

# Iterative Multiuser Detection and Decoding with Spatially Coupled Interleaving

Keigo Takeuchi, *Member, IEEE* and Shuhei Horio

**Abstract**—Spatially coupled (SC) interleaving is proposed to improve the performance of iterative multiuser detection and decoding (MUDD) for quasi-static fading multiple-input multiple-output systems. The linear minimum mean-squared error (LMMSE) demodulator is used to reduce the complexity and to avoid error propagation. Furthermore, sliding window MUDD is proposed to circumvent an increase of the decoding latency due to SC interleaving. Theoretical and numerical analyses show that SC interleaving can improve the performance of the iterative LMMSE MUDD for regular low-density parity-check codes.

**Index Terms**—spatial coupling, multiple-input multiple-output (MIMO) systems, iterative multiuser detection and decoding, sliding window decoding, density evolution.

## I. INTRODUCTION

**S**PATIAL coupling is a sophisticated technique for boosting the belief-propagation (BP) decoding threshold up to the optimal one [1]. The basic idea of spatially coupled (SC) low-density parity-check (LDPC) codes in [1] is as follows: An SC LDPC code is constructed as a chain of  $L$  conventional LDPC codes. The point is to introduce an irregular structure at both ends that allows the decoder to attain reliable information at the ends. When the code length  $M$  of each section in the chain is sufficiently long, the reliable information can propagate toward the center of the chain regardless of the chain length  $L$ . The influence of the irregularity at both ends is negligible when  $1 \ll L \ll M$ . Consequently, the BP-based iterative decoding can achieve the optimal performance.

In this letter, spatial coupling is utilized to improve the performance of BP-based iterative multiuser detection and decoding (MUDD) for quasi-static fading multiple-input multiple-output (MIMO) systems. One should not confuse the terminology “spatial coupling” with coupling in the physical space. Coupling is actually made in the time domain. Iterative MUDD algorithms based on approximate BP have been proposed in [2], [3] and analyzed via density evolution in the large-system limit where the numbers of transmit and receive antennas tend to infinity at the same rate [4]. These low-complexity algorithms can achieve excellent decoding performance in the large-system limit compared to non-iterative receivers. We propose SC interleaving to improve the performance of iterative MUDD via an improvement of the decoding threshold.

The main contribution of this letter is to incorporate spatial coupling into bit-interleaved coded modulation (BICM),

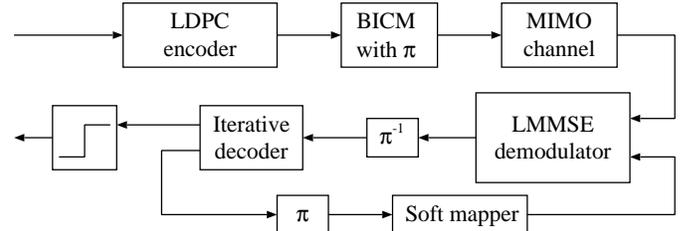


Fig. 1. MIMO system.  $\pi$  and  $\pi^{-1}$  represent the SC interleaver and its deinterleaver, respectively.

instead of encoding. It is possible to combine SC interleaving with any code that allows the decoder to use efficient BP decoding. This compatibility of SC interleaving is suitable for practical systems that use several codes as options.

As related works, it was proposed in [5]–[8] to combine spatial coupling with spread spectrum modulation. The main difference between this letter and the previous works is that practical codes with low decoding complexity are used in this letter, whereas the uncoded case [5]–[7] or the information-theoretically optimal codes [8] were considered in the previous works. In [8], the code rate was controlled to avoid the occurrence of error propagation. In this letter, we consider LDPC codes with a fixed rate, and investigate the influence of error propagation in iterative MUDD.

## II. SYSTEM MODEL

### A. MIMO System

We consider an MIMO system with  $K$  transmit antennas and  $N$  receive antennas operating over a frequency-flat quasi-static fading channel, shown in Fig. 1. A binary information stream is encoded with a  $(d_v, d_c)$ -regular LDPC code of code length  $M$  [9]. After SC interleaving, which will be presented shortly, the interleaved stream is modulated and divided into  $K$  data streams. Gray-mapped quadrature phase shift keying (QPSK)  $\mathcal{C} = \{a + jb : a, b = \pm 1/\sqrt{2}\}$  is used. The obtained data symbols  $\{\mathbf{x}_t = (x_{1,t}, \dots, x_{K,t})^T \in \mathcal{C}^K\}$  are directly transmitted from  $K$  transmit antennas at time  $t$ . The corresponding received vector  $\mathbf{y}_t \in \mathbb{C}^N$  is given by

$$\mathbf{y}_t = \mathbf{H}\mathbf{x}_t + \mathbf{n}_t, \quad \mathbf{n}_t \sim \mathcal{CN}(\mathbf{0}, N_0\mathbf{I}_N). \quad (1)$$

In (1),  $\{\mathbf{n}_t\}$  denote independent additive white Gaussian noise (AWGN) vectors with covariance  $N_0\mathbf{I}_N$ . The channel matrix  $\mathbf{H} = (\mathbf{h}_1, \dots, \mathbf{h}_K) \in \mathbb{C}^{N \times K}$  is assumed to be independent of the time index  $t$  and to be known to the receiver. The

Manuscript received June 25, 2013.

K. Takeuchi and S. Horio are with the Department of Communication Engineering and Informatics, the University of Electro-Communications, Tokyo 182-8585, Japan (e-mails: ktakeuchi@uec.ac.jp, h1231079@edu.cc.uec.ac.jp).

Digital Object Identifier \*\*\*

former assumption, i.e. the assumption of quasi-static fading is consistent with the latter assumption: It is possible for the receiver to estimate the channel matrix by utilizing the known training symbols sent from the transmitter. For simplicity, independent and identically distributed (i.i.d.) Rayleigh fading is assumed: The channel matrix  $\mathbf{H}$  has independent circularly symmetric complex Gaussian (CSCG) random elements with variance  $1/K$ . Note that there may be dependencies between the elements of  $\mathbf{H}$  in practice.

### B. Spatially Coupled Interleaver

We shall define an SC interleaver with section size  $M$ , chain length  $L$ , and coupling width  $W$ . The section size  $M$  is equal to the code length of the used code. The overall length of interleaving is  $ML$ . The chain is coupled circularly, and each section is connected to  $(W-1)$  neighboring sections uniformly and randomly. We impose a constraint under which each data symbol consists of bits in the same codeword, by defining a constrained interleaver of length  $M$  that maps consecutive integers  $\{2i, 2i+1\}$  to consecutive integers  $\{2j, 2j+1\}$  for  $i, j \in \{0, \dots, M/2-1\}$ . This constraint simplifies density evolution analysis. Input (respectively output) index  $m \in \mathcal{M} = \{0, \dots, M-1\}$  within section  $l \in \mathcal{L} = \{0, \dots, L-1\}$  corresponds to the  $(Ml+m)$ th input (respectively output) for the SC interleaver. See Fig. 2 for an example of the SC interleaver. In particular, the SC interleaver with  $W=1$  reduces to  $L$  uncoupled random interleavers.

**Definition 1** (SC Interleaver). *An SC interleaver  $\pi$  is a bijection from  $\mathcal{M} \times \mathcal{L}$  onto  $\mathcal{M} \times \mathcal{L}$ . Let  $\{\pi_l^{\text{in}} : l \in \mathcal{L}\}$  and  $\{\pi_l^{\text{out}} : l \in \mathcal{L}\}$  denote  $L$  independent random interleavers of length  $M$  and  $L$  independent random constrained interleavers of length  $M$ , respectively. For  $(m, l) \in \mathcal{M} \times \mathcal{L}$ , the SC interleaver  $\pi(m, l)$  is given by*

$$\pi(m, l) = (\pi_{l'}^{\text{out}}(\pi_l^{\text{in}}(m)), l'), \quad (2)$$

with  $l' = (l - (\lfloor \pi_l^{\text{in}}(m)/2 \rfloor)_W)_L$ , in which  $(i)_n \in \{0, \dots, n-1\}$  denotes the remainder for the division of  $i \in \mathbb{Z}$  by  $n \in \mathbb{N}$ .

A known sequence of length  $(W-1)M$  is sent in the first  $(W-1)$  sections. In general, the decoder can decode the codewords in sections close to the first sections with smaller error probability than in distant sections. When  $M$  is sufficiently large, the reliable information at the first sections may spread over the whole system. Consequently, it is possible to decode the codewords in distant sections with almost the same error probability as in sections close to the first sections.

## III. ITERATIVE MUDD

### A. Sliding Window MUDD

SC LDPC codes can be efficiently decoded by sliding window (SW) decoding [10]. We propose SW MUDD to reduce the decoding delay compared to iterative MUDD with parallel scheduling, in which messages are collectively sent to the decoder (respectively demodulator) after estimating the data symbols (respectively codewords) for *all* sections. In SW MUDD, the codeword at section  $l$  is decoded in the order<sup>1</sup>

<sup>1</sup> In order to avoid unnecessary latency, the codewords should be sent in the same order after transmission of a known sequence.

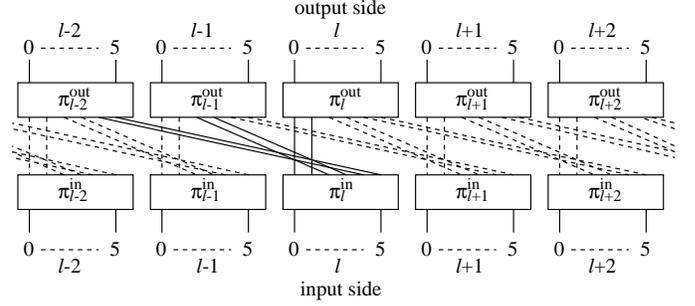


Fig. 2. SC interleaver for  $M=6$  and  $W=3$ .  $\pi_l^{\text{in}}$  and  $\pi_l^{\text{out}}$  denote a random interleaver and a random constrained interleaver for section  $l$ , respectively.

$l = W-1, L-1, W, L-2, \dots$ . In the decoding stage for section  $l$ , we update log likelihood ratios (LLRs) exchanged through the edges that are connected to the coded bits in section  $l$ , shown by the solid edges in Fig. 2, whereas the other LLRs are fixed to the current values. In each iteration, the demodulator calculates LLRs and sends them to the decoder, which uses the passed LLRs to calculate extrinsic LLRs to be fed back to the demodulator. After convergence or  $I$  iterations, the decoder outputs the decoding results, and the SW MUDD proceeds to the next stage. Note that the SW MUDD is suboptimal, since we do not update LLRs passed through the edges that are connected to the following sections.

The extrinsic LLRs passed from the decoder toward the demodulator are calculated by using the conventional sum-product algorithm with the number of iterations  $J$  [9]. The LLRs passed in the opposite direction are updated by using the linear minimum mean-squared error (LMMSE) demodulator, since the optimal demodulator is infeasible in terms of the complexity. In this letter, we refer to iterations in the decoder and in the MUDD as inner and outer iterations, respectively.

### B. LMMSE Demodulator

Let  $L_{k,t}^{\text{dec}} \in \mathbb{C}$  denote the a priori complex LLR for the data symbol  $x_{k,t} \in \mathcal{C}$  fed back from the decoder in an outer iteration round. The real and imaginary parts of  $L_{k,t}^{\text{dec}}$  correspond to the LLRs for those of the data symbol, respectively. In the initial outer iteration for the SW scheduling, the a priori LLR is equal to that passed from the decoder in the preceding stage if it exists. Otherwise, the LLR is set to zero. The corresponding a priori probability  $p(x_{k,t})$  is given by  $p(x_{k,t}) = p(\Re[x_{k,t}])p(\Im[x_{k,t}])$ , in which

$$p\left(\Re[x_{k,t}] = \pm \frac{1}{\sqrt{2}}\right) = \frac{e^{\pm \Re[L_{k,t}^{\text{dec}}]/2}}{e^{\Re[L_{k,t}^{\text{dec}}]/2} + e^{-\Re[L_{k,t}^{\text{dec}}]/2}}, \quad (3)$$

where a positive LLR implies that  $\Re[x_{k,t}] = 1/\sqrt{2}$  is more likely. The real part of the mean  $\hat{x}_{k,t}$  with respect to  $p(x_{k,t})$  is given by  $\Re[\hat{x}_{k,t}] = 2^{-1/2} \tanh(\Re[L_{k,t}^{\text{dec}}]/2)$ . The a priori probability and mean of  $\Im[x_{k,t}]$  are defined in the same manner. Thus, the a priori variance  $(1 - |\hat{x}_{k,t}|^2)$  of the QPSK symbol  $x_{k,t}$  is given by  $\sigma^2(\Re[L_{k,t}^{\text{dec}}], \Im[L_{k,t}^{\text{dec}}])$ , with

$$\sigma^2(L_r, L_i) = 1 - \frac{1}{2} \left\{ \tanh^2\left(\frac{L_r}{2}\right) + \tanh^2\left(\frac{L_i}{2}\right) \right\}. \quad (4)$$

We focus on the  $k$ th symbol at time  $t$ , and shall derive the LMMSE estimation of  $x_{k,t}$  based on the *extrinsic* information  $\{p(x_{k',t}) : k' \neq k\}$ . The use of the true a priori probability  $p(x_{k',t})$  results in the optimal nonlinear demodulator with high complexity. In order to reduce the complexity, the a priori probabilities  $p(x_{k',t})$  for all  $k' \neq k$  are approximated by proper complex Gaussian distributions with mean  $\hat{x}_{k',t}$  and variance  $(1 - |\hat{x}_{k',t}|^2)$  that are equal to those of  $x_{k',t}$  for  $p(x_{k',t})$ , respectively. On the other hand,  $p(x_{k,t})$  is approximated by a CSCG distribution with unit variance. These approximations result in the approximate posterior probability density function (pdf) of  $x_{k,t}$  given by

$$p(x_{k,t} | \mathbf{y}_t, \mathbf{H}) = \frac{1}{\pi \xi_{k,t}} \exp\left(-\frac{|x_{k,t} - \bar{x}_{k,t}|^2}{\xi_{k,t}}\right), \quad (5)$$

with

$$\bar{x}_{k,t} = \xi_{k,t} \mathbf{h}_k^H \boldsymbol{\Sigma}_{k,t}^{-1} \left( \mathbf{y}_t - \sum_{k' \neq k} \mathbf{h}_{k'} \hat{x}_{k',t} \right), \quad (6)$$

$$\xi_{k,t} = \left( 1 + \mathbf{h}_k^H \boldsymbol{\Sigma}_{k,t}^{-1} \mathbf{h}_k \right)^{-1}. \quad (7)$$

In these expressions,  $\boldsymbol{\Sigma}_{k,t}$  is given by

$$\boldsymbol{\Sigma}_{k,t} = N_0 \mathbf{I}_N + \sum_{k' \neq k} (1 - |\hat{x}_{k',t}|^2) \mathbf{h}_{k'} \mathbf{h}_{k'}^H. \quad (8)$$

Expression (5) implies that the complex LLR  $L_{k,t}^{\text{dem}} \in \mathbb{C}$  for  $x_{k,t}$  sent from the LMMSE demodulator to the decoder is given by

$$L_{k,t}^{\text{dem}} = 2\sqrt{2} \mathbf{h}_k^H \boldsymbol{\Sigma}_{k,t}^{-1} \left( \mathbf{y}_t - \sum_{k' \neq k} \mathbf{h}_{k'} \hat{x}_{k',t} \right), \quad (9)$$

of which real and imaginary parts correspond to the LLRs for those of  $x_{k,t}$ , respectively.

Unless  $N_0$  is zero, *soft* information about the data symbols, i.e. a finite LLR (9) is fed forward to the decoder. Note that the LMMSE demodulator reduces to the zero-forcing (ZF) demodulator when  $N_0$  in (8) is approximated by a sufficiently small value. Since the second term in (8) is not invertible for  $N > K - 1$ , the LLR (9) diverges when the ZF demodulator is used. In other words, *hard* information about the data symbols is sent to the decoder. The hard information may result in error propagation, so that the LMMSE demodulator is used to avoid error propagation in this letter.

#### IV. DENSITY EVOLUTION

##### A. Asymptotic Analysis

We follow [4] to present the density evolution of the iterative MUDD in the large-system limit after taking the infinite code length limit  $M \rightarrow \infty$ . In the large-system limit, the numbers of transmit and receive antennas tend to infinity with their ratio  $\alpha = K/N$  kept constant. It is known that the large-system analysis can provide a good prediction for the starting location of the so-called *waterfall* regime.

Let  $p_l^{\text{dec}}(L)$  denote the asymptotic pdf of the real LLRs emitted from the decoder for section  $l$  as  $M \rightarrow \infty$ . The pdf  $p_l^{\text{dec}}(L)$  can be analyzed with the Gaussian approximation of the LLRs [9]. Thus, we mainly present the large-system analysis of the demodulator.

##### B. LMMSE Demodulator

The analysis of the LMMSE demodulator is based on [11]. We focus on section  $l$ , and suppose that  $\{p_{l'}^{\text{dec}}(L) : l' = l, \dots, l + W - 1\}$  have been fed back from the decoder. Let  $(k, t)$  denote any couple that corresponds to indices included in section  $l$  at the output side of the SC interleaver. It is proved that the LLR (9) sent to the decoder is statistically equivalent to that for the interference-free complex AWGN channel

$$z_l = x_l + n_l, \quad n_l \sim \mathcal{CN}(0, \sigma_l^2), \quad (10)$$

with  $x_l \in \mathcal{C}$  denoting the input symbol for section  $l$ . In (10), the noise variance  $\sigma_l^2$  will be defined shortly. The complex LLR  $L_l^{\text{AWGN}} \in \mathbb{C}$  for the AWGN channel (10) with the uniform a priori probability  $p(x_l) = 1/|\mathcal{C}|$  is given by

$$L_l^{\text{AWGN}} = 2\sqrt{2} \frac{z_l}{\sigma_l^2}. \quad (11)$$

The distribution of the LLR (11) is statistically equivalent to that of the original LLR (9) in the large-system limit, when  $\sigma_l^2$  is given as the solution to a fixed-point equation.

**Theorem 1.** *Focus on section  $l$ , and suppose that  $\{p_{l'}^{\text{dec}}(L) : l' = l, \dots, l + W - 1\}$  have been fed back from the decoder. For any couple  $(k, t)$  that corresponds to indices included in section  $l$ , the LLR (9) given  $\mathbf{H}$ ,  $\{\hat{x}_{k',t}\}$ , and  $x_{k,t} = x$  converges in distribution to (11) given  $x_l = x$  with probability 1 in the large-system limit after taking  $M \rightarrow \infty$ . In evaluating (11),  $\sigma_l^2$  is given by the solution to the fixed-point equation,*

$$\sigma_l^2 = \alpha \left( N_0 + \frac{1}{W} \sum_{w=0}^{W-1} \text{MSE}_{l+w}(\sigma_l^2) \right), \quad (12)$$

with

$$\text{MSE}_{l'}(\sigma_l^2) = \int_{\mathbb{R}^2} \frac{\sigma^2(L_r, L_i) \sigma_l^2}{\sigma^2(L_r, L_i) + \sigma_l^2} p_{l'}^{\text{dec}}(L_r) p_{l'}^{\text{dec}}(L_i) dL_r dL_i, \quad (13)$$

where the a priori variance  $\sigma^2(L_r, L_i)$  is given by (4).

*Proof:* See Appendix A. ■

The function (13) corresponds to the average power of the interference due to the data symbols associated with the  $l'$ th codeword. Thus, the interference power tends to zero when the mass of  $p_{l'}^{\text{dec}}(L)$  concentrates at  $\pm\infty$ . This situation occurs when the  $l'$ th decoder sends the correct hard decision.

Theorem 1 implies that the asymptotic multiuser efficiency (ME) for section  $l$  is given by  $\alpha N_0 / \sigma_l^2$ , and that the analysis of the decoder for section  $l$  reduces to that of decoder for the LDPC-coded AWGN channel (10) with  $W$  signal-to-noise ratio (SNR) levels  $\{1/\sigma_{l'}^2 : l' = l - (W - 1), \dots, l\}$ . This problem can be solved with the Gaussian approximation of the LLRs [9]. See Appendix B for the details.

#### V. NUMERICAL RESULTS AND CONCLUDING REMARKS

The performance of the SC interleaving is compared to that of the conventional random interleaving with  $W = 1$ . In all numerical results, we used (3, 6)-regular LDPC codes [9].

Figure 3 shows the evolution of the asymptotic ME based on Theorem 1. We used the parallel scheduling in order to clarify the behavior of the iterative MUDD. ME close to one

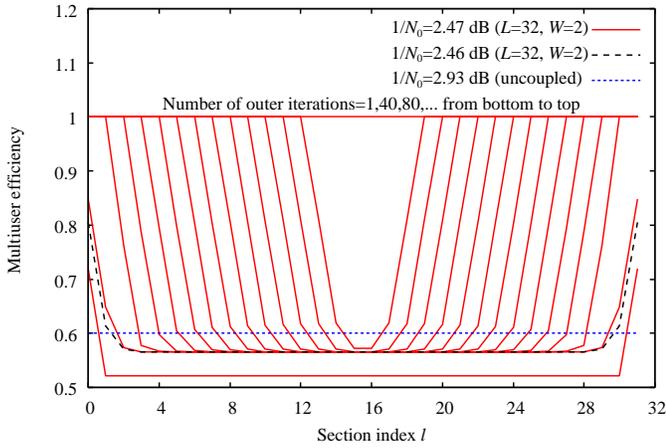


Fig. 3. Multiuser efficiency for the iterative LMMSE MUDD with parallel scheduling for  $\alpha = K/N = 1$  and the number of inner iterations  $J = 100$ .

TABLE I  
DECODING THRESHOLDS FOR SUFFICIENTLY LARGE  $L$ ,  $\alpha = K/N = 1$ ,  
AND THE NUMBER OF INNER ITERATIONS  $J = 100$ .

	$W = 1$	$W = 2$	$W = 3$	$W = 4$
Parallel	2.94 dB	2.47 dB	2.36 dB	2.31 dB
SW	2.94 dB	2.60 dB	2.55 dB	2.53 dB

implies that the inter-stream interference has been eliminated. Consequently, the systems can enjoy the interference-free performance. The ME for the conventional interleaving tends to a value distant from one after sufficiently many iterations when SNR  $1/N_0 = 2.93$  dB. On the other hand, the ME for the SC interleaving can still converge to one for  $1/N_0 = 2.47$  dB, whereas it cannot for  $1/N_0 = 2.46$  dB. These observations imply that the decoding threshold is between 2.46 dB and 2.47 dB, which is defined as the minimum SNR such that the ME converges to one after infinite outer iterations.

Table I lists the decoding thresholds for the parallel and SW scheduling. The chain length  $L$  was set to sufficiently large values to eliminate the influence of the rate loss due to spatial coupling. We find that the SC interleaving with  $W \geq 2$  can improve the decoding threshold compared to the conventional interleaving with  $W = 1$ . Furthermore, the SW scheduling with a decoding delay of  $O(W)$  is slightly inferior to the parallel scheduling with a delay of  $O(L)$ . This implies that there is a tradeoff between the performance and the delay.

Figure 4 shows the average bit error rates (BERs) in decoding for the  $16 \times 16$  MIMO system. Note that the iterative MUDD converges quickly for the SNR regime above the decoding thresholds, although the maximum numbers of inner and outer iterations were set to 100 and 300, respectively. As a fair comparison between the conventional and SC interleavers in terms of the overall rate, we also plotted the BERs for the case of no channel state information (CSI): A random training binary sequence of length 16384 was transmitted for the (non-iterative) LMMSE channel estimation in the MIMO system with the conventional interleaving. For the SC interleaving, on the other hand, 12.5 % of the sequence was used as the training sequence for spatial coupling, and the remaining sequence

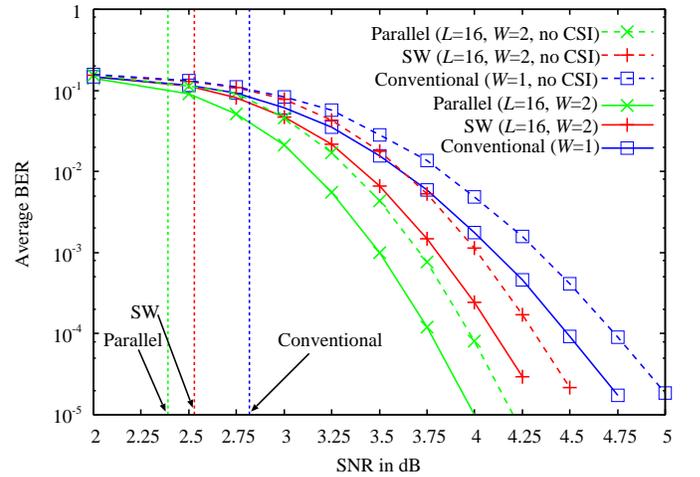


Fig. 4. BER versus SNR  $1/N_0$  for the  $16 \times 16$  MIMO system. Code length  $M = 2048$ , and the numbers of inner and outer iterations  $J \leq 100$  and  $I \leq 300$ , respectively. The vertical lines show the decoding thresholds for the three schemes.

was utilized for the LMMSE channel estimation. We find that the SC interleaving can provide performance gains of 0.5 dB and 0.8 dB for the SW and parallel scheduling, respectively, compared to the conventional random interleaving.

#### APPENDIX A PROOF OF THEOREM 1

In the proof of Theorem 1, we treat the channel matrix  $\mathbf{H}$  and the soft decisions  $\{\hat{x}_{k',t}\}$  as deterministic variables, and omit conditioning with respect to these variables. It is known that the LLR for linear receivers converges in distribution to a Gaussian random variable with probability 1 in the large-system limit, e.g., see [12], [13]. Thus, it is sufficient to evaluate the conditional mean and variance of the LLR (9).

We first calculate the conditional mean  $\mathbb{E}[L_{k,t}^{\text{dem}} | x_{k,t} = x]$ . Substituting (1) into (9) yields

$$\frac{L_{k,t}^{\text{dem}}}{2\sqrt{2}} = \mathbf{c}_{k,t}^H \mathbf{h}_k x_{k,t} + \sum_{k' \neq k} \mathbf{c}_{k,t}^H \mathbf{h}_{k'} (x_{k',t} - \hat{x}_{k',t}) + \mathbf{c}_{k,t}^H \mathbf{n}_t, \quad (14)$$

with  $\mathbf{c}_{k,t} = \mathbf{\Sigma}_{k,t}^{-1} \mathbf{h}_k$  denoting the LMMSE filter. Expression (14) implies

$$\mathbb{E}[L_{k,t}^{\text{dem}} | x_{k,t} = x] = 2\sqrt{2} \mathbf{h}_k^H \mathbf{\Sigma}_{k,t}^{-1} \mathbf{h}_k x, \quad (15)$$

where we have used the assumption that the a priori mean of  $x_{k',t}$  is equal to  $\hat{x}_{k',t}$  for all  $k' \neq k$ .

We next evaluate the conditional variance  $\mathbb{V}[L_{k,t}^{\text{dem}} | x_{k,t}]$ . Using the fact that  $\{x_{k',t}\}$  are regarded as independent random variables for all  $k' \neq k$  in the limit  $M \rightarrow \infty$ , because of random interleaving, we obtain

$$\begin{aligned} \mathbb{V}\left[\frac{L_{k,t}^{\text{dem}}}{2\sqrt{2}} \middle| x_{k,t}\right] &= \sum_{k' \neq k} |\mathbf{c}_{k,t}^H \mathbf{h}_{k'}|^2 (1 - |\hat{x}_{k',t}|^2) + N_0 \|\mathbf{c}_{k,t}\|^2 \\ &= \mathbf{c}_{k,t}^H \mathbf{\Sigma}_{k,t} \mathbf{c}_{k,t}, \end{aligned} \quad (16)$$

with (8). Substituting  $\mathbf{c}_{k,t} = \mathbf{\Sigma}_{k,t}^{-1} \mathbf{h}_k$ , we obtain

$$\mathbb{V}[L_{k,t}^{\text{dem}} | x_{k,t}] = 8 \mathbf{h}_k^H \mathbf{\Sigma}_{k,t}^{-1} \mathbf{h}_k. \quad (17)$$

As observed from (11), on the other hand, the LLR (11) conditioned on  $x_l = x$  for the AWGN channel has mean  $2\sqrt{2}x/\sigma_l^2$  and variance  $8/\sigma_l^2$ . Thus, it is sufficient to prove that  $\mathbf{h}_k^H \Sigma_{k,t}^{-1} \mathbf{h}_k$  converges to  $1/\sigma_l^2$  in the large-system limit, given by the solution to the fixed-point equation (12).

It is worth noting that the quantity  $\mathbf{h}_k^H \Sigma_{k,t}^{-1} \mathbf{h}_k$  is equal to the signal-to-interference ratio (SIR)  $\text{sir}_{k,t}$  for the LLR (9) in the limit  $M \rightarrow \infty$ . In fact, from (14) we obtain

$$\text{sir}_{k,t} = \frac{8|\mathbf{c}_{k,t}^H \mathbf{h}_k|^2}{\mathbb{V}[L_{k,t}^{\text{dem}}|x_{k,t}]} = \mathbf{h}_k^H \Sigma_{k,t}^{-1} \mathbf{h}_k, \quad (18)$$

where we have used  $\mathbf{c}_{k,t} = \Sigma_{k,t}^{-1} \mathbf{h}_k$  and (16). The SIR (18) depends on the channel matrix  $\mathbf{H}$  and the soft decisions  $\{\hat{x}_{k',t}\}$  via (8). Tse and Hanly [11] used random matrix theory to prove<sup>2</sup> that the SIR (18) converges *in probability* to a deterministic value in the large-system limit. Here, we shall present a bit stronger statement [14].

**Theorem 2** ([11]). *Suppose that  $\sigma_l^2$  is the solution to the fixed-point equation*

$$\sigma_l^2 = \alpha \left( N_0 + \int \frac{x\sigma_l^2}{x + \sigma_l^2} dF(x) \right), \quad (19)$$

where  $F(x)$  represents the limiting empirical distribution of the a priori variances  $\{1 - |\hat{x}_{k',t}|^2\}$ ,

$$F(x) = \lim_{K \rightarrow \infty} \frac{1}{K-1} \sum_{k' \neq k} \chi(1 - |\hat{x}_{k',t}|^2 \leq x), \quad (20)$$

with  $\chi$  denoting the indicator function. Then, the SIR (18) converges almost surely to  $1/\sigma_l^2$  in the large-system limit.

Note that the fixed-point equation (19) for the MIMO system is slightly different from the so-called Tse-Hanly equation [11] for code-division multiple-access (CDMA) systems, because of the difference in power normalization.

In order to complete the proof of Theorem 1, we evaluate the limiting empirical distribution (20). From (4), (20) reduces to

$$F(x) = \lim_{K \rightarrow \infty} \frac{1}{K-1} \sum_{k' \neq k} \chi(\sigma^2(\Re[L_{k',t}^{\text{dec}}], \Im[L_{k',t}^{\text{dec}}]) \leq x). \quad (21)$$

Recall that we are focusing on section  $l$  at the output side of the SC interleaver. From the construction of the SC interleaver, the  $W$  decoders from sections  $l$  to  $l+W-1$  feed the LLRs back to the demodulator with equal probability in the large-system limit. Furthermore, the assumption of the random bit-interleaving implies that the LLRs  $\{\Re[L_{k',t}^{\text{dec}}], \Im[L_{k',t}^{\text{dec}}] : k' \neq k\}$  are independent random variables, since we have first taken the limit  $M \rightarrow \infty$ . From the law of large numbers, the empirical distribution (21) converges almost surely to

$$F(x) = \frac{1}{W} \sum_{w=0}^{W-1} \int_{\mathbb{R}^2} \chi(\sigma^2(L_r, L_i) \leq x) \cdot p_{l+w}^{\text{dec}}(L_r) p_{l+w}^{\text{dec}}(L_i) dL_r dL_i, \quad (22)$$

<sup>2</sup> More precisely, they considered the case of real random variables. However, the result is easily generalized to the case of complex random variables.

which implies that the fixed-point equation (19) reduces to (12). Note that the subscripts of the two pdfs in (22) coincide with each other, since each data symbol consists of bits in the same codeword.

## APPENDIX B DENSITY EVOLUTION FOR DECODER

We shall present the DE analysis for the  $l$ th decoder. In Section IV-B, we have proved that the analysis of the  $l$ th decoder reduces to that of the decoder for the LDPC-coded complex AWGN channel with  $W$  SNR levels  $\{1/\sigma_{l'}^2 : l' = l - (W - 1), \dots, l\}$ . Since it is infeasible to trace the exact distribution of LLRs, we follow [9] to approximate the distributions of the LLRs by Gaussian distributions. As shown in Appendix A, the LLR (11) conditioned on  $x_l$  has mean  $2\sqrt{2}x_l/\sigma_l^2$  and variance  $8/\sigma_l^2$ . Since QPSK is used, the products  $\Re[L_i^{\text{AWGN}}] \Re[x_l]$  and  $\Im[L_i^{\text{AWGN}}] \Im[x_l]$  are independent of each other, and follow the real Gaussian distribution with mean  $2/\sigma_l^2$  and variance  $4/\sigma_l^2$ . In the Gaussian approximation of LLRs, the distributions of the LLRs are approximated by this constrained Gaussian distributions.

Without loss of generality, we assume transmission of all-zero codeword. We follow [9] to use the entropy  $h = \psi(m)$  as the parameter that determines the constrained Gaussian distribution, instead of mean  $m > 0$ , given by

$$\psi(m) = \int_{\mathbb{R}} S \left( \frac{e^{L/2}}{e^{L/2} + e^{-L/2}} \right) \frac{1}{\sqrt{4\pi m}} e^{-\frac{(L-m)^2}{4m}} dL, \quad (23)$$

where  $S(p)$  denotes the binary entropy function

$$S(p) = -p \log_2 p - (1-p) \log_2 (1-p). \quad (24)$$

As seen from (3), the entropy  $h = \psi(m)$  is regarded as the average entropy of a binary random variable characterized by a Gaussian-distributed LLR  $L$  with mean  $m$  and variance  $2m$ . Since (23) is monotonically decreasing, the inverse function  $m = \psi^{-1}(h)$  exists.

Let  $h_{c,l}^{(j)}$  denote the entropy for the LLR emitted from each check node in inner iteration  $j$ . Recall that we are focusing on the  $l$ th decoder. We approximate the pdf of the LLR by the Gaussian pdf with mean  $\psi^{-1}(h_{c,l}^{(j)})$  and variance  $2\psi^{-1}(h_{c,l}^{(j)})$ . Since each variable node sends to a check node the sum of LLRs sent by the demodulator and by the other  $(d_v - 1)$  check nodes connected to the variable node, we evaluate the entropy  $h_{v,l,l'}^{(j)}$  for the LLR emitted from a variable node that is connected to the AWGN channel with SNR  $1/\sigma_{l'}^2$  as

$$h_{v,l,l'}^{(j)} = \psi \left( \frac{2}{\sigma_{l'}^2} + (d_v - 1)\psi^{-1}(h_{c,l}^{(j-1)}) \right), \quad (25)$$

with  $h_{c,l}^{(0)} = 1$ . Thus, the average entropy  $h_{v,l}^{(j)}$  emitted from the variable nodes is given by

$$h_{v,l}^{(j)} = \frac{1}{W} \sum_{w=0}^{W-1} \psi \left( \frac{2}{\sigma_{l-w}^2} + (d_v - 1)\psi^{-1}(h_{c,l}^{(j-1)}) \right). \quad (26)$$

Here, we approximate the distribution of the LLRs emitted from the variable nodes by  $\mathcal{N}(\psi^{-1}(h_{v,l}^{(j)}), 2\psi^{-1}(h_{v,l}^{(j)}))$ , although the *true* distribution is the mixture of Gaussian

distributions under the first Gaussian approximation. Note that the same approximation is made in the DE analysis of irregular LDPC codes.

In order to calculate the entropy  $h_{c,l}^{(j)}$ , we use the duality between variable nodes with entropy  $h$  and check nodes with entropy  $(1-h)$  [9]. Exchanging the roles of variable nodes and check nodes, and repeating the derivation of (25), we obtain

$$h_{c,l}^{(j)} = 1 - \psi \left( (d_c - 1) \psi^{-1} (1 - h_{v,l}^{(j)}) \right). \quad (27)$$

The two expressions (26) and (27) correspond to the DE equations for the decoder. The entropy  $h_l^{\text{dec}}$  for the LLR passed from the  $l$ th decoder to the demodulator is given by

$$h_l^{\text{dec}} = \psi \left( d_v \psi^{-1} (h_{c,l}^{(J)}) \right), \quad (28)$$

where  $J$  denotes the total number of inner iterations. This implies that the asymptotic pdf  $p_l^{\text{dec}}(L)$  emitted from the  $l$ th decoder is given by

$$p_l^{\text{dec}}(L) = \frac{1}{\sqrt{4\pi\psi^{-1}(h_l^{\text{dec}})}} e^{-\frac{(L - \psi^{-1}(h_l^{\text{dec}}))^2}{4\psi^{-1}(h_l^{\text{dec}})}}, \quad (29)$$

where  $\psi^{-1}$  denotes the inverse function of (23).

#### REFERENCES

- [1] S. Kudekar, T. Richardson, and R. Urbanke, "Threshold saturation via spatial coupling: Why convolutional LDPC ensembles perform so well over the BEC," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 803–834, Feb. 2011.
- [2] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1046–1061, Jul. 1999.
- [3] J. Boutros and G. Caire, "Iterative multiuser joint decoding: Unified framework and asymptotic analysis," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1772–1793, Jul. 2002.
- [4] G. Caire, R. R. Müller, and T. Tanaka, "Iterative multiuser joint decoding: Optimal power allocation and low-complexity implementation," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1950–1973, Sep. 2004.
- [5] K. Takeuchi, T. Tanaka, and T. Kawabata, "Improvement of BP-based CDMA multiuser detection by spatial coupling," in *Proc. 2011 IEEE Int. Symp. Inf. Theory*, Saint Petersburg, Russia, Aug. 2011, pp. 1489–1493.
- [6] —, "Performance improvement of iterative multiuser detection for large sparsely-spread CDMA systems by spatial coupling," submitted to *IEEE Trans. Inf. Theory*, 2012, [Online]. Available: <http://arxiv.org/abs/1206.5919>.
- [7] C. Schlegel and D. Truhachev, "Multiple access demodulation in the lifted signal graph with spatial coupling," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2459–2470, Apr. 2013.
- [8] D. Truhachev, "Achieving AWGN multiple access channel capacity with spatial graph coupling," *IEEE Commun. Lett.*, vol. 16, no. 5, pp. 585–588, May 2012.
- [9] T. Richardson and R. Urbanke, *Modern Coding Theory*. New York: Cambridge University Press, 2008.
- [10] A. E. Pusane, A. J. Feltström, A. Sridharan, M. Lentmaier, K. Sh. Zigangirov, and D. J. Costello, Jr., "Implementation aspects of LDPC convolutional codes," *IEEE Trans. Commun.*, vol. 56, no. 7, pp. 1060–1069, Jul. 2008.
- [11] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth and user capacity," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 641–657, Mar. 1999.
- [12] D. Guo, S. Verdú, and L. K. Rasmussen, "Asymptotic normality of linear multiuser receiver outputs," *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 3080–3095, Dec. 2002.
- [13] D. Guo and S. Verdú, "Randomly spread CDMA: Asymptotics via statistical physics," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 1983–2010, Jun. 2005.
- [14] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*. Hanover, MA USA: Now Publishers, 2004.