the order of 100 interleavers, for each family of interleavers, we show the correlation values of only the first 5 interleavers. The numbers in the first column and first row indicate the sequential number of the interleaver.

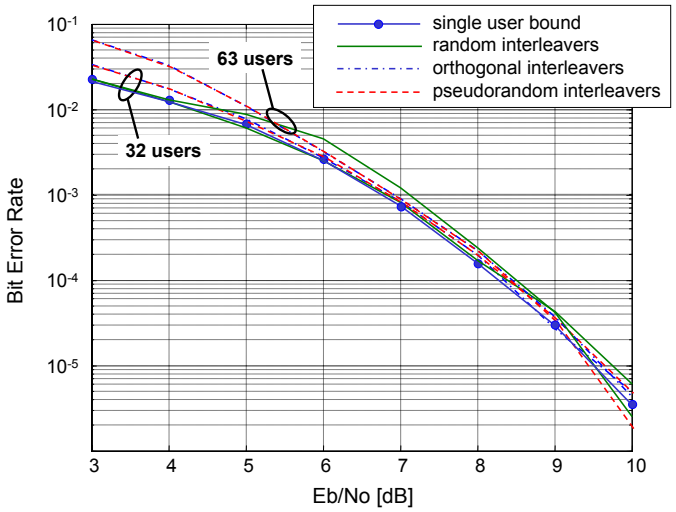

Fig. 5. Simulation results for 32 and 63 users

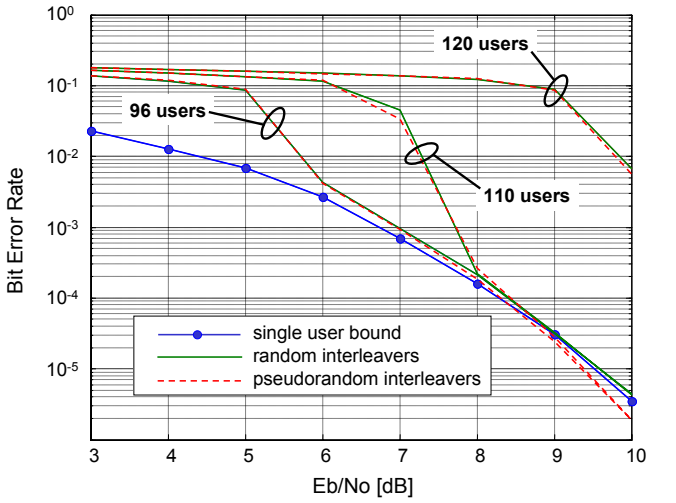

Fig. 6. Simulation results for 96, 110 and 120 users

The *peak basis correlation* values $\mathrm{P}(\pi_i, \pi_j)$ for *random interleavers* are:

	1	2	3	4	5
1	16384	1852	1756	1732	1692
2	1856	16384	1724	1804	1784
3	1780	1828	16384	1768	1732
4	1884	1772	1888	16384	1772
5	1716	1716	1796	1828	16384

The *peak basis correlation* values $\mathrm{P}(\pi_i, \pi_j)$ for *orthogonal interleavers* are:

	1	2	3	4	5
1	16384	0	0	0	0
2	0	16384	0	0	0
3	0	0	16384	0	0
4	0	0	0	16384	0
5	0	0	0	0	16384

The *peak basis correlation* values $P(\pi_i, \pi_j)$ for *pseudo random interleavers* are:

	1	2	3	4	5
1	16384	1760	1836	1696	1764
2	2060	16384	1756	1716	1824
3	1764	1760	16384	1752	1768
4	1804	1788	1876	16384	1692
5	1708	1716	1724	1744	16384

The correlation values for nested interleavers are very similar to those for pseudo random interleavers and therefore we do not show them here.

VI. Conclusion

There are two main contributions of this work. First we have determined what we mean by "good interleavers". We have defined orthogonal interleavers and we have shown that they lead to perfect decorrelation of users in a noiseless IDMA system. In Theorem 1 we proved that the number of orthogonal interleavers cannot be larger than the spreading length of the system. Furthermore, we have suggested a method of bounding the correlation between arbitrary interleavers. This method is useful to measure how "close to orthogonality" a set of non-orthogonal interleavers is.

The second main contribution is the construction of the three different types of interleavers presented in Section IV. The orthogonal, pseudo random and nested interleavers meet the design criteria of simplicity and fast generation on the one hand and low cross-correlation on the other hand, as stated in Section I. The simulations in Section V show that the performances of these interleavers are very similar to the performance of random interleavers. Thus, a working IDMA system can be built with each of the three proposed interleaving techniques. In [1], it is shown that IDMA with random interleavers can support more users than a comparable CDMA system. Since the performances of the interleavers described in Section IV are very similar to the performance of random interleavers, we can affirm that IDMA using orthogonal, pseudo random or nested interleavers, can support more users than a conventional CDMA system.

References

[1] L. Ping, L. Liu, W. K. Leung, *A Simple Approach to Near-Optimal Multiuser Detection: Interleave-Division Multiple-Access*, WCNC 2003 - IEEE Wireless Communications and Networking Conference, vol. 4, no.1, Mar 2003, pp. 391-396.

[2] J. S. Lee, L. E. Miller, *CDMA Systems Engineering Handbook*, Artech House, 1998.

[3] B. Vucetic, J. Yuan, *Turbo Codes: Principles and Applications*, Kluwer Academic Publishers, Third Printing, 2002.

[4] C. Berrou, A. Glavieux, *Near Optimum Error Correcting Codind and Decoding: Turbo-Codes*, IEEE Transactions on Communications, vol 44, no. 10, October 1996, pp. 1261-1271.

[5] W. Wesley Peterson, E. J. Weldon Jr., *Error-Correcting Codes*, The MIT Press, Second Edition, 1972.

[6] L. Ping, L. Liu, K. Wu, W. K. Leung, *Approaching the Capacity of Multiple Access Channels Using Interleaved Low-Rate Codes*, IEEE Communication Letters, vol. 8, no. 1, Jan 2004, pp. 4-6.

Appendix A. Proof of Theorem 1

Let $\{\pi_i : i \in \{1, 2, \ldots, K_{max}\}\}$ be a set of orthogonal interleavers. Define the following ℓ-dimensional vectors

$$\begin{aligned} w_1 &= (1, 1, \ldots, 1)^T, \\ w_2 &= (1, -1, 1, 1, \ldots, 1)^T, \\ w_3 &= (1, 1, -1, 1, \ldots, 1)^T, \\ &\vdots \\ w_\ell &= (1, 1, 1, 1, \ldots, -1)^T. \end{aligned}$$

One can easily prove that the vectors w_i are linearly independent. Hence, every vector in $\mathbb{R}^\ell$ can be written as a linear combination of these vectors. Since $W \subset \mathbb{R}^\ell$, every word $w \in W$ can be written as a linear combination of the vectors $w_1, w_2, \ldots, w_\ell$. Now let $x_1 \in W$ and $x_2 \in W$ be two words. We can write $x_1 = \sum_{m=1}^{\ell} \alpha_m w_m$ and $x_2 = \sum_{n=1}^{\ell} \beta_n w_n$ with $\alpha_i \in \mathbb{R}$ and $\beta_i \in \mathbb{R}$ for all $i \in \{1, 2, \ldots, \ell\}$. The orthogonality condition (2) for two different interleavers π_i and π_j with $i \leq K_{max}$ and $j \leq K_{max}$ for the words x_1 and x_2 becomes

$$\left\langle \pi_i\left(f\left(\sum_{m=1}^{\ell} \alpha_m w_m\right)\right) \, , \, \pi_j\left(f\left(\sum_{n=1}^{\ell} \beta_n w_n\right)\right) \right\rangle = 0. \tag{14}$$

Because of linearity, this equation is equivalent to

$$\sum_{m,n=1}^{\ell} \alpha_m \beta_n \left\langle \pi_i(f(w_m)), \pi_j(f(w_n)) \right\rangle = 0. \tag{15}$$

A sufficient condition for equality (15) to hold, is that the following equation holds

$$\left\langle (\pi_i(f(w_m))), \pi_j(f(w_n)) \right\rangle = 0, \tag{16}$$

for all $i \in \{1, 2, \ldots, K_{max}\}$ and $j \in \{1, 2, \ldots, K_{max}\}$ with $i \neq j$, and for all $m \in \{1, 2, \ldots, \ell\}$ and $n \in \{1, 2, \ldots, \ell\}$. The condition in (16) is also necessary for (2) to hold, since all w_m and w_n are words in W.

The vectors $\pi_j(f(w_1)), \pi_j(f(w_2)), \ldots, \pi_j(f(w_\ell))$ are linearly independent because $w_1, w_2, \ldots, w_\ell$ are linearly independent and f and π_j are linear functions. This means that for every j, the vectors $\pi_j(f(w_1)), \pi_j(f(w_2)), \ldots, \pi_j(f(w_\ell))$ generate via linear combinations a vector subspace of dimension ℓ in $\mathbb{R}^{\ell S}$. It can be proven by contradiction that all K_{max} such subspaces are orthogonal to each other and intersect each other only at one point, the origin. This implies that all vectors of the form $\pi_i(w_m)$ with $i \in \{1, 2, \ldots, K_{max}\}$ and $m \in \{1, 2, \ldots, \ell\}$ are linearly independent. There are $\ell \cdot K_{max}$ such vectors, hence

$$\ell \cdot K_{max} \leq \dim(\mathbb{R}^{\ell S}) = \ell \cdot S, \tag{17}$$

and $K_{max} \leq S$.