

# Joint Iterative Decoding of Spatially Coupled Low-Density Parity-Check Codes for Position Errors in Racetrack Memories\*

Ryo SHIBATA<sup>†a)</sup>, Nonmember, Gou HOSOYA<sup>††</sup>, and Hiroyuki YASHIMA<sup>††</sup>, Members

**SUMMARY** Racetrack memory (RM) has attracted much attention. In RM, insertion and deletion (ID) errors occur as a result of an unstable reading process and are called position errors. In this paper, we first define a probabilistic channel model of ID errors in RM with multiple read-heads (RHs). Then, we propose a joint iterative decoding algorithm for spatially coupled low-density parity-check (SC-LDPC) codes over such a channel. We investigate the asymptotic behaviors of SC-LDPC codes under the proposed decoding algorithm using density evolution (DE). With DE, we reveal the relationship between the number of RHs and achievable information rates, along with the iterative decoding thresholds. The results show that increasing the number of RHs provides higher decoding performances, although the proposed decoding algorithm requires each codeword bit to be read only once regardless of the number of RHs. Moreover, we show the performance improvement produced by adjusting the order of the SC-LDPC codeword bits in RM.

**key words:** insertion and deletion errors, spatially coupled low-density parity-check code, forward-backward algorithm, iterative decoding

## 1. Introduction

Recently, Racetrack memory (RM) has attracted much attention [1]. Because it has the advantages of an ultra-high storage density and fast data access speed, and is a non-volatile memory, RM is expected to be used as a worthy alternative to hard disk drives in storage devices.

Figure 1 shows the structure of RM. RM has magnetic domains along magnetic nanowires, and each domain can store a bit of data as a magnetization direction. Because the read-head (RH) is fixed and RM has a tape-like structure, the stored data can be read by shifting the domains, which is called a shift operation. The shift operation is accomplished by applying a shift current that moves the domains in one direction. The stored data can then be read bit-by-bit by the RH. Meanwhile, it is also possible to shift by more (less) than a single step by applying a stronger (weaker) current. However, such unstable operations cause imperfect shifts in the domains, called position errors [2]. These errors can be modeled as insertion and deletion errors. A deletion error is an event where the domains are shifted by more than the

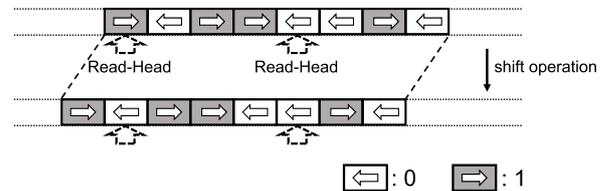


Fig. 1 Racetrack memory with multiple read-heads and shift operation.

range of one domain, and thus one of the domains is not read, which results in the deletion of the bit stored in this domain. On the other hand, an insertion error is an event where the domains were not successfully shifted. Thus, the same domain is re-read, which results in the insertion of the same bit. At first glance, the problem does not seem to be different from conventional insertion and deletion errors, and the codes for them should be applicable. However, another approach to deal with this problem involves the utilization of the distinctive features of RM, where multiple RHs can be used to read domains, and insertion and deletion errors occur simultaneously in each RH. In [3], the algebraic coding for the insertion and deletion errors with multiple RHs was studied, where reading the stored data with each of  $a$  RHs made it possible to correct  $a - 1$  deletions with at most a single bit of redundancy. However, the algebraic coding in [3] required that all  $a$  of the RHs read the codeword. Meanwhile, we are also interested in modern (probabilistic) coding theory, e.g., low-density parity-check (LDPC) codes [4], [5], for the insertion and deletion errors in RM with multiple RHs.

LDPC codes for conventional insertion and deletion channels were studied in [6]–[9]. The concatenation of the outer LDPC codes and inner watermark codes was investigated in [6], whereas the concatenation of LDPC codes and marker codes was investigated in [7], [8]. Recently, it was proven by Kudekar et al. that spatially coupled LDPC (SC-LDPC) [10], [11] codes are capacity achieving on memoryless binary-input symmetric-output channels under message-passing decoding [11]. Consequently, the principle of spatial coupling has attracted significant attention and has also been applied in channels with memory. Indeed, it was shown in [9] that SC-LDPC codes have a much better performance than LDPC block (LDPC-BLK) codes.

The methods used in [6]–[9] can be viewed as LDPC codes for a hidden Markov channel [12], where decodings can be performed by utilizing the message-passing algorithm (belief propagation) on a factor graph [13], in which a chan-

Manuscript received January 29, 2018.

Manuscript revised June 22, 2018.

<sup>†</sup>The author is with Dept. of Management Science, Graduate School of Engineering, Tokyo University of Science, Tokyo, 125-8585 Japan.

<sup>††</sup>The authors are with Dept. of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science, Tokyo, 125-8585 Japan.

\*A portion of this paper was presented in [14] and [15].

a) E-mail: ryo12049@gmail.com

DOI: 10.1587/transfun.E101.A.2055

nel and codes graphs are joined together.

In this paper, we define a channel model of the insertion and deletion errors in RM with multiple RHs, where substitution errors also occur, as hidden Markov channels, and then propose the joint iterative decoding of SC-LDPC codes for such a channel. The proposed decoding algorithm requires that each codeword bit be read only once, regardless of the number of RHs. From now on, this channel is referred to as the “multi-heads insertion and deletion (MHID) channel.” The main contributions of this paper are 1) the MHID channel model, 2) joint iterative decoding for these channels, 3) density evolution analysis for computing the thresholds of SC-LDPC codes over these channels, and 4) performance improvement produced by adjusting the order of the bits in the SC-LDPC codewords in RM.

The remainder of this paper is organized as follows. In Sect. 2, we describe the MHID channel model. In Sect. 3, we propose a coding scheme for MHID channels, which includes the joint iterative decoding, as well as the order of the codeword bits in RM. Section 4 explains the density evolution analysis for MHID channels. Section 5 shows the numerical results for the iterative decoding thresholds and achievable information rates for MHID channels, along with simulation results. Furthermore, the performance improvement produced by adjusting the order of the SC-LDPC codeword bits in RM is shown in Sect. 5. Section 6 concludes the paper.

## 2. MHID Channels

For integers  $i$  and  $j$  such that  $i \leq j$ , a set of consecutive integers  $\{i, \dots, j\}$  is denoted by  $[i, j]$ . When  $j < i$ ,  $[i, j]$  equals the empty set  $\emptyset$ . For a given vector  $\mathbf{v} = (v_1, v_2, \dots, v_k)$  and  $i, j \in [1, k]$ , a subvector of  $\mathbf{v}$  is denoted by  $[\mathbf{v}]_i^j = (v_i, v_{i+1}, \dots, v_j)$ , where  $i \leq j$ . When  $j < i$ ,  $[\mathbf{v}]_i^j$  equals the empty string  $\epsilon$ . Let  $\mathbb{N}$ ,  $\mathbb{R}$ , and  $\mathbb{Z}$  be sets of all natural numbers, all real numbers, and all integers, respectively. From now on, we refer to data before reading by the RH as a transmitted sequence and data after reading by the RH as a received sequence.

We provide the MHID channels described in section 1 as hidden Markov channels. In RM, one shift operation moves all the domains, and it is possible to employ multiple RHs to read the domains. Hence, the insertion and deletion errors occur at the same timing in each RH. The MHID channels introduced here certainly have these features. In addition to position errors (insertion and deletion errors), we consider substitution errors such as caused by head sensing flaws in RM [2]. Let  $a$  denote the number of RHs on RM. Let  $\mathbf{w}^k = (w_1^k, \dots, w_n^k)$  denote a transmitted sequence of length  $n$  on the  $k$ th RH, where  $w_t^k \in \{0, 1\}$  and  $k \in [1, a]$ . Let  $\mathbf{r}^k = (r_1^k, \dots, r_{n'}^k)$  denote a received sequence of length  $n'$  on the  $k$ th RH, where  $r_t^k \in \{0, 1\}$  and  $k \in [1, a]$ . It is assumed that  $\mathbf{w}^1, \dots, \mathbf{w}^a$  are independent and uniformly distributed (i.u.d.) sequences and independent from each other. When transmitting the  $t$ th bits of the transmitted sequences, either one of the following events occurs in MHID channels as:

### 1) Deletion

$w_t^k$  is not transmitted with probability  $P_{\text{del}}$ .

### 2) Insertion

$w_t^k$  is transmitted twice with probability  $P_{\text{ins}}$ , having a substitution error with probability  $P_{\text{subs}}$  at each time.

### 3) Transmission

$w_t^k$  is transmitted with probability  $P_{\text{trans}} = 1 - P_{\text{del}} - P_{\text{ins}}$ , having a substitution error with probability  $P_{\text{subs}}$ .

It is assumed that the insertion, deletion, and transmission events are all independent and identically distributed (i.i.d.) and that the positions of the deletion and insertion errors are unknown to the receiver.

While the unobservable insertions and deletions make the length of the received sequence  $n'$  vary and maybe differ from the length of the transmitted sequence  $n$ ,  $n'$  is equivalent to the number of sensed bits by RH through shift operations for reading  $\mathbf{w}^k$ . Now, we express the difference of the lengths as a state of a hidden Markov model. Let  $s_t$  denote the channel state at time  $t \in [0, n]$ , where  $s_t \in [-S, +S]$ ,  $S \in \mathbb{N}$  and  $s_0 = 0$ . Channel state transitions are defined as follows:

- 1) When the insertion event occurs,  $s_t = s_{t-1} + 1$ .
- 2) When the deletion event occurs,  $s_t = s_{t-1} - 1$ .
- 3) When the transmission event occurs,  $s_t = s_{t-1}$ .

Thus,  $s_t$  is an event where both the first  $t$  bits of the transmitted sequences are transmitted and the first  $t + s_t$  bits of the received sequences are received. In other words, all of the RHs read  $t + s_t$  bits while  $t$  bits being transmitted. The length of the received sequence can be expressed as  $n' = n + s_n$ . This gives the final state  $s_n = n' - n$ . Therefore, the initial and final state probabilities are given by the following:

$$p(s_0) = \begin{cases} 1 & (s_0 = 0) \\ 0 & (\text{otherwise}), \end{cases} \quad (1)$$

$$p(s_n) = \begin{cases} 1 & (s_n = n' - n) \\ 0 & (\text{otherwise}). \end{cases} \quad (2)$$

The state transition probability for  $t \in [1, n]$  is expressed as follows:

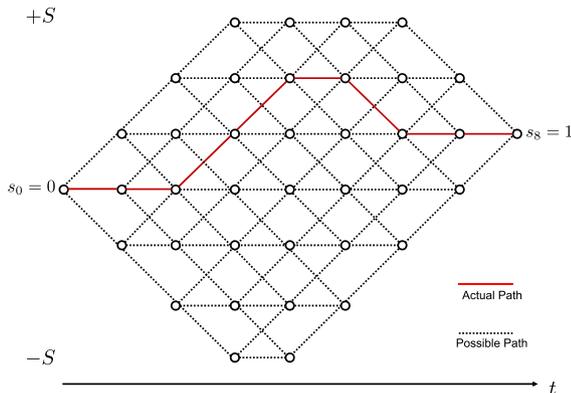
$$p(s_t | s_{t-1}) = \begin{cases} P_{\text{ins}} & (s_t = s_{t-1} + 1, |s_{t-1}| < S) \\ P_{\text{ins}} + P_{\text{del}} & (s_t = s_{t-1} + 1, s_{t-1} = -S) \\ P_{\text{del}} & (s_t = s_{t-1} - 1, |s_{t-1}| < S) \\ P_{\text{ins}} + P_{\text{del}} & (s_t = s_{t-1} - 1, s_{t-1} = +S) \\ P_{\text{trans}} & (s_t = s_{t-1}) \\ 0 & (\text{otherwise}). \end{cases} \quad (3)$$

The output probability for  $t \in [1, n]$  and  $k \in [1, a]$  is given by the following:

$$p([\mathbf{r}^k]_{t+s_{t-1}}^{t+s_t} | w_t^k, s_{t-1}, s_t)$$

**Table 1** Example of inputs, outputs, events and channel states in MHID channels when  $a = 2$  and  $n = 8$ .

| $t$   | 0 | 1 | 2 | 3  | 4  | 5 | 6          | 7 | 8 |
|-------|---|---|---|----|----|---|------------|---|---|
| $w^1$ |   | 0 | 1 | 0  | 0  | 1 | 0          | 1 | 1 |
| $w^2$ |   | 1 | 1 | 0  | 1  | 0 | 0          | 1 | 0 |
| Event |   | T | T | I  | I  | T | D          | T | T |
| $s$   | 0 | 0 | 0 | 1  | 2  | 2 | 1          | 1 | 1 |
| $r^1$ |   | 0 | 1 | 00 | 00 | 1 | $\epsilon$ | 1 | 1 |
| $r^2$ |   | 1 | 1 | 00 | 11 | 0 | $\epsilon$ | 1 | 0 |

**Fig. 2** Trellis diagram corresponding to Table 1.

$$= \begin{cases} \prod_{i \in [0,1]} p(r_{t+s_t-i}^k | w_t^k) & (s_t = s_{t-1} + 1) \\ 1 & (s_t = s_{t-1} - 1) \\ p(r_{t+s_t}^k | w_t^k) & (s_t = s_{t-1}) \\ 0 & (\text{otherwise}), \end{cases} \quad (4)$$

where

$$p(r_j^k | w_z^k) = \begin{cases} 1 - P_{\text{subs}} & (r_j^k = w_z^k) \\ P_{\text{subs}} & (\text{otherwise}). \end{cases} \quad (5)$$

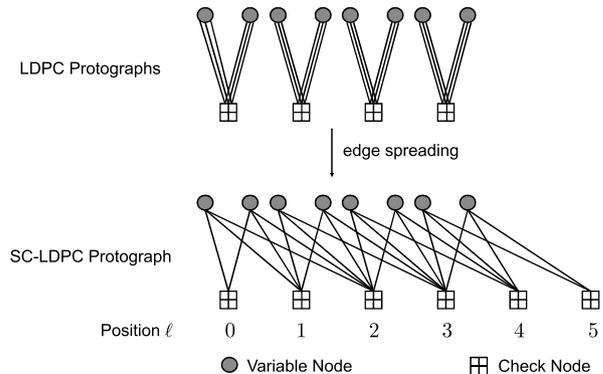
An example of the transmitted and received sequences with  $a = 2$  and  $n = 8$  is provided in Table 1, where the insertion, deletion, and transmission events are denoted by the letters ‘‘I,’’ ‘‘D,’’ and ‘‘T,’’ respectively. Figure 2 shows a trellis diagram corresponding to Table 1. As shown in Table 1, each event has occurred simultaneously for all  $k \in \{1, 2\}$ .

### 3. Coding for MHID Channels

In this section, we first briefly review SC-LDPC codes. Then, we propose a coding scheme for MHID channels, which includes a description of the transmitted sequences, joint iterative decoding algorithm, and the order of the codeword bits in RM.

#### 3.1 Protograph-Based $(d_v, d_c, L)$ SC-LDPC Codes

We introduce protograph-based  $(d_v, d_c, L)$  SC-LDPC codes [10]. These codes are obtained by coupling  $(d_v, d_c)$ -regular LDPC codes, where  $d_c$  is the check node degree and  $d_v$  is the variable node degree,  $\frac{d_c}{d_v} = k \in \mathbb{N}$ . A protograph [16] is

**Fig. 3**  $(3, 6)$  LDPC protographs and  $(3, 6, 4)$  SC-LDPC protograph.

a relatively small bipartite graph from which a larger graph can be obtained by using a copy-and-permute procedure.

To construct an  $(d_v, d_c, L)$  SC-LDPC protograph,  $L$  copies of an  $(d_v, d_c)$ -regular LDPC protograph are coupled using a procedure called edge spreading, where  $L$  is the coupling length. The  $L$  protographs are indexed by  $\ell$ . We connect each of  $k$  variable nodes at position  $\ell$  to the check nodes at positions  $\ell, \ell + 1, \dots, \ell + d_v - 1$ . For instance, Fig. 3 shows the  $(3, 6, 4)$  SC-LDPC protograph construction. Once the coupled protograph is produced,  $M$ -fold graph-lifting provides  $(d_v, d_c, L)$  SC-LDPC codes with a code length  $N = kLM$  and a code rate of

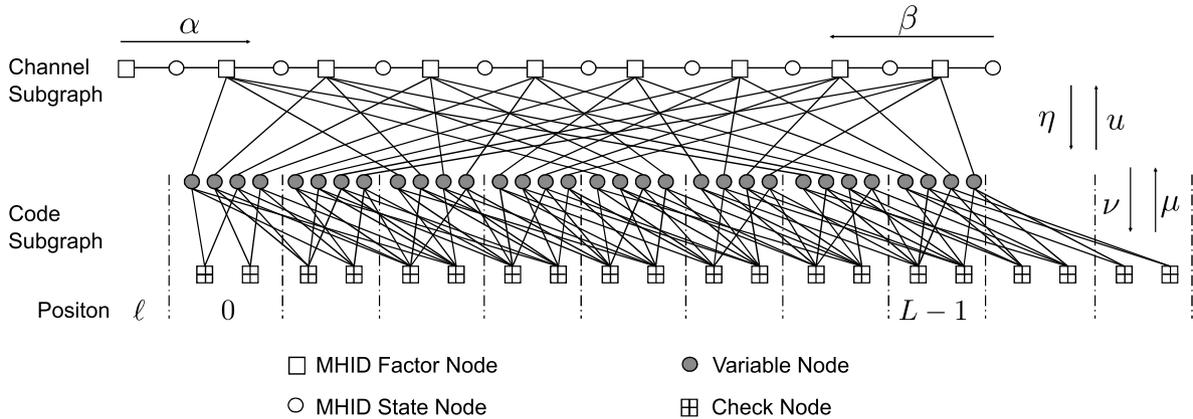
$$R_{(d_v, d_c, L)} = 1 - \frac{(L + d_v - 1)}{kL}. \quad (6)$$

The code rate  $R_{(d_v, d_c, L)}$  tends to the uncoupled rate  $1 - \frac{d_v}{d_c}$  when  $L \rightarrow \infty$ . The  $(d_v, d_c, L)$  SC-LDPC protograph has a so-called *structured irregularity*, where the check nodes located at the ends are connected to a smaller number of variable nodes than those in the middle. This irregularity leads the message-passing decoding to propagate in a wave-like fashion from the ends toward the center [11].

#### 3.2 Transmitted Sequences

A codeword of the SC-LDPC code  $C_E$  is denoted by  $\mathbf{x}^E = (x_1^E, \dots, x_N^E)$ . A permutation on a set  $[1, N]$  is denoted by  $\pi = (\pi(1) \pi(2) \dots \pi(N))$  in one-line notation. The stored data in RM are denoted by  $\mathbf{x} = (x_1, \dots, x_N)$ , which is the rearrangement of the codeword  $\mathbf{x}^E$  using the permutation  $\pi$ , i.e.,  $\mathbf{x} = (x_1, \dots, x_N) = (x_{\pi(1)}^E, \dots, x_{\pi(N)}^E)$ . We assume that  $a$  RHs are fixed on RM at equal intervals in a range of  $\mathbf{x}$  and  $N$  is divisible by  $a$ .

Through shift operations, each of the  $a$  RHs reads a part of  $\mathbf{x}$  via the MHID channels. The  $k$ th ( $k = 1, \dots, a$ ) transmitted sequence (data to be read by the  $k$ th RH) is denoted by  $\mathbf{w}^k = [\mathbf{x}]_{(k-1)n+1}^{kn}$ , where  $n = \frac{N}{a}$ . Thus, each bit of  $\mathbf{x}$  is read only once, regardless of the number of RHs,  $a$ . The index of  $\mathbf{x}$  corresponding to  $w_t^k$  is given by the map  $\xi : [1, a] \times [1, n] \rightarrow [1, N]$ ,  $\xi_t^k = (k-1)n + t$ , i.e.,  $w_t^k = x_{\xi_t^k} = x_{\pi(\xi_t^k)}^E$  and  $\mathbf{w}^k = [\mathbf{x}]_{\xi_1^k}^{\xi_n^k} = (x_{\pi(\xi_1^k)}^E, \dots, x_{\pi(\xi_n^k)}^E)$ .



**Fig. 4** Joint factor graph of channel subgraph and code subgraph, where (3, 6, 8) SC-LDPC code with  $M = 2$ , number of RHs  $a = 4$ , and  $\pi = (1\ 2\ \dots\ N)$ .

**Example 1:** When  $N = 16$ ,  $a = 2$ , and  $\mathbf{x} = (0100101111010010)$ , we obtain  $\mathbf{w}^1 = (01001011)$  and  $\mathbf{w}^2 = (11010010)$ .

### 3.3 Joint Iterative Decoding

To estimate the codeword  $\mathbf{x}^E$  from the received sequence  $\mathbf{r}^k$  ( $k = 1, \dots, a$ ), we evaluate  $\hat{x}_t \in \{0, 1\}$  that maximizes a posteriori probability  $p(x_t^E | \mathbf{r}^1, \dots, \mathbf{r}^a)$ .  $p(x_t^E | \mathbf{r}^1, \dots, \mathbf{r}^a)$  can be computed by marginalizing a joint probability  $p(\mathbf{x}^E, \mathbf{r}^1, \dots, \mathbf{r}^a, \mathbf{s})$ , where  $\mathbf{s} = (s_0, s_1, \dots, s_n)$ . In this subsection, we introduce the factorization of the joint probability  $p(\mathbf{x}^E, \mathbf{r}^1, \dots, \mathbf{r}^a, \mathbf{s})$ , and give its counterpart factor graph. After that, we propose a joint iterative decoding algorithm for SC-LDPC codes over MHID channels.

Based on the definition of the conditional probability and conditional independence between random variables, the joint probability  $p(\mathbf{x}^E, \mathbf{r}^1, \dots, \mathbf{r}^a, \mathbf{s})$  can be factorized as follows:

$$\begin{aligned}
 & p(\mathbf{x}^E, \mathbf{r}^1, \dots, \mathbf{r}^a, \mathbf{s}) \\
 &= \prod_{k \in [1, a]} p(\mathbf{r}^k | \mathbf{x}^E, \mathbf{r}^{k+1}, \dots, \mathbf{r}^a, \mathbf{s}) p(\mathbf{x}^E, \mathbf{s}) \\
 &= \prod_{k \in [1, a]} p(\mathbf{r}^k | x_{\pi(\xi_1^k)}^E, \dots, x_{\pi(\xi_n^k)}^E, \mathbf{s}) p(\mathbf{x}^E, \mathbf{s}) \\
 &= \prod_{k \in [1, a]} p(\mathbf{r}^k | x_{\pi(\xi_1^k)}^E, \dots, x_{\pi(\xi_n^k)}^E, \mathbf{s}) p(\mathbf{x}^E) p(\mathbf{s}), \quad (7)
 \end{aligned}$$

where in the second step we have used the fact that  $\mathbf{r}^k$  depends only on  $\mathbf{s}$  and  $\mathbf{w}^k = (x_{\pi(\xi_1^k)}^E, \dots, x_{\pi(\xi_n^k)}^E)$ . We can further express (7) based on the definition of MHID channels, a Markov property, and a code indicator function. The probabilities  $p(\mathbf{s})$  and  $p(\mathbf{r}^k | x_{\pi(\xi_1^k)}^E, \dots, x_{\pi(\xi_n^k)}^E, \mathbf{s})$  are expressed as follows:

$$\begin{aligned}
 & p(\mathbf{r}^k | x_{\pi(\xi_1^k)}^E, \dots, x_{\pi(\xi_n^k)}^E, \mathbf{s}) \\
 &= \prod_{t \in [1, n]} p([\mathbf{r}^k]_{t+s_{t-1}}^{t+s_t} | x_{\pi(\xi_t^k)}^E, s_{t-1}, s_t), \quad (8)
 \end{aligned}$$

$$p(\mathbf{s}) = p(s_0) \prod_{t \in [1, n]} p(s_t | s_{t-1}). \quad (9)$$

Let  $H$  denote a parity-check matrix with  $m$  rows and  $N$  columns for the SC-LDPC code  $C_E$ , where  $\mathbf{h}^i$  is the  $i$ th row of  $H$ . A code indicator function for  $C_E$  is denoted by,  $i \in [1, m]$ ,

$$h_i(\mathbf{x}^E) = \begin{cases} 1 & (\mathbf{x}^E (\mathbf{h}^i)^T = 0) \\ 0 & (\text{otherwise}), \end{cases} \quad (10)$$

where  $T$  denotes the transpose of a matrix. The probability  $p(\mathbf{x}^E)$  is expressed as follows:

$$p(\mathbf{x}^E) = \frac{1}{|C_E|} \prod_{j \in [1, m]} h_j(\mathbf{x}^E), \quad (11)$$

where  $|C_E|$  is the total number of codewords in  $C_E$ . Substituting (8), (9), and (11) into (7), we can express the joint probability  $p(\mathbf{x}^E, \mathbf{r}^1, \dots, \mathbf{r}^a, \mathbf{s})$  as follows:

$$\begin{aligned}
 & p(\mathbf{x}^E, \mathbf{r}^1, \dots, \mathbf{r}^a, \mathbf{s}) = \\
 & p(s_0) \prod_{t \in [1, n]} p(s_t | s_{t-1}) \prod_{k \in [1, a]} p([\mathbf{r}^k]_{t+s_{t-1}}^{t+s_t} | x_{\pi(\xi_t^k)}^E, s_{t-1}, s_t) \\
 & \quad \times \frac{1}{|C_E|} \prod_{j \in [1, m]} h_j(\mathbf{x}^E). \quad (12)
 \end{aligned}$$

Figure 4 shows the joint factor graph induced by the factorization (12). The graph is composed of two subgraphs for MHID channels and SC-LDPC codes, where the edge connections between the two subgraphs depend on the order of the codeword bits in RM (i.e., the permutation  $\pi$ ). When using  $a$  RHs, an MHID factor node is connected to  $a$  different variable nodes transmitted at the same time instant.

Next, we describe the joint iterative decoding algorithm. The decoding is a message-passing algorithm on the joint factor graph, where the forward-backward (FB) algorithm in the channel subgraph and the log-domain sum-product decoding in the code subgraph are executed alternately and iteratively.

In the channel subgraph, we compute the messages  $\eta$  using the FB algorithm, where the state transitions are estimated by considering both the received sequences for all RHs and the messages  $u$  (these are available from the previous iteration of the sum-product decoding). Using the messages  $u$ , the priori probabilities of the FB algorithm are computed by

$$p(x_{\pi(\xi_t^k)}^E = 0) = \frac{\exp[u_{\pi(\xi_t^k)}]}{1 + \exp[u_{\pi(\xi_t^k)}]}, \quad (13)$$

$$p(x_{\pi(\xi_t^k)}^E = 1) = \frac{1}{1 + \exp[u_{\pi(\xi_t^k)}]}. \quad (14)$$

Let  $\alpha(t, s_t)$  and  $\beta(t, s_t)$  be the forward and backward probabilities, respectively. These can be evaluated recursively as follows:

$$\begin{aligned} \alpha(t, s_t) = & \sum_{s_{t-1}} \prod_{k \in [1, a]} \sum_{x_{\pi(\xi_t^k)}^E} p(x_{\pi(\xi_t^k)}^E) p([\mathbf{r}^k]_{t+s_{t-1}}^{t+s_t} | x_{\pi(\xi_t^k)}^E, s_{t-1}, s_t) \\ & \times p(s_t | s_{t-1}) \alpha(t-1, s_{t-1}), \end{aligned} \quad (15)$$

$$\begin{aligned} \beta(t, s_t) = & \sum_{s_{t+1}} \prod_{k \in [1, a]} \sum_{x_{\pi(\xi_{t+1}^k)}^E} p(x_{\pi(\xi_{t+1}^k)}^E) p([\mathbf{r}^k]_{t+1+s_t}^{t+1+s_{t+1}} | x_{\pi(\xi_{t+1}^k)}^E, s_t, s_{t+1}) \\ & \times p(s_{t+1} | s_t) \beta(t+1, s_{t+1}). \end{aligned} \quad (16)$$

From the definition of the MHID channels, the initial and final conditions are given as follows:

$$\alpha(0, s_0) = \begin{cases} 1 & (s_0 = 0) \\ 0 & (\text{otherwise}), \end{cases} \quad (17)$$

$$\beta(n, s_n) = \begin{cases} 1 & (s_n = n' - n) \\ 0 & (\text{otherwise}). \end{cases} \quad (18)$$

Consequently, the messages  $\eta$  can be computed as follows:

$$\eta_{\pi(\xi_t^k)} = \ln \frac{\gamma_t^k(0)}{\gamma_t^k(1)}, \quad (19)$$

where  $\gamma_t^k$  is evaluated using  $\alpha(t, s_t)$  and  $\beta(t, s_t)$  as follows:

$$\begin{aligned} \gamma_t^k(z) = & \sum_{s_t} \sum_{s_{t-1}} \alpha(t-1, s_{t-1}) \beta(t, s_t) p(s_t | s_{t-1}) \\ & \times p([\mathbf{r}^k]_{t+s_{t-1}}^{t+s_t} | x_{\pi(\xi_t^k)}^E = z, s_{t-1}, s_t) \\ & \times \prod_{i \in [1, a] \setminus k} \sum_{x_{\pi(\xi_t^i)}^E} p(x_{\pi(\xi_t^i)}^E) p([\mathbf{r}^i]_{t+s_{t-1}}^{t+s_t} | x_{\pi(\xi_t^i)}^E, s_{t-1}, s_t). \end{aligned} \quad (20)$$

After the FB algorithm, we perform the sum-product decoding with the messages  $\eta$  in (19) as inputs. The messages  $\mu$  and  $\nu$  are computed for all the check nodes and all the variable nodes, and then the messages  $u$  are computed. Then,

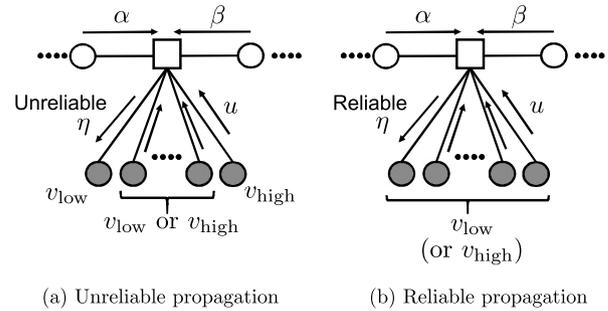


Fig. 5 Edge connections between MHID factor node and variable nodes.

the next iteration of the FB algorithm is performed. The algorithm is iteratively executed until the estimated codeword fulfills the parity check equations or the number of iterations reaches a predetermined number.

### 3.4 Order of Codeword Bits in RM

As described in 3.3, the order of the codeword bits in RM determines the edge connections between the MHID factor nodes and variable nodes. In this paper, we consider two types of orderings. The first is the ordering given by the permutation  $\pi_1 = (1 \ 2 \ \dots \ N)$ , i.e., the stored data  $\mathbf{x}$  have the same order as  $\mathbf{x}^E$ , and the factor graph given by  $\pi_1$  is depicted in Fig. 4. The second is an appropriate order that considers the wave-like decoding of SC-LDPC codes. Let  $v_{\text{low}}$  ( $v_{\text{high}}$ ) be a variable node that connects to low (high) degree check nodes of  $(d_v, d_c, L)$  SC-LDPC codes. Namely, a variable node at positions  $\ell = 0, \dots, d_v - 2, L - d_v + 1, \dots, L - 1$ , is  $v_{\text{low}}$ , and a variable node at other positions is  $v_{\text{high}}$ . In the joint factor graph, when  $a \geq 2$ , an MHID factor node connects to  $a$  different variable nodes transmitted at the same time instant. As shown in Fig. 5(a), if both  $v_{\text{low}}$  and  $v_{\text{high}}$  are connected to the same MHID factor node, the wave-like decoding would stop halfway through because the messages  $\eta$  into  $v_{\text{low}}$  would be updated by unreliable messages  $u$  from  $v_{\text{high}}$ . To avoid such unreliable propagations, we consider edge connections so that an MHID factor node is connected to only  $v_{\text{low}}$ 's or  $v_{\text{high}}$ 's, as shown in Fig. 5(b). The joint factor graph shown in Fig. 6 obtains such edge connections by connecting variable nodes in turn from left to right to MHID factor nodes with assumption of  $\frac{kM}{a} \in \mathbb{N}$ . The counterpart order is given by the permutation  $\pi_2$ . The element of the permutation  $\pi_2$  corresponding to  $\xi_t^k$  is expressed as

$$\pi_2(\xi_t^k) = k + (t-1)a. \quad (21)$$

The ordering given by  $\pi_2$  can exploit the structured irregularity of SC-LDPC codes and cause the joint iterative decoding algorithm to proceed in a wave-like fashion from the ends toward the center.

## 4. Density Evolution

The density evolution (DE) [5], [18] is known as a technique to analyze the convergence behavior and performance

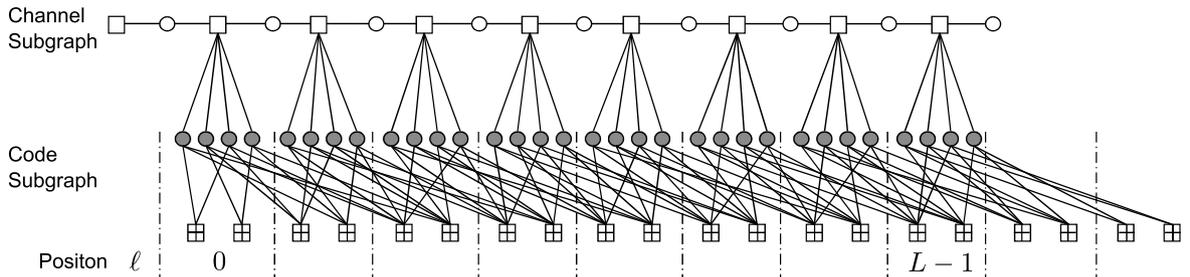


Fig. 6 Joint factor graph for (3, 6, 8) SC-LDPC code with  $M = 2$ , number of RHs  $a = 4$ , and  $\pi_2$ .

of LDPC codes under message-passing decodings. This method involves tracking the probability density function of the messages exchanged in factor graphs throughout the iterations.

In this section, we introduce the DE analysis of  $(d_v, d_c, L)$  SC-LDPC codes under the proposed joint iterative decoding. Tracking the densities of the messages  $\mu$  and  $\nu$  is the same as the DE analysis of  $(d_v, d_c, L)$  SC-LDPC codes [10]. Furthermore, in our joint iterative decoding, it is necessary to track the densities of the messages  $u$  and  $\eta$  in each iteration. The density of  $u$  outgoing from the  $t$ th variable node at position  $\ell$  in the  $i$ th iteration is obtained by convolving the densities of  $\mu$  incoming to the  $t$ th variable node at position  $\ell$ , in the  $i$ th iteration. Because no closed-form expression is known for evolving densities through Markov processes, we adopt a Monte Carlo method to obtain the densities of  $\eta$ . This approach for intersymbol interference channels and LDPC-BLK codes was suggested in [19]. We perform the FB algorithm (15)–(20) with the realizations of  $u$  obtained from their densities in the  $(i - 1)$ th iteration. Here, the transmitted sequences must be randomly chosen i.u.d. sequences, and their length  $n$  must be very long to avoid trellis boundary effects. After obtaining the messages  $\eta$ , we create histograms of  $\eta$  for each variable node in positions  $\ell = 0, \dots, L - 1$ , which are regarded as the densities of  $\eta$  in the  $i$ th iteration. With the DE, we can evaluate the iterative decoding threshold, which is the threshold for the probability of decoding an arbitrarily small error when  $i \rightarrow \infty$  for  $(d_v, d_c, L)$  SC-LDPC codes under the joint iterative decoding with a given number of  $a$ . Note that the iterative decoding threshold here is the maximum value of the sum of the insertion and deletion error probabilities  $P_{\text{ins}} + P_{\text{del}}$  in the MHID channels.

## 5. Results and Discussion

This section presents numerical results, including the achievable information rate for MHID channels and iterative decoding thresholds. We also show the simulation results. Throughout this section, we refer to the MHID channel with  $P_{\text{ins}} = P_{\text{del}}$  and  $S = 4$ .

### 5.1 Achievable Information Rate

Let us compute the achievable information rate for MHID

channels for a given number of RHs  $a$ . In this paper, we consider the achievable information rate with i.u.d. inputs, which is called the symmetric information rate (SIR). The SIR can be viewed as a lower bound on the channel capacity. For channels with memory, the SIR can be numerically evaluated using the Arnold–Loeliger method [17], which is also applicable to MHID channels. The SIR for MHID channels is given by

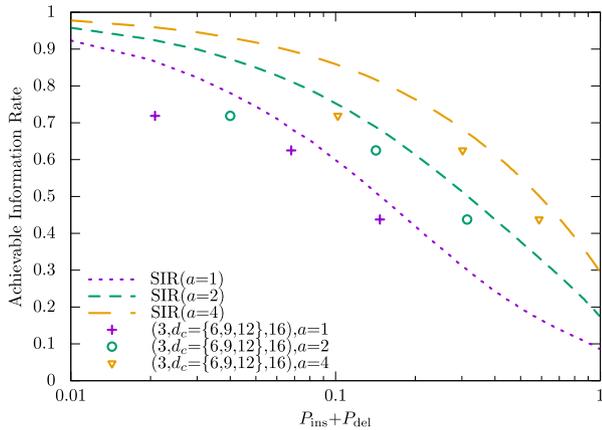
$$\begin{aligned} I_{\text{SIR}} &= \lim_{n \rightarrow \infty} \frac{1}{an} I(\mathbf{w}^1, \dots, \mathbf{w}^a; \mathbf{r}^1, \dots, \mathbf{r}^a) \\ &= \lim_{n \rightarrow \infty} \frac{1}{an} (-\log p(\mathbf{w}^1, \dots, \mathbf{w}^a) - \log p(\mathbf{r}^1, \dots, \mathbf{r}^a) \\ &\quad + \log p(\mathbf{w}^1, \dots, \mathbf{w}^a, \mathbf{r}^1, \dots, \mathbf{r}^a)). \end{aligned} \quad (22)$$

For details of the evaluation of (22), see [17].

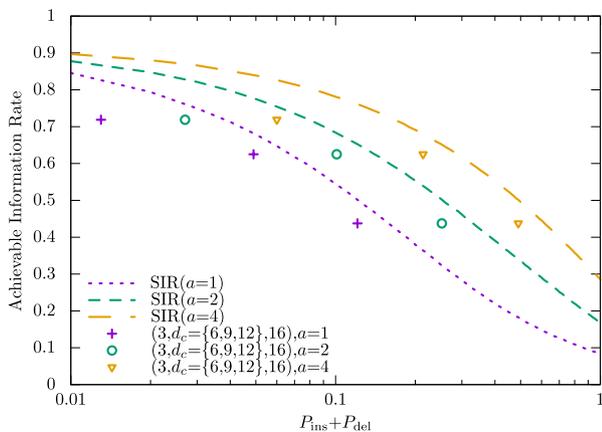
Figures 7, 8, and 9 show the SIRs for MHID channels with different numbers of RHs  $a$  evaluated using the Arnold–Loeliger method [17]. From Figs. 7 and 8, we observe that the SIRs with  $a = 1$  rapidly decreases when the insertion and deletion error probabilities increase. Meanwhile, an increase in the number of RHs greatly improves the SIRs and causes a gradual decrease in the SIRs, which enables reliable communication in an RM influenced by position errors. Unlike the behavior of the insertion and deletion errors ( $P_{\text{ins}}$  and  $P_{\text{del}}$ ), the substitution errors ( $P_{\text{subs}}$ ) dramatically degrades the SIRs of the MHID channels, even if the number of RHs increases, as shown in Fig. 9.

### 5.2 Density Evolution

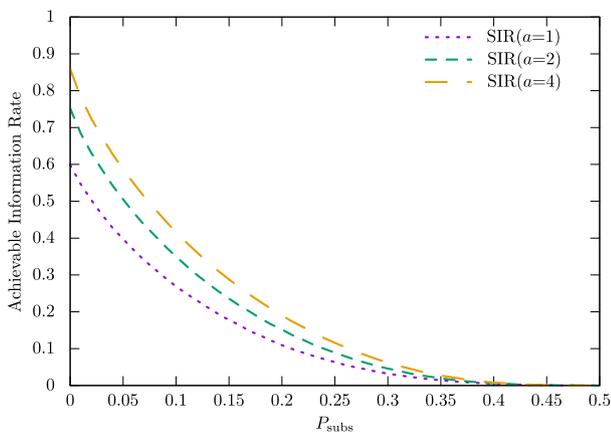
Table 2 lists the iterative decoding thresholds for different numbers of RHs,  $a = 1, 2, 4$ . For any LDPC codes listed in Table 2, the thresholds increase as the number of RHs increases. An MHID factor node is connected to  $a$  different variable nodes when using  $a$  RHs. Thus, the FB algorithm estimates each state transition using  $a$  messages from connected variable nodes. Moreover, a message  $\eta$  is updated by  $a - 1$  different messages  $u$  (the sum-product decoding results in the previous iteration), in addition to the forward and backward probabilities. Such estimations of the state transitions and messages  $\eta$  enhance the reliability of the FB algorithm, and thus improve the overall joint iterative decoding performance. Therefore, the proposed joint decoding algorithm can exploit the feature of MHID channels, where insertion and deletion errors occur simultaneously in each RH.



**Fig. 7** Symmetric information rates and thresholds for  $(3, d_c = \{6, 9, 12\}, 16)$  SC-LDPC codes with  $\pi_2$  when  $P_{\text{subs}} = 0$ .



**Fig. 8** Symmetric information rates and thresholds for  $(3, d_c = \{6, 9, 12\}, 16)$  SC-LDPC codes with  $\pi_2$  when  $P_{\text{subs}} = 0.01$ .



**Fig. 9** Symmetric information rates versus  $P_{\text{subs}}$  when  $P_{\text{ins}} + P_{\text{del}} = 0.1$ .

For comparing the decoding performance for the MHID channels between SC-LDPC codes and LDPC-BLK codes, we searched degree distributions of irregular LDPC codes that have higher thresholds. The search was performed in a similar way to [20]. For any number of RH  $a \in \{1, 2, 4\}$ , as far as we searched, degree distributions that give higher

**Table 2** Thresholds of LDPC-BLK codes and SC-LDPC codes with various numbers of RHs  $a = 1, 2, 4$ .

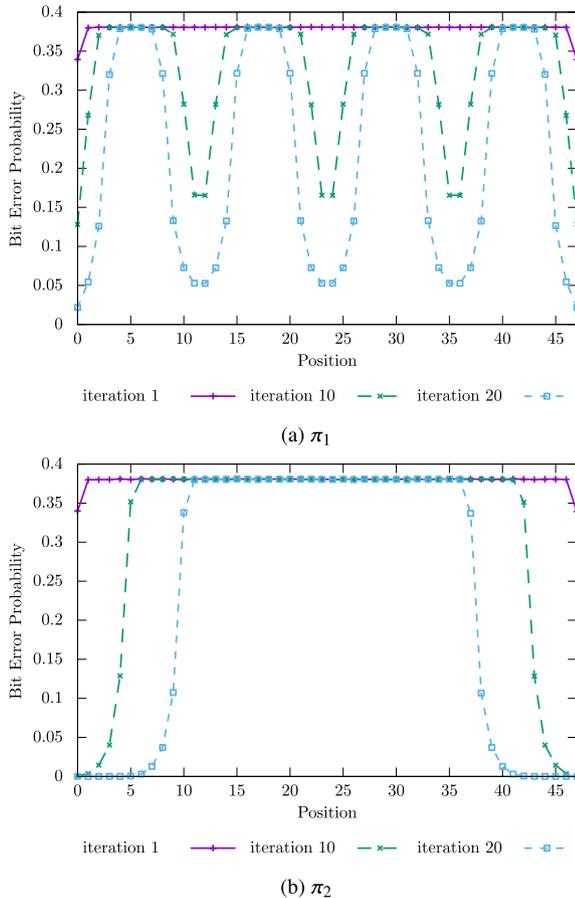
| Ensemble          | Threshold ( $P_{\text{ins}} = P_{\text{del}}$ ) |         |         |                          |         |         |
|-------------------|---|---------|---------|--------------------------|---------|---------|
|                   | $P_{\text{subs}} = 0$                           |         |         | $P_{\text{subs}} = 0.01$ |         |         |
|                   | $a = 1$   | $a = 2$ | $a = 4$ | $a = 1$                  | $a = 2$ | $a = 4$ |
| (3,6)-regular     | 0.0006  | 0.0011  | 0.0022  | 0.0005                   | 0.0010  | 0.0018  |
| MHID irregular    | 0.0021  | 0.0041  | 0.0094  | 0.0020                   | 0.0036  | 0.0069  |
| $(3,6,16), \pi_1$ | 0.147   | 0.228   | 0.342   | 0.121                    | 0.196   | 0.292   |
| $(3,6,16), \pi_2$ | 0.147   | 0.314   | 0.587   | 0.121                    | 0.252   | 0.490   |
| $(3,9,16), \pi_1$ | 0.068   | 0.093   | 0.129   | 0.049                    | 0.075   | 0.101   |
| $(3,9,16), \pi_2$ | 0.068   | 0.142   | 0.302   | 0.049                    | 0.101   | 0.214   |

thresholds could only be found by increasing the fraction of degree-2 variable nodes. For example, the degree distributions,  $\lambda(x) = 0.642566x + 0.357143x^{29}$ ,  $\rho(x) = x^5$  (“MHID irregular” in Table 2), were obtained under the constraints that the code rate of 0.5, the maximum variable node degree of 30, and the check node degree only of 6. Such a tendency of the search for good degree distributions can also be seen in the search for better degree distributions of LDPC codes concatenated with markers for insertion and deletion channels [8]. Unfortunately, the high fraction of degree-2 variable node not only results in a reduced minimum distance for the ensemble but also is not expected to produce a good code concerning short cycles. Furthermore, the irregular LDPC codes optimized for AWGN channels [5] are not suitable for the MHID channels under the joint iterative decoding. For example, the listed code in [5] with the code rate of 0.5 and the maximum variable node degree of 30 has the threshold  $P_{\text{ins}} + P_{\text{del}} < 0.0001$  even when  $a = 4$  and  $P_{\text{subs}} = 0$ . Meanwhile, as you can see from Table 2, the SC-LDPC codes remarkably outperform the LDPC-BLK codes.

One reason is that the  $(d_v, d_c, L)$  SC-LDPC codes are known to have superior performance compared to LDPC-BLK codes. Another reason depends on the structured irregularity of the  $(d_v, d_c, L)$  SC-LDPC codes. The messages  $u$  from  $v_{\text{low}}$  have reliable information and assist in estimating the state transitions they engage in. On the other hand, the messages  $\eta$  updated by the messages  $u$  from  $v_{\text{low}}$  assist in the sum-product decoding. Namely, the FB algorithm and the sum-product decoding enhance each other with the low degree check nodes, which improves the thresholds of the  $(d_v, d_c, L)$  SC-LDPC codes.

We observe that the thresholds differ depending on the order of the SC-LDPC codewords in RM, i.e., permutations  $\pi_1, \pi_2$ , and the thresholds with  $\pi_2$  is much higher than those with  $\pi_1$ . When using SC-LDPC codes, considering the order of the codeword bits in RM is effective in improving the decoding performance, while it does not apply to unstructured LDPC codes defined by degree distribution pair.

Figures 7 and 8 compare the thresholds with the SIRs, for different  $P_{\text{subs}}$ , where the thresholds of the  $(3, d_c = \{6, 9, 12\}, 16)$  SC-LDPC codes with  $a = 1, 2, 4$  and  $\pi_2$  are plotted. We observe that there is a gap between the threshold and SIR, and the gap increases as increasing code rate. However, the SC-LDPC codes are much closer to the SIRs



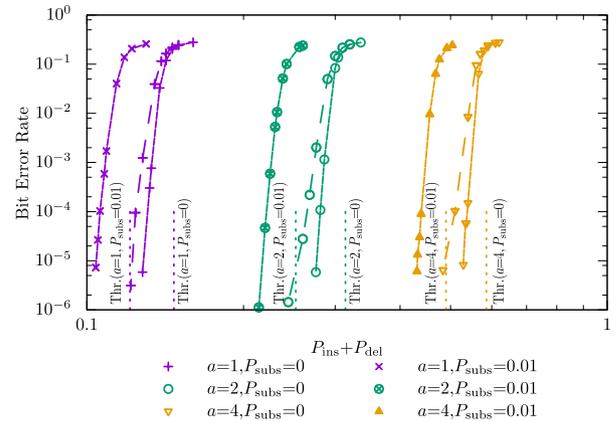
**Fig. 10** Average bit error probability of variable nodes at each position for (3, 6, 48) SC-LDPC codes, where  $a = 4$ ,  $P_{\text{ins}} + P_{\text{del}} = 0.3$ , and  $P_{\text{subs}} = 0$ .

than the LDPC-BLK codes listed in Table 2 even if higher code rate. It is an open problem whether SC-LDPC codes may approach the SIRs of MHID channels.

In Fig. 10, we track the average bit error probability of variable nodes from the (3, 6, 48) SC-LDPC codes at each position for some iterations. When using the ordering by  $\pi_1$ , the joint iterative decoding proceeds as can be seen in Fig. 10(a) and thus would stop halfway through as described in 3.4. On the other hand, as shown in Fig. 10(b), the ordering by  $\pi_2$  can utilize the structured irregularity and thus make the joint iterative decoding proceed in a wave-like fashion from the ends toward the center.

### 5.3 Simulation Result

Figure 11 shows the bit error rate for the (3, 6, 16) SC-LDPC codes. We observe that the bit error rates decrease with their thresholds as turning points. Therefore, the thresholds obtained by the DE can well predict the performance of the joint decoding performance. Moreover, the performance improvement by increasing the number of RHs can be verified for various code lengths.



**Fig. 11** Bit error rate curves for (3, 6, 16) SC-LDPC codes of  $N = 65536$  (solid curves) and  $N = 32768$  (dashed curves) with  $\tau_2$ .

## 6. Conclusion

In this paper, we proposed an SC-LDPC code coding scheme for MHID channels. First, we characterized the channel model, which is the counterpart of shift operations in an RM with multiple RHs. We proposed a joint iterative decoding algorithm for MHID channels. This is a message-passing algorithm for the factor graph, where two subgraphs of MHID channels and SC-LDPC codes are joined together. The DE analysis and simulation results showed a performance improvement, even though each codeword bit was read only once regardless of the number of RHs. Furthermore, we observed that the SC-LDPC codes had superior performance for MHID channels compared to the LDPC-BLK codes. Considering the order of the codeword bits in an RM, we could utilize the structured irregularity of SC-LDPC codes and obtained a significant improvement in the decoding performance.

## Acknowledgments

This work was supported by JSPS KAKENHI Grant Numbers JP16K18108, JP17K06443. The authors would like to thank the reviewers for their constructive comments that have helped improving the overall quality of the paper.

## References

- [1] S. Mittal, "A survey of techniques for architecting processor components using domain wall memory," *ACM J. Emerg. Technol. Comput. Syst.*, vol.13, no.2, pp.1–29, Nov. 2016.
- [2] C. Zhang, G. Sun, X. Zhang, W. Zhang, W. Zhao, T. Wang, Y. Liang, Y. Liu, Y. Wang, and J. Shu, "Hi-fi playback: Tolerating position errors in shift operations of racetrack memory," *Proc. ACM/IEEE Int. Symp. Computer Architecture*, pp.694–706, June 2015.
- [3] Y.M. Chee, H.M. Kiah, A. Vardy, V.K. Vu, and E. Yaakobi, "Coding for racetrack memories," *Proc. IEEE Int. Symp. Inf. Theory*, pp.619–623, July 2017.
- [4] R.G. Gallager, "Low-density parity-check codes," *IRE Trans. Inf. Theory*, vol.8, pp.21–28, Jan. 1962.
- [5] T.J. Richardson, M.A. Shokrollahi, and R.L. Urbanke, "Design

- of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol.47, no.2, pp.619–637, Feb. 2001.
- [6] M.C. Davey and D.C. MacKay, "Reliable communication over channel with insertions, deletions and substitutions," *IEEE Trans. Inf. Theory*, vol.47, no.2, pp.687–698, Feb. 2001.
- [7] F. Wang, D. Fertonani, and T.M. Duman, "Symbol-level synchronization and LDPC code design for insertion/deletion channel," *IEEE Trans. Commun.*, vol.59, no.5, pp.1287–1297, May 2011.
- [8] E.A. Ratzler, "Marker codes for channel with insertions and deletions," *Ann. Telecommun.*, vol.60, no.1-2, pp.29–44, Jan.–Feb. 2005.
- [9] R. Goto, K. Kasai, and H. Kaneko, "Coding of insertion-deletion-substitution channels without markers," *Proc. IEEE Int. Symp. Inf. Theory*, pp.635–639, July 2016.
- [10] M. Lentmaier, A. Sridharan, D. Costello, and K. Zigangirov, "Iterative decoding threshold analysis for LDPC convolutional codes," *IEEE Trans. Inf. Theory*, vol.56, no.10, pp.5274–5289, Oct. 2010.
- [11] 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, Jan. 2011.
- [12] J. Garcia-Frias, "Decoding of low-density parity-check codes over finite-state binary Markov channel," *IEEE Trans. Commun.*, vol.52, no.11, pp.1840–1843, Nov. 2004.
- [13] F.R. Kschischang, B.J. Frey, and H.A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol.47, no.2, pp.498–519, Feb. 2001.
- [14] R. Shibata, G. Hosoya, and H. Yashima, "Joint state estimation and decoding of low-density parity-check codes for racetrack memories," *IEICE Technical Report*, IT2017–45, Sept. 2017 (in Japanese).
- [15] R. Shibata, G. Hosoya, and H. Yashima, "Performance analysis of spatially-coupled low-density parity-check codes for racetrack memories," *Proc. SITA2017*, pp.529–534, Nov. 2017 (in Japanese).
- [16] J. Thorpe, "Low-density parity-check (LDPC) codes constructed from protographs," *IPN Progress Report* 42-154, Aug. 2003.
- [17] D.M. Arnold, H.A. Loeliger, P.O. Vontobel, A. Kavčić, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Trans. Inf. Theory*, vol.52, no.8, pp.3490–3508, Aug. 2006.
- [18] S.-Y. Chung, G.D. Forney, Jr., T.J. Richardson, and R.L. Urbanke, "On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit," *IEEE Commun. Lett.*, vol.5, no.2, pp.58–60, Feb. 2001.
- [19] A. Kavčić, X. Ma, and M. Mitzenmacher, "Binary intersymbol interference channels: Gallager codes density evolution and code performance bounds," *IEEE Trans. Inf. Theory*, vol.49, no.7, pp.1636–1652, July 2003.
- [20] N. Varnica and A. Kavčić, "Optimized low-density parity-check codes for partial response channels," *IEEE Commun. Lett.*, vol.7, no.4, pp.168–170, April 2003.



**Gou Hosoya** received the B.E. degree, M.E. degree, and D.E. degree in Industrial and Management Systems Engineering from Waseda University, Tokyo, Japan, in 2002, 2004, and 2008, respectively. He is currently a junior associate professor at the Department of Information and Computer Technology, Faculty of Engineering at Tokyo University of Science, Tokyo, Japan. His research interests are coding theory and information theory. He is a member of the IEEE.



**Hiroyuki Yashima** received the B.E., M.E., and Ph.D. degrees from Keio University, Japan, in 1981, 1987, 1990, respectively, all in electrical engineering. From 1990 to 2003, he was an associate professor at the Department of Information and Computer Sciences, Saitama University, Japan. From 1994 to 1995, he was a visiting scholar at the University of Victoria, Canada. In 2003, he joined Tokyo University of Science, Japan, as a professor at the Department of Management Science. Currently, he is a professor at the Department of Information and Computer Technology. His research interests include modulation and coding, optical communication systems, satellite communication systems, and spread spectrum communication systems. He is a member of the IEEE.



**Ryo Shibata** received the B.E. degree and M.E. degree from Tokyo University of Science in 2016 and 2018, respectively. He is currently pursuing the D.E. degree in Tokyo University of Science. His research interest includes coding theory.