

A Simple, Unified Approach to Nearly Optimal Multiuser Detection and Space-Time Coding¹

Li Ping, K. Y. Wu, Lihai Liu, and W. K. Leung

Department of Electronic Engineering, City University of Hong Kong, Hong Kong, email: eeliping@cityu.edu.hk

Abstract – Techniques using interleaving as the basic means for signal separation are introduced for both multiple access systems and multiple transmit antenna systems. A very low-cost chip-by-chip iterative detection algorithm is presented. The proposed schemes can achieve nearly optimal performance for systems with a large numbers of users or transmit antennas.

I. INTRODUCTION

The performance of multiple access systems [1]-[4] and multiple transmit antenna systems [5]-[6] is limited by interference, i.e., multiple access interference (MAI) for the former and cross antenna interference (CAI) for the latter.

This paper reports a simple and unified treatment for both MAI and CAI problems. We adopt interleaving as the basic means for signal separation as opposed to other means based on time, frequency or spreading sequences. A very low-cost chip-by-chip iterative detection algorithm is presented. The proposed schemes can achieve nearly optimal performance for systems with large numbers of users or transmit antennas. More insights into the proposed schemes can be found in [7][8].

II. IDMA AND MULTIUSER DETECTION

This section is concerned with an interleave-division-multiple-access (IDMA) scheme [7]. IDMA inherits many advantages from CDMA, in particular, diversity against fading and mitigation of other-cell-user interference.

A. The Transmitter Structure

Fig. 1(a) shows the transmitter structure of an IDMA system. The information stream $\mathbf{d}^{(m)}$ for user m ($m=1, \dots, M$) is encoded using a code C into a coded sequence $\mathbf{c}^{(m)} = \{c_j^{(m)}, j=1, \dots, J\}$, followed by a chip level interleaver $\pi^{(m)}$ that maps $\mathbf{c}^{(m)}$ to $\mathbf{x}^{(m)} = \{x_j^{(m)}, j=1, \dots, J\}$. (We follow the convention of CDMA and call the basic elements in $\mathbf{c}^{(m)}$ and $\mathbf{x}^{(m)}$ “chips”.) C can be either the same or different for different users. It can be an FEC code, or a spreading sequence (spreading is in fact a special form of coding), or a combination of the two. From a performance point of view, it is advantageous to use a low-rate FEC code [9] that can provide an extra coding gain.

The key principle of IDMA is that the interleaver $\pi^{(m)}$ should be different for different users. In the following, we assume that these interleavers are generated independently and randomly.

In a multipath channel, after sampling at chip rate, the received signal from M users can be written as

$$r_i = \sum_{l=0}^{L-1} \sum_{m=1}^M x_{i-l}^{(m)} h_l^{(m)} + n_i, \quad i=1, 2, \dots \quad (1)$$

where L is the delay spread in terms of chips, $x_j^{(m)}$ the j th chip transmitted by the m th user, $\{h_l^{(m)}, l=0, \dots, L-1\}$ the channel reflection coefficients for user m and $\{n_i\}$ samples of a complex zero-mean additive white Gaussian noise (AWGN) with variance $\sigma^2 = N_0/2$ per real dimension.

For simplicity, we will focus on BPSK modulation ($x_j^{(m)} \in \{+1, -1\}$) and real $\{h_l^{(m)}\}$. The principle can be generalized to higher order modulations and complex channels.

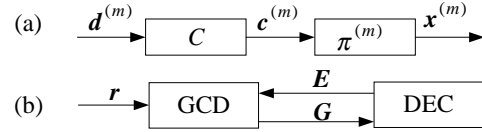


Fig. 1. Structure of the IDMA system. (a) The transmitter structure for user m , where $\pi^{(m)}$ is an interleaver. (b) The receiver structure.

B. The Receiver Structure

Fig. 1(b) illustrates an iterative IDMA multiuser detector. It consists of a Gaussian chip detector (GCD) and an *a posteriori* probability (APP) decoder (DEC). The GCD exchanges information with the DEC in a turbo-type manner [10]. The DEC makes hard decisions on the information bits in the last iteration.

The receiver is to estimate $\{x_j^{(m)}\}$ based on the received signal $\mathbf{r} \equiv \{r_i\}$, the channel parameters $\mathbf{h} \equiv \{h_l^{(m)}\}$ and the code constraint C . Finding an optimal solution is usually prohibitively complicated. We now consider a sub-optimal approach by first separating the conditions, i.e., \mathbf{r} , \mathbf{h} and C , and then combining the results using an iterative process. This greatly reduces the complexity involved.

Specifically, the constraint of C is ignored in the GCD. The output of the GCD $\mathbf{G} \equiv \{g_j^{(m)}, j=1, \dots, J, m=1, \dots, M\}$ is defined by the log-likelihood ratio (LLR) below,

$$g_j^{(m)} \equiv \log \left(\frac{p(\mathbf{r} | x_j^{(m)} = +1, \mathbf{h})}{p(\mathbf{r} | x_j^{(m)} = -1, \mathbf{h})} \right), \quad \forall m, j. \quad (2)$$

The DEC consists of M local APP decoders. The m th local APP decoder performs an APP decoding of C for user m using $\mathbf{g}^{(m)} = \{g_j^{(m)}, j=1, \dots, J\}$ (after appropriate interleaving) as its input. Its output is the so-called extrinsic LLR

¹ This work was fully supported by a City University of Hong Kong Strategic Grant (7001383).

$$e_j^{(m)} = \log \left(\frac{\Pr(x_j^{(m)} = +1 | \mathbf{g}^{(m)}, C)}{\Pr(x_j^{(m)} = -1 | \mathbf{g}^{(m)}, C)} \right) - g_j^{(m)} \quad \forall m, j. \quad (3)$$

We denote $\mathbf{E} \equiv \{e_j^{(m)}, j = 1, \dots, J, m = 1, \dots, M\}$, which forms the feedback information in Fig. 1(b) for the next iteration.

C. Detailed Description of the GCD

The GCD generates coarse estimations of $\{x_j^{(m)}, \forall j, m\}$. We ignore the constraint of C at this stage to maintain low complexity. Consider the j th chip of the m th user with $x_j^{(m)} \in \{+1, -1\}$ under BPSK modulation. We treat $x_j^{(m)}$ as a random variable and use $e_j^{(m)}$ (initialized to zero) to approximate the *a priori* LLR for $x_j^{(m)}$, i.e.,

$$\log \left(\frac{\Pr(x_j^{(m)} = +1)}{\Pr(x_j^{(m)} = -1)} \right) \approx e_j^{(m)}. \quad (4)$$

Based on (4), we have

$$\mu_j^{(m)} \equiv E(x_j^{(m)}) \approx \frac{\exp(e_j^{(m)}) - 1}{\exp(e_j^{(m)}) + 1} = \tanh(e_j^{(m)} / 2), \quad (5)$$

$$v_j^{(m)} \equiv \text{Var}(x_j^{(m)}) = 1 - (\mu_j^{(m)})^2. \quad (6)$$

According to (1) and denoting $\zeta_{i,j}^{(m)} = r_i - x_j^{(m)} h_{i-j}^{(m)}$, the i th chip of the received signal can be expressed as

$$r_i = x_j^{(m)} h_{i-j}^{(m)} + \zeta_{i,j}^{(m)}. \quad (7)$$

The Gaussian Approximation: Applying the central limit theorem, $\zeta_{i,j}^{(m)}$ in (7) can be approximated by a Gaussian random variable with mean and variance

$$E(\zeta_{i,j}^{(m)}) = E(r_i) - \mu_j^{(m)} h_{i-j}^{(m)}, \quad (8a)$$

$$\text{Var}(\zeta_{i,j}^{(m)}) = \text{Var}(r_i) - v_j^{(m)} |h_{i-j}^{(m)}|^2, \quad (8b)$$

where (based on (1))

$$E(r_i) = \sum_{j=i-L+1}^i \sum_{m=1}^M \mu_j^{(m)} h_{i-j}^{(m)}, \quad (9a)$$

$$\text{Var}(r_i) = \sigma^2 + \sum_{j=i-L+1}^i \sum_{m=1}^M v_j^{(m)} |h_{i-j}^{(m)}|^2. \quad (9b)$$

Applying the Gaussian approximation to (7), the LLR for $x_j^{(m)}$ based on r_i can be calculated as

$$\begin{aligned} g_{i,j}^{(m)} &\equiv \log \left(\frac{p(r_i | x_j^{(m)} = +1, \mathbf{h})}{p(r_i | x_j^{(m)} = -1, \mathbf{h})} \right) = \log \left(\frac{\phi(r_i - E(\zeta_{i,j}^{(m)}) - h_{i-j}^{(m)}, \text{Var}(\zeta_{i,j}^{(m)}))}{\phi(r_i - E(\zeta_{i,j}^{(m)}) + h_{i-j}^{(m)}, \text{Var}(\zeta_{i,j}^{(m)}))} \right) \\ &= 2h_{i-j}^{(m)} \cdot \frac{r_i - E(\zeta_{i,j}^{(m)})}{\text{Var}(\zeta_{i,j}^{(m)})} \end{aligned} \quad (10)$$

where $\phi(t, \sigma^2) = (2\pi\sigma^2)^{-1/2} \exp(-t^2/2\sigma^2)$ is the pdf of a Gaussian random variable. Considering L samples in \mathbf{r} related to $x_j^{(m)}$, $g_j^{(m)}$ in (2) can be approximated by

$$g_j^{(m)} \approx \sum_{i=j}^{j+L-1} g_{i,j}^{(m)} \quad (11)$$

This is an approximate soft-in/soft-out RAKE-type operation ignoring the correlation among interference samples.

D. The DEC Principle

The APP decoding in the DEC is a standard function so we will not discuss it in detail. A common assumption in all realistic APP decoders is that the distortion components in the input data should be i.i.d. For a conventional CDMA receiver, interference takes a frame-to-frame form, i.e., the chips within a frame of one user will interfere with the chips within a frame of another user. This results in dependency among the interference components in different chips, which will affect the performance of the DEC. Sophisticated matrix techniques can be applied to treat this problem [3], but the cost is quite high.

For an IDMA receiver, interference among different users is diversified due to different interleavers employed by different users. Consequently, the simple formulas in (10) and (11) can be applied. The interference components in the resultant $\{g_j^{(m)}\}$ are approximately independent so that $\{g_j^{(m)}\}$ can be used directly by the DEC.

E. A Summary of the IDMA Detection Algorithm

Initialization: Set $e_j^{(m)} = 0 \quad \forall m, j. \quad (12)$

Main iteration:
 $\mu_j^{(m)} \leftarrow \tanh(e_j^{(m)} / 2) \quad \forall m, j. \quad (13a)$

$v_j^{(m)} \leftarrow 1 - (\mu_j^{(m)})^2 \quad \forall m, j. \quad (13b)$

$E(r_i) \leftarrow \sum_{j=i-L+1}^i \sum_{m=1}^M \mu_j^{(m)} h_{i-j}^{(m)} \quad \forall i. \quad (14a)$

$\text{Var}(r_i) \leftarrow \sigma^2 + \sum_{j=i-L+1}^i \sum_{m=1}^M v_j^{(m)} |h_{i-j}^{(m)}|^2 \quad \forall i. \quad (14b)$

$E(\zeta_{i,j}^{(m)}) \leftarrow E(r_i) - \mu_j^{(m)} h_{i-j}^{(m)} \quad i=j, \dots, j+L-1, \forall m, j. \quad (15a)$

$\text{Var}(\zeta_{i,j}^{(m)}) \leftarrow \text{Var}(r_i) - v_j^{(m)} |h_{i-j}^{(m)}|^2 \quad i=j, \dots, j+L-1, \forall m, j. \quad (15b)$

$g_{i,j}^{(m)} \leftarrow 2h_{i-j}^{(m)} \cdot \frac{r_i - E(\zeta_{i,j}^{(m)})}{\text{Var}(\zeta_{i,j}^{(m)})} \quad i=j, \dots, j+L-1, \forall m, j. \quad (16a)$

$g_j^{(m)} \leftarrow \sum_{i=j}^{j+L-1} g_{i,j}^{(m)} \quad \forall m, j. \quad (16b)$

The APP decoding in the DEC is performed at this point. Then go back to (13) for the next iteration.

The normalized computational cost in (13) ~ (16) (excluding the APP decoding of C) is only about $6L$ additions, $4L+2$ multiplications and a tanh function per chip per user per iteration. This complexity is very low and is independent of user number M . If C is a turbo-type code [9], the DEC may contain internal iterations, which can be incorporated into the global iteration described above (e.g. one internal iteration per DEC per global iteration). Since other cost is so low, the overall receiver complexity is dominated by the APP decoding of C . In this case, the normalized cost per user is

roughly the same as that of a single-user decoder. Using the low-cost scheme in [9], the receiver complexity (per user) of an IDMA scheme can be even lower than that of a standard 16-state single-user turbo decoder.

F. Simulation Results

The above detection algorithm can be carried out by either a parallel schedule or a serial schedule [7]. The results (Fig. 2 – Fig. 4) in this subsection are simulated with the serial one. We will use N_{inf} for the information block length of each user, L for the number of multipaths, M for user number, S for the length of a spreading sequence, and R_c for the rate of C . We define $R_{\text{system}}=M \times R_c$, which is the overall measurement of spectrum efficiency. We observed that the structure of the spreading sequence has little impact on the IDMA performance, except that the numbers of +1 and -1 should be balanced, e.g., [+1, -1, +1, -1, ...].

Fig. 2 shows the BER performance of an IDMA scheme with $S=64$ over AWGN channels (so $R_c=1/64$). $N_{\text{inf}}=256$ for each user and the interleaver length is 64×256 . When $M \leq S$, 5 iterations are sufficient. (For $M > S$, more iterations should be used.) Near single-user performance is achievable for very large M . The detector converges even when M is nearly double S . For QPSK modulation, similar performance can be achieved for twice the number of users, by splitting the signals evenly in the real and imaginary parts.

Fig. 3 illustrates the BER performance of an IDMA scheme over quasi-static Rayleigh fading multipath channels. All users employ the same rate-1/2 convolutional code (generators = (1, 35/23)₈). $M=16$, $N_{\text{inf}}=512$ and $S=16$ (so $R_c=1/(2 \times 16)$). Fading coefficients remain constant within a frame of 16384 (= 512 \times 2 \times 16) chips. From Fig. 3, performance improves significantly with increasing path number due to the improved diversity.

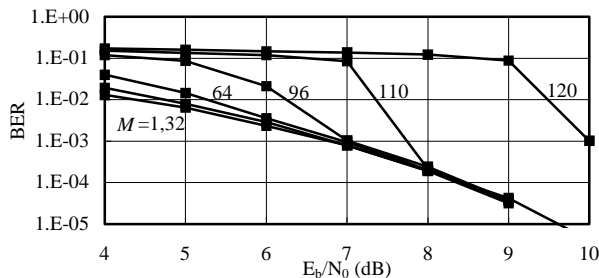


Fig. 2. Un-coded IDMA performance over AWGN channels. $S=64$.

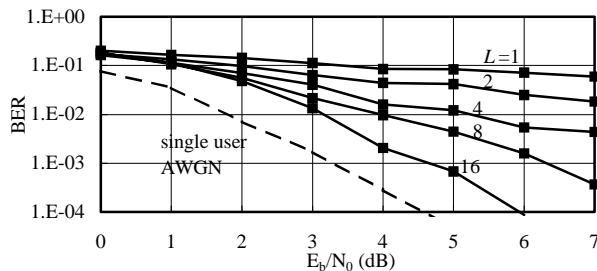


Fig. 3. Convolutionally coded IDMA performance over quasi-static Rayleigh fading channels. $N_{\text{inf}}=512$. Iteration number=5.

Fig. 4 contains the simulated results for an IDMA system over AWGN channels based on a low-rate (with rate 0.0581) turbo-Hadamard code [9] with $N_{\text{inf}}=4095$. There is no additional spreading sequence used, so $S=1$ (and $R_c=0.0581$). Performance of $\text{BER}=10^{-5}$ at $E_b/N_0 \approx 1.4$ dB is observed with $M=8$ ($R_{\text{system}}=0.465$). Recall that a turbo code with the same rate and information length requires a similar E_b/N_0 value in AWGN. Thus the performance shown is nearly optimal.

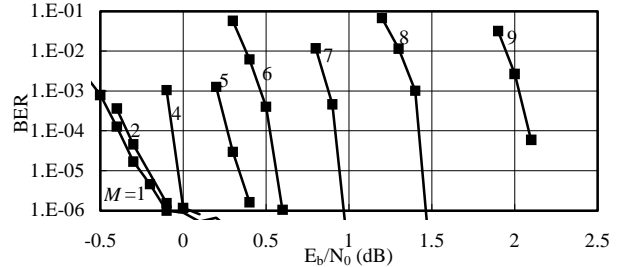


Fig. 4. Performance of an IDMA scheme based on a low-rate turbo-Hadamard code [9] over AWGN channels.

III. IDM-ST SPACE-TIME CODES

We now proceed to consider space-time (ST) code design for multiple transmit antenna systems [8]. The basic strategy is to reduce the correlation among the signals over different transmit antennas. It can be shown that this strategy is consistent with the rank and determinant criteria established in [5]. We take a random approach following the IDMA principle developed above.

Fig. 5 shows an interleave-division-multiplexing space-time (IDM-ST) coding scheme with N transmit antennas. The information sequence \mathbf{d} is first encoded by C into a chip sequence \mathbf{c} . To reduce correlation and increase rate, we adopt interleaving and stacking operations as follows. Assume that \mathbf{c} is segmented into F equal-length subsequences $\{\mathbf{c}_1, \dots, \mathbf{c}_f, \dots, \mathbf{c}_F\}$. (F is referred to as the stacking index.) Each \mathbf{c}_f is then independently interleaved N times, producing $\{\mathbf{x}^{(n,f)}, n=1, \dots, N\}$. Random interleavers are used here. Altogether, we need $N \times F$ independent random interleavers. The stacking operation is defined by an ordinary linear summation,

$$\mathbf{X}^{(n)} = \sum_{f=1}^F \mathbf{x}^{(n,f)}, \quad n=1, 2, \dots, N. \quad (17)$$

Here $\mathbf{X}^{(n)}$ is the signal transmitted over antenna n . We refer to the interleavers and stackers together as an IDM-ST code C_{ST} . The stacking operation is to increase rate and it does not cause any additional problem, since the related interference is resolved together with CAI and inter symbol interference.

The transmission schemes in Figs. 1(a) and 5 are closely related. Consider Fig. 1(a) with $M=N \times F$ users: each user transmits a signal sequence $\mathbf{x}^{(n,f)}, f=1, \dots, F, n=1, \dots, N$ but $\{\mathbf{x}^{(n,f)}, n=1, \dots, N\}$ result from the same input information. We assume that N transmit antennas are used and the F users transmitting $\{\mathbf{x}^{(n,f)}, f=1, \dots, F\}$ share the n th antenna. Then the systems in Figs.1(a) and 5 are equivalent.

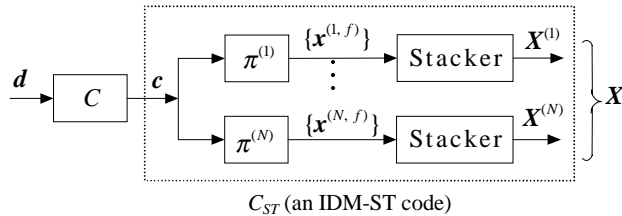


Fig. 5. The IDM-ST transmitter structure with N transmit antennas. Each $\pi^{(n)}$ consists of F independent random interleavers, each responsible for a segment $\mathbf{x}^{(n,f)}$.

The IDM-ST receiver is similar to the IDMA one with some modifications. Recall that $\{\mathbf{x}^{(n,f)}, n=1, \dots, N\}$ are different interleaved versions of \mathbf{c}_f . We can thus sum up the de-interleaved versions of the estimations of $\{\mathbf{x}^{(n,f)}, n=1, \dots, N\}$ from the GCD to generate the LLR vector for \mathbf{c}_f , which forms the input to the DEC. The DEC function should be modified accordingly, as it is necessary to estimate the mean and variance of every chip in every $\mathbf{x}^{(n,f)}$ (details omitted here). The computational cost of the above receiver is also very low (increasing linearly with N).

The capacity of a transmit diversity system with N transmit antennas and one receive antenna (referred to as an $N \times 1$ system) can be calculated as (for a fixed \mathbf{h}) [11][12],

$$C_{N \times 1} = \log(1 + (\|\mathbf{h}\|^2 / N) \cdot (E_s / N_0)) \quad (18)$$

where E_s is the average symbol energy on all transmit antennas. Refer to Fig. 5. Based on (18), we will say that C_{ST} is canonical if the performance of the overall system is that of C in an AWGN channel scaled by an energy factor of $\|\mathbf{h}\|^2 / N$ for any fixed \mathbf{h} . The canonical performance in a quasi-static fading channel can be estimated based on that of C in AWGN scaled by $\|\mathbf{h}\|^2 / N$ and averaged over \mathbf{h} .

Based on (18), the capacity of an $N \times 1$ system is achievable by the system in Fig. 5 using a C_{ST} with canonical performance and an FEC code C optimized for AWGN channels. Therefore the canonical performance serves as a benchmark for the performance of a full diversity system.

Let N_{inf} , N , S , F and R_c be defined as above. We define $R_{\text{system}} = F \times R_c$ as the overall spectrum efficiency. Complex signaling and one receive antenna are always assumed.

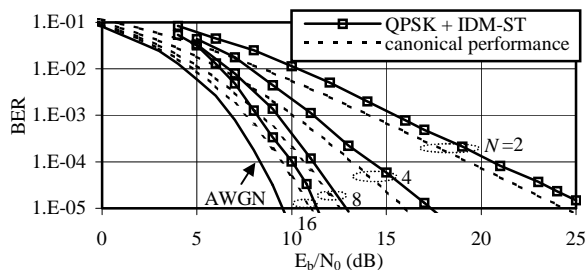


Fig. 6. Performance of un-coded IDM-ST schemes over quasi-static Rayleigh fading channels. $S=(8,4,2,1)$ for $N=(2,4,8,16)$, respectively. $R_c=1/S$. No ISI. $F=2S$ so $R_{\text{system}}=2$ bits/symbol. $N_{\text{inf}}=512$. Iteration number=10.

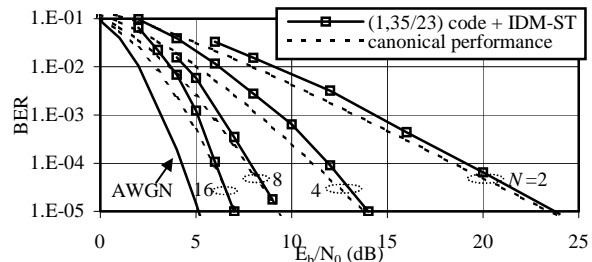


Fig. 7. Performance of convolutionally coded IDM-ST schemes with QPSK modulation over quasi-static Rayleigh fading channels. The convolutional code has rate=1/2 and generators=(1, 35/23)₈. $S=2$ and $R_c=1/2S$. No ISI. $F=2S$ so $R_{\text{system}}=1$ bit/symbol. $N_{\text{inf}}=1000$. Iteration number=5.

Figs. 6 and 7 illustrate the performance of IDM-ST schemes with and without coding over quasi-static Rayleigh fading channels. We can see that both schemes (especially the coded one) can nearly achieve the canonical performance. By increasing N from 2 to 16, substantial gain (about 15dB) is achieved at $\text{BER}=10^{-5}$. For $N>2$, the performance in Figs. 6 and 7 is also considerably better than that of existing schemes [5][6] (assuming comparable rates).

IV. CONCLUSIONS

Simple, nearly optimal interleave-based schemes have been introduced for both multiuser detection and space-time coding.

REFERENCES

- [1] M. Moher, "An iterative multiuser decoder for near-capacity communications," *IEEE Trans. Commun.*, pp. 870-880, July 1998.
- [2] M. C. Reed, C. B. Schlegel, P. D. Alexander, and J. A. Asenstorfer, "Iterative multiuser detection for CDMA with FEC: Near-single-user performance," *IEEE Trans. Commun.*, pp. 1693-1699, Dec. 1998.
- [3] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, pp. 1046-1061, July 1999.
- [4] F. N. Brannstrom, T. M. Aulin, and L. K. Rasmussen, "Iterative multiuser detection of trellis code multiple access using a posteriori probabilities," in *Proc. ICC 2001*, Finland, June 2001, pp. 11-15.
- [5] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance analysis and code construction," *IEEE Trans. Inform. Theory*, pp. 744-765, Mar. 1998.
- [6] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, pp. 1456-1467, July 1999.
- [7] Li Ping, L. H. Liu, and W. K. Leung, "Interleave-Division Multiple-Access: A Simple Approach to Near Optimal Multiuser Detection," submitted to *ICC 2003* (also available: www.ee.cityu.edu.hk/~liping).
- [8] Li Ping, K. Y. Wu, and W.K. Leung, "A Simple Approach to Near-Optimal Multiple Transmit Antenna Space-Time Codes," submitted to *ICC 2003* (also available: www.ee.cityu.edu.hk/~liping).
- [9] Li Ping, W. K. Leung, and K. Y. Wu, "Low-rate turbo-Hadamard codes," submitted for publication, available: www.ee.cityu.edu.hk/~liping.
- [10] 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.
- [11] G. J. Foschini and M. Gans, "On the limits of wireless communication in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311-335, Mar. 1998.
- [12] A. Narula, M. D. Trott, and G.W. Wornell, "Performance limits of coded diversity methods for transmitter antenna arrays," *IEEE Trans., Inform. Theory*, pp. 2418-2433, Nov. 1999.