PAPER *Special Section on Information Theory and Its Applications*

Density Evolution Analysis of Robustness for LDPC Codes over the Gilbert-Elliott Channel

Manabu KOBAYASHI[†a)], **Hideki YAGI**[††], **Toshiyasu MATSUSHIMA**[†††], ***Members***, ***and*** **Shigeichi HIRASAWA**[††††], ***Fellow***

SUMMARY In this paper, we analyze the robustness for low-density parity-check (LDPC) codes over the Gilbert-Elliott (GE) channel. For this purpose we propose a density evolution method for the case where LDPC decoder uses the mismatched parameters for the GE channel. Using this method, we derive the region of tuples of true parameters and mismatched decoding parameters for the GE channel, where the decoding error probability approaches asymptotically to zero.
key words: *LDPC codes, density evolution, Gilbert-Elliott channel, robustness*

1. Introduction

Low-density parity-check (LDPC) codes are a class of linear codes with very sparse parity-check matrices [1]. To analyze the performance of LDPC codes, the density-evolution (DE) algorithm has been proposed by Richardson et al. in [2]. This method can find the convergence behavior of LDPC codes under message-passing decoding, assuming very large code length. Futhermore irregular LDPC codes exhibit the performance extremely close to the Shannon limit for memoryless channels [3], [4].

For the Gilbert-Elliott (GE) channel or more general Markov channels, several message-passing decoding algorithms of LDPC codes have been proposed [5]–[8]. These algorithms result in significantly improved performance for channels with memory. Furthermore, Eckford et al. have analyzed performance of LDPC codes over the GE channel using DE, under an estimation-decoding strategy, in which intermadiate results from the iterative decoding algorithm are used to estimate a channel state [8]. This DE analysis provides performance thresholds which the decoder converges to zero error probability over the GE channel. Recently, Eckford et al. proposed a design method of irregular LDPC codes using approximate DE for Markov channels [9]. The best of LDPC codes designed in [9] achieves roughly 95% of the capacity of the corresponding GE channel.

These conventional DE algorithms assume that the receiver knows true parameters of the probability density function (PDF) or the probability mass function (PMF) of the channel noise. However, it is difficult for the receiver to know true channel parameters in general. Therefore the message-passing decoder usually uses estimated parameters of the channel noise. When estimated channel parameters differ from the true ones, how much does the estimation error of the channel parameters give the influence to the performance threshold? If we can know this influence, it may be possible to design an estimation method of channel parameters and the LDPC decoder.

For turbo codes over the additive white Gaussian noise channel, Summers et al. have studied the sensitivity of decoder performance to misestimation of the SNR by computer simulations [10]. Moreover, for periodic scalar fading channels, Jones et al. have shown the robustness of LDPC codes by demonstrating the efect of misestimation of the channel parameters [11]. In previous works, however, only memoryless channels have been studied.

In this paper, we propose a DE algorithm for the case where the LDPC decoder uses estimated channel parameters over the GE channel. This is an extension of the DE algorithm proposed by Eckford et al. [8]. Furthermore, using the proposed DE alogrithm, we show some regions of tuples of true and estimated channel parameters over the GE channel, where the decoding error probability approaches asymptotically to zero. Consequently, we show that the message-passing decoder in [8] has the robustness for LDPC codes over the GE channel.

2. System Models and GE-LDPC Decoder [8]

In this section, we describe the channel and decoder models (which are necessary for further discussion).

Throughout this paper, random variables will be denoted with upper case letters, and specific values from the corresponding sample space with corresponding lower case letters. We will use $p_X(x)$ to represent the PMF of a discrete random variable X. Likewise, the PDF of a continuous random variable X will be denoted as $f_X(x)$. When no confusion can arise, we will let the argument of the PMF or PDF identify the corresponding random variable, simply writing $p(x)$ for $p_X(x)$, $f(x)$ for $f_X(x)$ and so on.

Manuscript received January 25, 2008.
Manuscript revised April 23, 2008.
[†]The author is with the Department of Information Science, School of Engineering, Shonan Institute of Technology, Fujisawa-shi, 251-8511 Japan.
[††]The author is with Media Network Center, Waseda University, Tokyo, 169-8050 Japan.
[†††]The author is with the School of Fundamental Science and Engineering, Waseda University, Tokyo, 169-8555 Japan.
[††††]The author is with the School of Creative Science and Engineering, Waseda University, Tokyo, 169-8555 Japan.
a) E-mail: kobayasi@info.shonan-it.ac.jp
DOI: 10.1093/ietfec/e91–a.10.2754

Definition 1: Let C be a binary (d_v, d_c)-regular LDPC code[†] of length n with an $m \times n$ parity-check matrix H. The kth row of H is denoted by $\boldsymbol{h}^k$. Let $h_k : \{0,1\}^n \to \{0,1\}$ be the indicator function for $\boldsymbol{h}^k$ as follows:

$$h_k(\boldsymbol{x}) = \begin{cases} 1, & \boldsymbol{h}^k \boldsymbol{x}^{\mathrm{T}} = 0, \\ 0, & \text{otherwise}, \end{cases} \tag{1}$$

where $\boldsymbol{x}^{\mathrm{T}}$ denotes the transpose of $\boldsymbol{x}$. $h(\boldsymbol{x})$ is defined as $h(\boldsymbol{x}) = \prod_{k=1}^{m} h_k(\boldsymbol{x})$. □

From this definition we have $h(\boldsymbol{x}) = 1$ if $\boldsymbol{x} \in C$, and $h(\boldsymbol{x}) = 0$ otherwise. Assuming that each codeword of C is equally likely to be selected by the transmitter, we have $p(\boldsymbol{x}) = h(\boldsymbol{x})/|C|$, where $|C|$ is the number of codewords of C.

Definition 2: A codeword $\boldsymbol{X} \in C$ is transmitted to the channel and the receiver receives a channel output $\boldsymbol{Y} \in \{0,1\}^n$ such that

$$\boldsymbol{Y} = \boldsymbol{X} + \boldsymbol{Z}, \tag{2}$$

where $\boldsymbol{Z} \in \{0,1\}^n$ is a noise sequence and $+$ denotes the componentwise modulo-2 addition. The noise sequence $\boldsymbol{Z}$ arises from a two-states hidden Markov process at the GE channel. Given the state space $\mathcal{S} = \{G, B\}$ and a state sequence $\boldsymbol{S} \in \mathcal{S}^n$, η_s denotes the crossover probability at the state $s \in \mathcal{S}$, i.e. $\eta_G = \Pr(Z_i = 1|S_i = G)$ and $\eta_B = \Pr(Z_i = 1|S_i = B)$. The state transition probabilities are denoted by $g = \Pr(S_{i+1} = G|S_i = B)$ and $b = \Pr(S_{i+1} = B|S_i = G)$. Let P be the state transition matrix, i.e.,

$$P = \begin{bmatrix} 1-b & b \\ g & 1-g \end{bmatrix}. \tag{3}$$

□

Then the PMF of the state sequence $\boldsymbol{S} \in \mathcal{S}^n$ can be factored as $p(\boldsymbol{s}) = p(s_1)\prod_{j=1}^{n-1} p(s_{j+1}|s_j)$. The joint PMF of a channel output $\boldsymbol{Y}$, a codeword $\boldsymbol{X}$ and a state sequence $\boldsymbol{S}$ is given by

$$\begin{aligned} p(\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{s}) &= p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{s})p(\boldsymbol{s})p(\boldsymbol{x}) \\ &= \frac{1}{|C|}\left(\prod_{i=1}^{n} p(y_i|x_i, s_i)\right) \\ &\quad \cdot \left(p(s_1)\prod_{j=1}^{n-1} p(s_{j+1}|s_j)\right)\left(\prod_{k=1}^{m} h_k(\boldsymbol{x})\right). \end{aligned} \tag{4}$$

This probabilistic model results in a factor graph formed by connecting the LDPC factor graph and the GE Markov chain factor graph so that edges are created to connect the appropriate symbol-variable nodes and channel factor nodes. For this factor graph, the message-passing decoder utilizing sum-product algorithm (SPA) can effectively decode the GE channel noise [6], [8]. This decoder will be referred to as the GE-LDPC decoder [8]. The GE-LDPC decoder is depicted graphically in Fig. 1.

Definition 3: For $d \in \mathcal{R}$ and $y \in \{0,1\}$, let $\gamma(d, y)$ be defined as

$$\gamma(d, y) = \frac{1}{2}\left[1 + \phi(y)\tanh\left(\frac{d}{2}\right)\right], \tag{5}$$

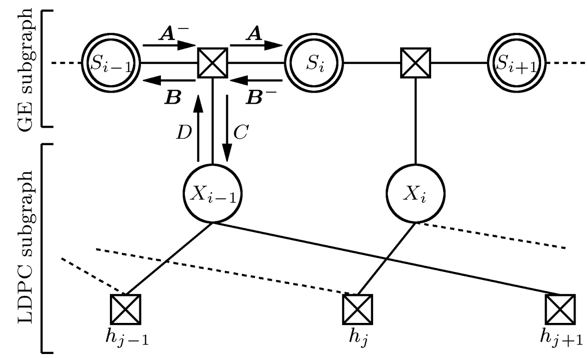

Fig. 1 A GE-LDPC decoder graph and the massage flow through the Markov subgraph.

where $\phi(x) = (-1)^x, x \in \{0,1\}$. Furthermore, let $N = \mathrm{diag}[\eta_G, \eta_B]$ and

$$E(d, y) = N(1 - \gamma(d, y)) + (I - N)\gamma(d, y), \tag{6}$$

where I is an identity matrix. □

In Fig. 1, the GE-LDPC decoder calculates the current forward message vector $\boldsymbol{A} = (A_1, A_2)^{\mathrm{T}}$ by using the previous forward message vector $\boldsymbol{A}^- = (A_1^-, A_2^-)^{\mathrm{T}}$, the channel output Y and the extrinsic information D which is the output of SPA. More precisely, $\boldsymbol{A}$ is calculated as

$$\boldsymbol{A} = \frac{P^{\mathrm{T}}E(D, Y)\boldsymbol{A}^-}{\boldsymbol{u}^{\mathrm{T}}P^{\mathrm{T}}E(D, Y)\boldsymbol{A}^-}, \tag{7}$$

where, by introducing a sequence $\boldsymbol{u}$, the denominator is a normalization constant to satisfy that $A_1 + A_2 = 1$. The forward message vector $\boldsymbol{A}$ indicates the marginal state probability estimated by using the extrinsic informations from the start to the current time, and A_1 and A_2 imply the estimated probabilities corresponding to the state G and B at the current time, conditioned by some received symbols, respectively. Similarly, the current backward message vector $\boldsymbol{B} = (B_1, B_2)^{\mathrm{T}}$ is calculated by using the backward message vector $\boldsymbol{B}^- = (B_1^-, B_2^-)^{\mathrm{T}}$ at the next time, the channel output Y and the extrinsic information D from the SPA as follows:

$$\boldsymbol{B} = \frac{E(D, Y)P^{\mathrm{T}}\boldsymbol{B}^-}{\boldsymbol{v}^{\mathrm{T}}E(D, Y)P^{\mathrm{T}}\boldsymbol{B}^-}, \tag{8}$$

where, by introducing a sequence $\boldsymbol{v}$, the denominator is a normalization constant to satisfy that $B_1 + B_2 = 1$. The backward message vector $\boldsymbol{B}$ indicates the marginal state probability estimated by using the extrinsic informations from the current time to the end. Using these forward and backward message vectors, the channel message C, which is the input to the SPA, is calculated as

$$C = \phi(Y)\log\frac{(\boldsymbol{A}^-)^{\mathrm{T}}(I - N)P\boldsymbol{B}^-}{(\boldsymbol{A}^-)^{\mathrm{T}}NP\boldsymbol{B}^-}. \tag{9}$$

The extrinsic information D is calculated by the SPA in

[†]Although we assume regular LDPC codes for simplicity, the extension to irregular LDPC codes is straightforward.

the same manner as messages from a symbol-variable node to a parity-check node.

In this way, GE-LDPC decoder iterates the following steps: (i) calculate the forward and the backward messages using the channel output Y and the extrinsic information D from the SPA, (ii) calculate the channel message C according to (9), (iii) input the channel message C to the SPA and (iv) calculate the extrinsic information D by using the SPA. Eckford et al. have shown that GE-LDPC decoder is asymptotically optimal as the code length $n \to \infty$ since the subgraphs of the GE-LDPC factor graph are cycle-free for some fixed depth with $Pr \to 1$ as the code length $n \to \infty$ [8]. Furthermore, Eckford et al. have proposed a DE algorithm for the GE-LDPC decoder with true channel parameters over the GE channel [8].

3. DE Algorithm for GE-LDPC Decoder Using Estimated Channel Parameters

Conventional DE algorithms assume that the receiver knows true parameters of the channel noise. However, it is difficult for the receiver to know true channel parameters in general. Therefore the message-passing decoder usually uses estimated parameters of the channel noise. In this section, we propose a DE algorithm over the GE channel for the case where the GE-LDPC decoder uses estimated channel parameters which differ from true ones.

3.1 Preliminary

Definition 4: Let $\theta = (g, b, \eta_G, \eta_B)$ be a channel parameter vector for the GE channel and let the parameter space Θ be the set of all channel parameter vectors. And let $\tilde{\theta} = (\tilde{g}, \tilde{b}, \tilde{\eta}_G, \tilde{\eta}_B)$ and $\hat{\theta} = (\hat{g}, \hat{b}, \hat{\eta}_G, \hat{\eta}_B), \tilde{\theta}, \hat{\theta} \in \Theta$, denote the true and the estimated channel parameter vector, respectively. The PMFs corresponding to $\tilde{\theta}$ and $\hat{\theta}$ are denoted by $\tilde{p}$ and $\hat{p}$, respectively (e.g. $\tilde{p}(Z = 1|S = s) = \tilde{\eta}_s, \hat{p}(Z = 1|S = s) = \hat{\eta}_s, s \in \mathcal{S}, \tilde{p}(S = G) = \tilde{g}/(\tilde{g} + \tilde{b}), \hat{p}(S = G) = \hat{g}/(\hat{g} + \hat{b})$ and so on). In an analogous way, random variables corresponding to $\tilde{\theta}$ and $\hat{\theta}$ are denoted by $\tilde{A}$ and $\hat{A}$, respectively. □

Two conditions are required for the efficient application of DE analysis: (i) the independence assumption and (ii) the symmetry condition. As described above, the subgraphs of the GE-LDPC factor graph are cycle-free for some fixed depth with $Pr \to 1$ as the code length $n \to \infty$ [8]. Since the subgraphs are independent of the channel parameter vector, the independence assumption is fulfilled. The proof of the symmetry condition is given in the Appendix. From the symmetry condition, we can assume that the all-zero codeword is transmitted. Note that received vector $\boldsymbol{Y}$ is equal to the noise vector $\boldsymbol{Z}$ from (2) in this case.

Next, an overview of this section is presented here. We assume that the GE-LDPC decoder uses an estimated channel parameter vector $\hat{\theta}$. Then GE-LDPC decoder updates the massages $\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \hat{C}$ and D as described in the previous section. From $\hat{A}_2 = 1 - \hat{A}_1$ and $\hat{B}_2 = 1 - \hat{B}_1$, it is sufficient to consider the messages $\hat{A}_1$ and $\hat{B}_1$. Let $f(\hat{A}_1)$, $f(\hat{B}_1)$, $f(\hat{C})$ and $f(D)$ be the PDFs of $\hat{A}_1$, $\hat{B}_1$, $\hat{C}$ and D, respectively. In the DE analysis, we want to update these PDFs from the previous PDFs along the message passing of the decoder as follows: (i) calculate $f(\hat{A}_1)$ from $f(\hat{A}_1^-)$ and $f(D)$, and likewise calculate $f(\hat{B}_1)$ from $f(\hat{B}_1^-)$ and $f(D)$, (ii) calculate $f(\hat{C})$ from $f(\hat{A}_1)$ and $f(\hat{B}_1)$, and (iii) calculate $f(D)$ from $f(\hat{C})$. Unfortunately, steps (i) and (ii) cannot be calculated since $\hat{\theta}$ differs from the true parameter vector $\tilde{\theta}$. The detail will be described in the following subsection. Therefore, we give up updating $f(\hat{A}_1)$ and $f(\hat{B}_1)$, and we introduce the joint PDFs $f(\tilde{A}_1, \hat{A}_1)$ and $f(\tilde{B}_1, \hat{B}_1)$. Then the current PDF $f(\tilde{A}_1, \hat{A}_1)$ can be calculated from the previous PDF $f(\tilde{A}_1^-, \hat{A}_1^-)$ and $f(D)$. Likewise $f(\tilde{B}_1, \hat{B}_1)$ can be calculated. As a result, we can obtain $f(\hat{C})$ by utilizing $f(\tilde{A}_1, \hat{A}_1)$ and $f(\tilde{B}_1, \hat{B}_1)$. Thus, the outline of the DE is as follows: (i) calculate $f(\tilde{A}_1, \hat{A}_1)$ from $f(\tilde{A}_1^-, \hat{A}_1^-)$ and $f(D)$, and likewise calculate $f(\tilde{B}_1, \hat{B}_1)$ from $f(\tilde{B}_1^-, \hat{B}_1^-)$ and $f(D)$, (ii) calculate $f(\hat{C})$ from $f(\tilde{A}_1, \hat{A}_1)$ and $f(\tilde{B}_1, \hat{B}_1)$, and (iii) calculate $f(D)$ from $f(\hat{C})$.

3.2 PDFs of Forward and Backward Messages

Given a channel parameter vector $\theta = (g, b, \eta_G, \eta_B) \in \Theta$ and the previous foward message A_1^-, which is the element of $\boldsymbol{A}^- = (A_1^-, A_2^-)^{\mathrm{T}}$, the current foward message A_1, which is the element of $\boldsymbol{A} = (A_1, A_2)^{\mathrm{T}}$, can be calculated as the following function of θ and A_1^- from (7) (see [8]):

$$A_1 = \mathcal{A}_1(\theta, A_1^-) = \frac{N_1^A(\theta) + N_2^A(\theta)A_1^-}{D_1^A(\theta) + D_2^A(\theta)A_1^-}, \tag{10}$$

where

$$\begin{aligned}
N_1^A(\theta) &= g(\eta_B + (1 - 2\eta_B)\gamma(D, Y)), \\
N_2^A(\theta) &= (1 - b)(\eta_G + (1 - 2\eta_G)\gamma(D, Y)) \\
&\quad -g(\eta_B + (1 - 2\eta_B)\gamma(D, Y)), \\
D_1^A(\theta) &= \eta_B + (1 - 2\eta_B)\gamma(D, Y), \\
D_2^A(\theta) &= (\eta_G + (1 - 2\eta_G)\gamma(D, Y)) \\
&\quad -\eta_B - (1 - 2\eta_B)\gamma(D, Y).
\end{aligned} \tag{11}$$

Although $\mathcal{A}_1(\theta, A_1^-)$ depends on D and Y, we omit them for simplicity in this paper.

Let $\mathcal{A}_1^-(\theta, A_1)$ denote the inverse function for $\mathcal{A}_1(\theta, A_1^-)$, and from (10) we have

$$A_1^- = \mathcal{A}_1^-(\theta, A_1) = \frac{D_1^A(\theta)A_1 - N_1^A(\theta)}{N_2^A(\theta) - D_2^A(\theta)A_1}. \tag{12}$$

Then the partial derivative $\mathcal{A}_1^{-\prime}(\theta, A_1) = \frac{\partial \mathcal{A}_1^-(\theta, A_1)}{\partial A_1}$ is given by

$$\mathcal{A}_1^{-\prime}(\theta, A_1) = \frac{D_1^A(\theta)N_2^A(\theta) - D_2^A(\theta)N_1^A(\theta)}{(N_2^A(\theta) - D_2^A(\theta)A_1)^2}. \tag{13}$$

Since we assume that the decoder uses an estimated channel parameter vector $\hat{\theta}$, we consider to calculate the PDF of $\hat{A}_1$, denoted by $f(\hat{A}_1)$. To obtain the current

PDF $f(\hat{A}_1)$ from the previous PDF $f(\hat{A}_1^-)$, we assume that $p(S|\hat{A}_1^-)$ can be calculated for $S \in \mathcal{S}$. Then we can obtain $f(\hat{A}_1^-|S)$ as follows:

$$f(\hat{a}_1^-|s) = \frac{p(s|\hat{a}_1^-)f(\hat{a}_1^-)}{p(s)}. \tag{14}$$

Noting that $\hat{A}_1^-$ is independent of D and Y given S, we have $f(\hat{A}_1^-|S) = f(\hat{A}_1^-|S, Y, D)$. Since $\hat{A}_1^- = \mathcal{A}_1^-(\hat{\theta}, \hat{A}_1)$ is determined by $\hat{A}_1$ given S, Y and D, $f(\hat{A}_1|S, Y, D)$ can be expressed by $f(\hat{A}_1^-|S) = f(\hat{A}_1^-|S, Y, D)$. Using the transformation of variables, we have

$$f(\hat{a}_1|s, y, d) = f(\mathcal{A}_1^-(\hat{\theta}, \hat{a}_1)|s)|\mathcal{A}_1^{-\prime}(\hat{\theta}, \hat{a}_1)|. \tag{15}$$

Since the PMFs of Y and S are determined by true channel parameters, marginalizing over true channel parameters and the PDF $f(D)$ of D from the SPA, $f(\hat{A}_1)$ is expressed as

$$f(\hat{a}_1) = \sum_{s\in\mathcal{S}} \sum_{y\in\{0,1\}} \tilde{p}(s)\tilde{p}(y|s) \int_d f(\hat{a}_1|s, y, d)f(d)\mathrm{d}d. \tag{16}$$

Therefore, if we can obtain $p(S|\hat{A}_1^-)$, $f(\hat{A}_1)$ can be calculated from $f(\hat{A}_1^-)$.

If we assume that $\hat{\theta} = \tilde{\theta}$, we have $p(S = G|\hat{a}_1^-) = \hat{a}_1^-$ and $p(S = B|\hat{a}_1^-) = 1 - \hat{a}_1^-$ because $\hat{a}_1^- = \tilde{a}_1^-$, and $\tilde{a}_1^-$ implies the probability at the state G. For the case where $\hat{\theta} \neq \tilde{\theta}$, however, the conditional probability satisfies $p(S = G|\hat{a}_1^-) = \tilde{a}_1^-$, and $\tilde{a}_1^-$ is not determined by only $\hat{a}_1^-$. Since $p(S|\hat{A}_1^-)$ cannot be calculated by using only $\hat{A}_1^-$, we cannot calculate (14), (15) and (16). Therefore, we give up updating $f(\hat{A}_1)$.

Here we introduce the joint PDF $f(\tilde{A}_1^-, \hat{A}_1^-)$ of $\tilde{A}_1^-$ and $\hat{A}_1^-$. Given this joint PDF, we will describe a method to obtain the next joint PDF $f(\tilde{A}_1, \hat{A}_1)$. First, from the above argument $p(S|\tilde{A}_1^-, \hat{A}_1^-)$ is given by

$$p(s|\tilde{a}_1^-, \hat{a}_1^-) = \begin{cases} \tilde{a}_1^-, & s = G, \\ 1 - \tilde{a}_1^-, & s = B. \end{cases} \tag{17}$$

Then using Bayes' rule, we have

$$f(\tilde{a}_1^-, \hat{a}_1^-|s) = p(s|\tilde{a}_1^-, \hat{a}_1^-)f(\tilde{a}_1^-, \hat{a}_1^-)/p(s). \tag{18}$$

Next, we consider to obtain the conditional joint PDF $f(\tilde{A}_1, \hat{A}_1|S, Y, D)$. Given S, Y and D, $\tilde{A}_1^-$ and $\hat{A}_1^-$ are determined by $\tilde{A}_1$ and $\hat{A}_1$ from (12), respectively, as follows:

$$\tilde{A}_1^- = \mathcal{A}_1^-(\tilde{\theta}, \tilde{A}_1), \qquad \hat{A}_1^- = \mathcal{A}_1^-(\hat{\theta}, \hat{A}_1). \tag{19}$$

Since $\tilde{A}_1^-$ and $\hat{A}_1^-$ are conditionally independent of Y and D given $S \in \mathcal{S}$, note that $f(\tilde{A}_1^-, \hat{A}_1^-|S, Y, D) = f(\tilde{A}_1^-, \hat{A}_1^-|S)$. Therefore, using the transformation of variables for (19), $f(\tilde{A}_1, \hat{A}_1|S, Y, D)$ can be expressed by $f(\tilde{A}_1^-, \hat{A}_1^-|S)$. From (19) the Jacobian matrix J is given by

$$J = \begin{bmatrix} \partial\tilde{A}_1^-/\partial\tilde{A}_1 & \partial\tilde{A}_1^-/\partial\hat{A}_1 \\ \partial\hat{A}_1^-/\partial\tilde{A}_1 & \partial\hat{A}_1^-/\partial\hat{A}_1 \end{bmatrix}, \tag{20}$$

where $\partial\hat{A}_1^-/\partial\tilde{A}_1 = \partial\tilde{A}_1^-/\partial\hat{A}_1 = 0$. Jacobian $|J|$ is given by

$$|J| = \mathcal{A}_1^{-\prime}(\tilde{\theta}, \tilde{A}_1)\mathcal{A}_1^{-\prime}(\hat{\theta}, \hat{A}_1). \tag{21}$$

Therefore we obtain $f(\tilde{A}_1, \hat{A}_1|S, Y, D)$ as follows:

$$\begin{aligned} f(\tilde{a}_1, \hat{a}_1|s, y, d) &= f(\mathcal{A}_1^-(\tilde{\theta}, \tilde{a}_1), \mathcal{A}_1^-(\hat{\theta}, \hat{a}_1)|s) \\ &\quad \cdot|\mathcal{A}_1^{-\prime}(\tilde{\theta}, \tilde{a}_1)\mathcal{A}_1^{-\prime}(\hat{\theta}, \hat{a}_1)|. \end{aligned} \tag{22}$$

At last, marginalizing over true channel parameters and the PDF $f(D)$ of D from the SPA, $f(\tilde{A}_1, \hat{A}_1)$ is expressed as

$$\begin{aligned} f(\tilde{a}_1, \hat{a}_1) &= \sum_{s\in\mathcal{S}} \sum_{y\in\{0,1\}} \tilde{p}(s)\tilde{p}(y|s) \\ &\quad \cdot \int_d f(\tilde{a}_1, \hat{a}_1|s, y, d)f(d)\mathrm{d}d. \end{aligned} \tag{23}$$

In the sequel, we can update the current PDF $f(\tilde{A}_1, \hat{A}_1)$ from the previous PDF $f(\tilde{A}_1^-, \hat{A}_1^-)$ and the input PDF $f(D)$, according to (18), (22) and (23).

By an analogous argument, we can update the joint PDF $f(\tilde{B}_1, \hat{B}_1)$ of the current backward message from the joint PDF $f(\tilde{B}_1^-, \hat{B}_1^-)$.

3.3 PDF of Channel Message

Next, we show a method to update the PDF $f(\hat{C})$ of the channel message $\hat{C}$ which is the input to the SPA.

Using $\hat{A}_1$ and $\hat{B}_1$ in (9), $\hat{C}$ is calculated as (see [8])

$$\hat{C} = \phi(Y) \log \frac{\hat{N}_1^C + \hat{N}_2^C\hat{A}_1}{\hat{D}_1^C + \hat{D}_2^C\hat{A}_1}, \tag{24}$$

where

$$\begin{aligned} \hat{N}_1^C &= (1 - \hat{\eta}_B)(1 - \hat{g} - \hat{B}_1 + 2\hat{g}\hat{B}_1), \\ \hat{N}_2^C &= (1 - \hat{\eta}_G)(\hat{b} + (1 - 2\hat{b})\hat{B}_1) \\ &\quad -(1 - \hat{\eta}_B)(1 - \hat{g} - \hat{B}_1 + 2\hat{g}\hat{B}_1), \\ \hat{D}_1^C &= \hat{\eta}_B(1 - \hat{g} - \hat{B}_1 + 2\hat{g}\hat{B}_1), \\ \hat{D}_2^C &= \hat{\eta}_G(\hat{b} + (1 - 2\hat{b})\hat{B}_1) \\ &\quad -\hat{\eta}_B(1 - \hat{g} - \hat{B}_1 + 2\hat{g}\hat{B}_1). \end{aligned} \tag{25}$$

Since $\hat{C}$ is determined by $Y, \hat{A}_1$ and $\hat{B}_1$ from (24), $\hat{A}_1$ is also determined by $Y, \hat{B}_1$ and $\hat{C}$. Given Y and $\hat{B}_1$, $\hat{A}_1$ can be expressed from (24) as a function of $\hat{C}$

$$\hat{A}_1 = \mathcal{A}_1^C(\hat{C}) = \frac{\hat{D}_1^C \exp\{\phi(Y)\hat{C}\} - \hat{N}_1^C}{\hat{N}_2^C - \hat{D}_2^C \exp\{\phi(Y)\hat{C}\}}. \tag{26}$$

Its derivative is

$$\begin{aligned} \mathcal{A}_1^{C\prime}(\hat{C}) &= \frac{d\mathcal{A}_1^C(\hat{C})}{d\hat{C}} \\ &= \frac{(\hat{D}_1^C\hat{N}_2^C - \hat{D}_2^C\hat{N}_1^C)\exp\{\phi(Y)\hat{C}\}}{(\hat{N}_2^C - \hat{D}_2^C \exp\{\phi(Y)\hat{C}\})^2}. \end{aligned} \tag{27}$$

Marginalizing $f(\tilde{A}_1, \hat{A}_1|S)$ which can be calculated in the same manner as (18), we obtain $f(\hat{A}_1|S)$ as follows:

$$f(\hat{a}_1|s) = \int_{\tilde{a}_1} f(\tilde{a}_1, \hat{a}_1|s)\mathrm{d}\tilde{a}_1. \tag{28}$$

Furthermore, since $\hat{A}_1$ is conditionally independent of Y and $\hat{B}_1$ given $S \in \mathcal{S}$, note that $f(\hat{A}_1|S,Y,\hat{B}_1) = f(\hat{A}_1|S)$. Therefore we have $f(\hat{C}|S,Y,\hat{B}_1)$ as follows:

$$f(\hat{c}|s,y,\hat{b}_1) = f(\mathcal{A}_1^C(\hat{c})|s)|\mathcal{A}_1^{C\prime}(\hat{c})|, \tag{29}$$

by using the transformation of variables.

In the sequel, we can obtain the PDF $f(\hat{C})$ by the marginalization as follows:

$$\begin{aligned} f(\hat{c}) = &\sum_{s\in\mathcal{S}}\sum_{y\in\{0,1\}} \tilde{p}(s)\tilde{p}(y|s) \\ &\cdot \int_{\hat{b}_1}\int_{\tilde{b}_1} f(\hat{c}_1|s,y,\hat{b}_1) f(\tilde{b}_1,\hat{b}_1|s)\mathrm{d}\tilde{b}_1\mathrm{d}\hat{b}_1. \end{aligned} \tag{30}$$

3.4 DE Algorithm for GE-LDPC Decoder Using Estimated Channel Parameter Vector

In this subsection, we show a DE algorithm for the case where the GE-LDPC decoder uses the estimated channel parameter vector $\hat{\theta}$ which deffers from the true one $\tilde{\theta}$.

Definition 5: Let $\delta(\cdot)$ be the Dirac delta function and let $\tilde{\eta}$ and $\hat{\eta}$ be the average inversion probabilities corresponding to $\tilde{\theta}$ and $\hat{\theta}$, respectively. These are given by

$$\tilde{\eta} = \frac{\tilde{g}\tilde{\eta}_G + \tilde{b}\tilde{\eta}_B}{\tilde{g}+\tilde{b}}, \qquad \hat{\eta} = \frac{\hat{g}\hat{\eta}_G + \hat{b}\hat{\eta}_B}{\hat{g}+\hat{b}}. \tag{31}$$

Furthermore let $\mathcal{F}(f(X))$ denote the Fourier transform of a function $f(X)$, and let $\mathcal{F}^{-1}$ denote the inverse transform of $\mathcal{F}$. $l_{\max}$ is the maximum iteration nunber and $P_{err}(j)$ represents the probability of symbol error after jth iteration of GE-LDPC decoding. □

The DE algorithm iterates until the iteration number exceeds $l_{\max}$ or $P_{err}(j) < \epsilon$, where ϵ is a small number. The DE algorithm proceeds as follows:

1) Let $j := 0$, and set the initial PDFs:

$$\begin{aligned} &f^{(0)}_{\tilde{A}_1,\hat{A}_1}(\tilde{a}_1,\hat{a}_1) = \delta\left(\tilde{a}_1 - \frac{\tilde{g}}{\tilde{g}+\tilde{b}}\right)\delta\left(\hat{a}_1 - \frac{\hat{g}}{\hat{g}+\hat{b}}\right), \\ &f^{(0)}_{\tilde{B}_1,\hat{B}_1}(\tilde{b}_1,\hat{b}_1) = \delta(\tilde{b}_1 - 1/2)\delta(\hat{b}_1 - 1/2), \\ &f^{(0)}_{\hat{C}}(\hat{c}) = \tilde{\eta}\delta(\hat{c} + \log(1-\hat{\eta})/\hat{\eta}) \\ &\qquad + (1-\tilde{\eta})\delta(\hat{c} - \log(1-\hat{\eta})/\hat{\eta}), \\ &f^{(0)}_P = f^{(0)}_{\hat{C}}. \end{aligned}$$

2) Using the PDF $f^{(j)}_P$ of the message P which is the output of a symbol-variable node, calcurate the PDF $f^{(j)}_Q$ of the massage Q which is the output of a parity-check node [2]. Letting $P_1, P_2, \ldots, P_{d_c-1}$ be $d_c - 1$ random variables with the PDF $f^{(j)}_P$, we obtain the PDF $f^{(j)}_Q$ by using the relation that the message Q is calculated by

$$\tanh(Q/2) = \prod_{i=1}^{d_c-1}\tanh(P_i/2). \tag{32}$$

3) Using the PDF $f^{(j)}_Q$, calculate the PDF $f^{(j)}_D$ of the extrinsic information D as follows:

$$f^{(j)}_D = \mathcal{F}^{-1}[\mathcal{F}(f^{(j)}_Q)^{d_v}]. \tag{33}$$

4) Calculate $f^{(j+1)}_{\tilde{A}_1,\hat{A}_1}$ from $f^{(j)}_D$ and $f^{(j)}_{\tilde{A}_1,\hat{A}_1}$, according to (23). Similarly, calculate $f^{(j+1)}_{\tilde{B}_1,\hat{B}_1}$ from $f^{(j)}_D$ and $f^{(j)}_{\tilde{B}_1,\hat{B}_1}$.

Calculate $f^{(j+1)}_{\hat{C}}$ from $f^{(j)}_{\tilde{A}_1,\hat{A}_1}$ and $f^{(j)}_{\tilde{B}_1,\hat{B}_1}$, according to (30).

5) From $f^{(j+1)}_{\hat{C}}$ and $f^{(j)}_Q$, update

$$f^{(j+1)}_P = \mathcal{F}^{-1}[\mathcal{F}(f^{(j+1)}_{\hat{C}})\mathcal{F}(f^{(j)}_Q)^{d_v-1}]. \tag{34}$$

6) If $j \geq l_{\max}$ or the following inequality holds, stop.

$$\begin{aligned} P_{err}(j) = &\int_{-\infty}^{0} \mathcal{F}^{-1}[\mathcal{F}(f^{(j+1)}_P)\mathcal{F}(f^{(j)}_Q)^{d_v}](x)\mathrm{d}x \\ &< \epsilon. \end{aligned} \tag{35}$$

Otherwise setting $j := j+1$, go to 2). □

Here, we briefly describe the modifications to calculate DE for irregular LDPC codes [8], [9]. Let λ_i and ρ_i be the probabilities that a given edge in an LDPC subgraph is connected to a symbol-variable node and check node of degree i, respectively. Then (34) is replaced with

$$f^{(j+1)}_P = \mathcal{F}^{-1}\left[\mathcal{F}(f^{(j+1)}_{\hat{C}})\sum_{i=1}^{v_{\max}}\lambda_i\mathcal{F}(f^{(j)}_Q)^{i-1}\right], \tag{36}$$

where $v_{\max}$ represents the maximum variable degree, and (33) is replaced with

$$f^{(j)}_D = \mathcal{F}^{-1}\left[\sum_{i=1}^{v_{\max}}\bar{\lambda}_i\mathcal{F}(f^{(j)}_Q)^{i}\right], \tag{37}$$

where $\bar{\lambda}_i$ denotes the probability that a given symbol-variable node in an LDPC subgraph has degree i, that is, $\bar{\lambda}_i = \frac{\lambda_i/i}{\sum_{j=1}^{v_{\max}}\lambda_j/j}$. Meanwhile, let $f^{(j)}_{Q,i}$ represent the PDF of a message at the output of a check node of degree i after jth iteration. In step 2), we calculate $f^{(j)}_Q$ by using $f^{(j)}_Q = \sum_{i=1}^{c_{\max}}\rho_i f^{(j)}_{Q,i-1}$, where $c_{\max}$ represents the maximum check degree.

4. Discussions

4.1 Relation between Conventional and Proposed DE

The conventional DE in [8] assumes that the decoder uses the true channel parameter vector $\tilde{\theta}$ and provides performance thresholds which the decoder converges to zero error probability over the GE channel. The proposed DE can analyze the case where an estimated channel parameter vector $\hat{\theta}$, which the decoder uses, differs from the true one.

If we set $\hat{\theta} = \tilde{\theta}$ in the proposed DE, the obtained thresholds are the same as the results of the conventional DE. More precisely, noting that $\tilde{a}_1 = \hat{a}_1$ holds for $\hat{\theta} = \tilde{\theta}$, we have

$$f(\tilde{a}_1, \hat{a}_1) = \begin{cases} f(\tilde{a}_1) = \int_{\hat{a}_1} f(\tilde{a}_1, \hat{a}_1) \mathrm{d}\hat{a}_1, & \tilde{a}_1 = \hat{a}_1, \\ 0, & \textit{otherwise}. \end{cases} \quad (38)$$

In this case the proposed DE results in the conventional one. Thus our DE is an extension of the conventional one.

As described later, however, note that the computational complexity of the proposed DE is much larger than that of conventional DE since the proposed DE needs to update the joint PDF $f(\tilde{A}_1, \hat{A}_1)$.

4.2 Computaional Complexity of DE

In this subsection, we discuss the computational complexity for the proposed DE algorithm. The DE needs to calculate the Fourier transform and the integration of some functions iteratively. In the practical case, quantizing the continuous random variables, it is sufficient to obtain the approximate PDF by the discrete Fourier transform and numerical integration. Therefore we show the computational complexity of one iteration for the DE algorithm by the order of the number of the quantization levels.

First, we consider the steps 2), 3) and 5) of the proposed DE algorithm in Sect. 3.4. These steps need the almost same complexity as the DE algorithm for the memoryless channel. We assume that the random variables $\hat{C}, P, Q$ and D are quantized by the same number of the quantized levels, denoted by n_P. Then the calculation of each step, dominated by FFT processing, needs the complexity $O(n_P \log n_P)$.

Next, we consider the step 4) in the proposed DE algorithm. We assume that the random variables $\tilde{A}_1, \hat{A}_1, \tilde{B}_1$ and $\hat{B}_1$ are quantized by the same number of the quantized levels, denoted by n_A. The complexities to calculate the PDFs according to (18) and (23) are $O(n_A^2)$ and $O(n_A^2 n_P)$, respectively, since the number of $f(\tilde{a}_1, \hat{a}_1)$ is n_A^2 by quantization and each integration with respect to d needs $O(n_P)^\dagger$.

Furthermore, we consider the complexity to calculate the PDF $f(\hat{C})$ according to (30). For all $\hat{a}_1$ and $\hat{b}_1$ we calculate

$$\begin{aligned} f(\hat{a}_1|s) &= \int_{\tilde{a}_1} f(\tilde{a}_1, \hat{a}_1|s) \mathrm{d}\tilde{a}_1, \\ f(\hat{b}_1|s) &= \int_{\tilde{b}_1} f(\tilde{b}_1, \hat{b}_1|s) \mathrm{d}\tilde{b}_1, \end{aligned} \quad (39)$$

and the total complexity is $O(n_A^2)$. Calculating these PDFs at once, the complexity to obtain $f(\hat{C})$ is $O(n_A n_P)$ to calculate the integration with respect to $\hat{b}_1$ for all $\hat{c}$.

In the sequal, the dominant complexity is $O(n_A^2 n_P)$ to update $f(\tilde{A}_1, \hat{A}_1)$ and $f(\tilde{B}_1, \hat{B}_1)$.

By an analogous argument, the conventional DE algorithm in [8] needs the complexity $O(n_A n_P)$. Therefore, comparing the order of the complexity for simplicity, the complexity of the proposed DE is about n_A times larger than that of the conventional DE.

4.3 Case of Memoryless Channel

Although our main target is the GE channel with memory, in this subsection we consider some memoryless channels. In the case of the memoryless channel, $f(\hat{C})$, which is the PDF of channel message, is invariant under the DE since $\hat{C}$ is the channel output. Therefore, it is sufficient to consider only the PDF of the channel output.

[Binary Symmetric Channel (BSC)]

Let $\tilde{\epsilon}$ and $\hat{\epsilon}$ be true and estimated crossover probabilities of the BSC, respectively. Then an estimated channel message is $\hat{C} = \phi(Y) \log \frac{1-\hat{\epsilon}}{\hat{\epsilon}}$. Assuming all-zero codeword $\boldsymbol{X}$ is transmitted, from $\Pr(Y = 0|X = 0) = 1 - \tilde{\epsilon}$, the PDF $f(\hat{C})$ of a channel message is

$$f(\hat{c}) = \tilde{\epsilon}\delta\left(\hat{c} + \log\frac{1-\hat{\epsilon}}{\hat{\epsilon}}\right) + (1-\tilde{\epsilon})\delta\left(\hat{c} - \log\frac{1-\hat{\epsilon}}{\hat{\epsilon}}\right). \quad (40)$$

[Additive White Gaussian Noise (AWGN) Channel]

We assume that the symbol of a transmitted codeword is $X \in \{+1, -1\}$, and let $\tilde{\sigma}^2$ and $\hat{\sigma}^2$ be true and estimated variances, which are channel parameters, for the AWGN channel, respectively. Then

$$\Pr(y|x) = \frac{1}{\sqrt{2\pi}\tilde{\sigma}} \exp\left\{\frac{-(y-x)^2}{2\tilde{\sigma}^2}\right\}, \quad (41)$$

and an estimated channel message is $\hat{C} = 2Y/\hat{\sigma}^2$. Therefore, assuming the all-one codeword $\boldsymbol{X}$ is transmitted, the PDF $f(\hat{C})$ of a channel message is

$$f(\hat{c}) = \frac{\hat{\sigma}^2}{2\sqrt{2\pi}\tilde{\sigma}} \exp\left\{\frac{-(\hat{\sigma}^2\hat{c}/2 - 1)^2}{2\tilde{\sigma}^2}\right\}, \quad (42)$$

by using the transformation of variables for (41).

5. DE Analysis of Robustness

In this section, we show the robustness for the GE-LDPC decoder using the proposed DE algorithm. More precisely, we show how much the estimation error of the channel parameters gives the influence to the performance thresholds when the estimated channel parameter vector differs from the true one.

First, we consider the region of the esimated channel parameter vectors for the GE-LDPC decoder, where the decoding error probability approaches asymptotically to zero and the true channel parameter vector is fixed.

In Fig. 2 we show such estimated parameter region for a (3,4)-regular rate-1/4 LDPC code, where we set $\tilde{\theta} = (\tilde{g} = 0.02, \tilde{b} = 0.02, \tilde{\eta}_G = 0.04, \tilde{\eta}_B = 0.37)$ and $\hat{g} = \hat{b}$. In this figure, we used $l_{\max} = 2000$ and $\epsilon = 1.0^{-8}$, and we show the upper and the lower boundaries of $\hat{\eta}_B$ corresponding to

†Note that by quantization each PDF is interpreted as PMF and integration is substituted with summation.

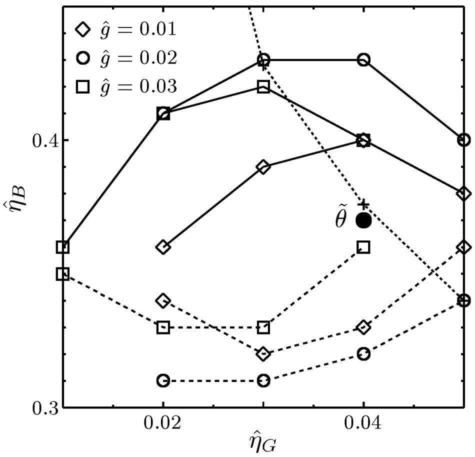

Fig. 2 Decoding region for the GE-LDPC decoder with $\hat{\theta}$ for a (3,4)-regular LDPC code, where $\tilde{\theta} = (\tilde{g} = 0.02, \tilde{b} = 0.02, \tilde{\eta}_G = 0.04, \tilde{\eta}_B = 0.37)$ is fixed and $\hat{g} = \hat{b}$. Solid and dashed lines represent the upper and the lower boundaries of $\hat{\eta}_B$ as the decoding region, respectively. Dotted line represents the true threshold.

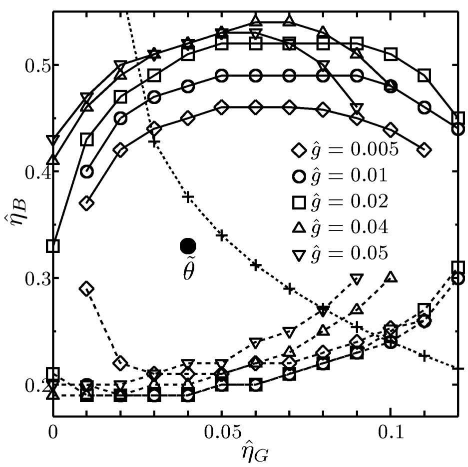

Fig. 3 Decoding region for the GE-LDPC decoder with $\hat{\theta}$ for a (3,4)-regular LDPC code, where $\tilde{\theta} = (\tilde{g} = 0.02, \tilde{b} = 0.02, \tilde{\eta}_G = 0.04, \tilde{\eta}_B = 0.33)$ is fixed and $\hat{g} = \hat{b}$. Solid and dashed lines represent the upper and the lower boundaries of $\hat{\eta}_B$ as the decoding region, respectively. Dotted line represents the true threshold.

$\hat{\eta}_G$ such that $P_{err}(l_{\max}) < \epsilon$. For reference, we also show the threshold for the case where $\hat{\theta} = \tilde{\theta} = (\tilde{g}, \tilde{b}, \hat{\eta}_G, \hat{\eta}_B)$. This threshold is depicted with a dotted line in Fig. 2 and we refer to as the *true threshold*. Similarly, in Fig. 3 we show the case where $\tilde{\theta} = (\tilde{g} = 0.02, \tilde{b} = 0.02, \tilde{\eta}_G = 0.04, \tilde{\eta}_B = 0.33)$ is fixed and $\hat{g} = \hat{b}$, where only $\tilde{\eta}_B$ is slightly smaller than the case in Fig. 2.

Figure 2 shows that the GE-LDPC decoder using $\hat{\theta}$ in a certain region can decode correctly even if $\tilde{\theta}$ is close to the true threshold. Furthermore, Fig. 3 shows that the decoding region of GE-LDPC decoder is much larger than that for the case in Fig. 2. Therefore the decoding region tends to expand as $\tilde{\theta}$ gets away from the true threshold. This can be explained by using the extrinsic information transfer (EXIT) chart. Let $\mathrm{P_e}(f)$ be the error probability of a

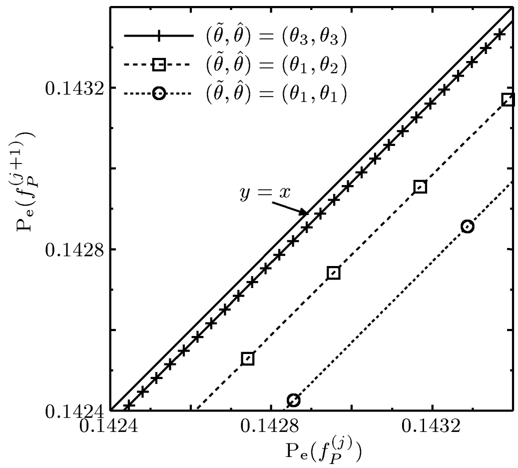

Fig. 4 The EXIT chart of the proposed DE for the cases where the tuples $(\tilde{\theta}, \hat{\theta})$ are set in (θ_1, θ_1), (θ_1, θ_2) and (θ_3, θ_3), respectively, where $\theta_1 = (g = 0.02, b = 0.02, \eta_G = 0.04, \eta_B = 0.37)$, $\theta_2 = (g = 0.01, b = 0.01, \eta_G = 0.03, \eta_B = 0.39)$ and $\theta_3 = (g = 0.02, b = 0.02, \eta_G = 0.04, \eta_B = 0.376)$.

PDF $f(X)$, i.e. $\mathrm{P_e}(f) = \int_{-\infty}^{0} f(x)\mathrm{d}x$. Furthermore, let P be a message from a symbol-variable node to a check node, and let $f_P^{(j)}(p)$ be the PDF of jth iteration of the proposed DE. If the proposed DE succeeds[†] for some tuple $(\tilde{\theta}, \hat{\theta})$, we have $\mathrm{P_e}(f_P^{(j+1)}) < \mathrm{P_e}(f_P^{(j)})$ for all j. When $\mathrm{P_e}(f_P^{(j)})$ and $\mathrm{P_e}(f_P^{(j+1)})$ are set in the x(horizontal)-axis and y(vertical)-axis of an EXIT chart[††], respectively, this implies that all points of $(\mathrm{P_e}(f_P^{(j)}), \mathrm{P_e}(f_P^{(j+1)}))$ are below the function $y = x$. In Fig. 4, we show some points of $(\mathrm{P_e}(f_P^{(j)}), \mathrm{P_e}(f_P^{(j+1)}))$ for the cases where the tuples $(\tilde{\theta}, \hat{\theta})$ are set in (θ_1, θ_1), (θ_1, θ_2) and (θ_3, θ_3), respectively, where $\theta_1 = (g = 0.02, b = 0.02, \eta_G = 0.04, \eta_B = 0.37)$, $\theta_2 = (g = 0.01, b = 0.01, \eta_G = 0.03, \eta_B = 0.39)$ and $\theta_3 = (g = 0.02, b = 0.02, \eta_G = 0.04, \eta_B = 0.376)$. Some points $(\mathrm{P_e}(f_P^{(j)}), \mathrm{P_e}(f_P^{(j+1)}))$ in Fig. 4 are selected such that $(\mathrm{P_e}(f_P^{(j)}) - \mathrm{P_e}(f_P^{(j+1)}))/\mathrm{P_e}(f_P^{(j)})$ are small, i.e. these points are relatively close to the function $y = x$. In Fig. 4 a solid line corresponding to $(\tilde{\theta}, \hat{\theta}) = (\theta_3, \theta_3)$ is very close to the function $y = x$ since θ_3 is very close to the true threshold. Therefore, in the case where $\tilde{\theta} = \theta_3$, if $\hat{\theta}$ is far from θ_3, then a sequence of $\mathrm{P_e}(f_P^{(j)})$ will converge to some positive fixed point since a trajectory of $(\mathrm{P_e}(f_P^{(j)}), \mathrm{P_e}(f_P^{(j+1)}))$ cannot pass through between the function $y = x$ and a solid line corresponding to $(\tilde{\theta}, \hat{\theta}) = (\theta_3, \theta_3)$ in Fig. 4. On the other hand, a dotted line correnponding to $(\tilde{\theta}, \hat{\theta}) = (\theta_1, \theta_1)$ is far from the function $y = x$ as compared with the case of $(\tilde{\theta}, \hat{\theta}) = (\theta_3, \theta_3)$. Therefore, a dashed line corresponding to $(\tilde{\theta}, \hat{\theta}) = (\theta_1, \theta_2)$ in Fig. 4 exists between the function $y = x$ and a dotted line corresponding to $(\tilde{\theta}, \hat{\theta}) = (\theta_1, \theta_1)$ although there is a difference between $\hat{\theta}(= \theta_2)$ and $\tilde{\theta}(= \theta_1)$. In this way, as $\tilde{\theta}$ gets away from the true threshold, the decoding region tends to expand. Inversely, as $\tilde{\theta}$ gets closer to the true threshold, more accuracy tends to be required in estimation.

Next, we consider the DE for irregular LDPC codes.

[†]This implies that $P_{err}(l_{\max}) < \epsilon$.
[††]This EXIT chart has been used in [12].

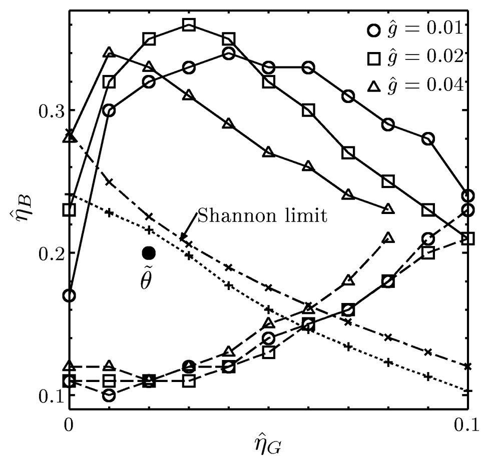

Fig. 5 Decoding region for the GE-LDPC decoder with $\hat{\theta}$ for a (λ, ρ)-irregular LDPC code, where $\tilde{\theta} = (\tilde{g} = 0.02, \tilde{b} = 0.02, \tilde{\eta}_G = 0.02, \tilde{\eta}_B = 0.2)$ is fixed and $\hat{g} = \hat{b}$. Solid and dashed lines represent the upper and the lower boundaries of $\hat{\eta}_B$ as the decoding region, respectively. Dotted line represents the true threshold.

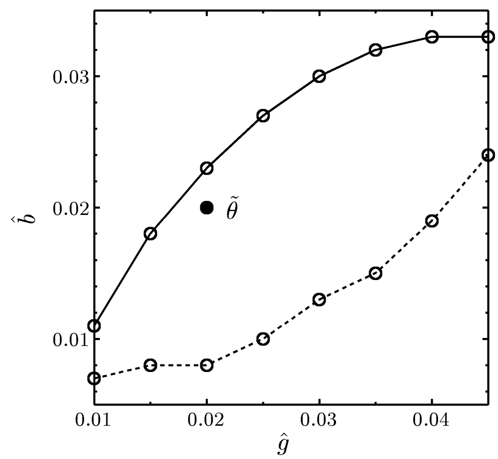

Fig. 6 Decoding region of $(\hat{g}, \hat{b})$ for the case where $\tilde{\theta} = (\tilde{g} = 0.02, \tilde{b} = 0.02, \tilde{\eta}_G = 0.04, \tilde{\eta}_B = 0.37)$, $\hat{\eta}_G = 0.04$ and $\hat{\eta}_B = 0.4$ are fixed. Solid and dashed lines represent the upper and the lower boundaries of $\hat{b}$ as the decoding region, respectively.

Let $\lambda(x)$ and $\rho(x)$ be polynomials representing the variable node and check node degree distributions, respectively, that is, $\lambda(x) = \sum_{i=1}^{v_{\max}} \lambda_i x^{i-1}$ and $\rho(x) = \sum_{i=1}^{c_{\max}} \rho_i x^{i-1}$. In Fig. 5 we show the result for a (λ, ρ)-irregular rate-1/2 LDPC code, where we set $\tilde{\theta} = (\tilde{g} = 0.02, \tilde{b} = 0.02, \tilde{\eta}_G = 0.02, \tilde{\eta}_B = 0.2)$ and $\hat{g} = \hat{b}$. In Fig. 5, we used the following degree distribution pair as in [3],

$$\begin{aligned}\lambda(x) = {} & 0.17120x + 0.21053x^2 + 0.00273x^3 \\ & + 0.00009x^6 + 0.15269x^7 + 0.09227x^8 \\ & + 0.02802x^9 + 0.01206x^{14} + 0.07212x^{29} \\ & + 0.25830x^{49}, \end{aligned} \tag{43}$$

$$\rho(x) = 0.33620x^8 + 0.08883x^9 + 0.57497x^{10}. \tag{44}$$

Figure 5 shows that the decoding region is sufficiently large for the irregular LDPC code as well as the result of Fig. 3.

These results show that an estimation error of channel parameters does not give a big influence to the decoding performance as $\tilde{\theta}$ goes from the true threshold, if the code length is large enough.

Moreover the decoding regions of tuple $(\hat{\eta}_G, \hat{\eta}_B)$ vary according to values of $\hat{g}(= \hat{b})$. We here discuss influence of the sign of the estimation error $\hat{g} - \tilde{g}$. That is, if $\hat{g} > \tilde{g}$, estimation error $\hat{g} - \tilde{g}$ is positive. In Figs. 3 and 5, if $\hat{g} > \tilde{g}$, for small $\hat{\eta}_G$ the decoding region of $\hat{\eta}_B$ tends to be large compared with the case of $\hat{g} < \tilde{g}$. As $\hat{\eta}_G$ is larger for $\hat{g} > \tilde{g}$, inversely, the decoding region of $\hat{\eta}_B$ tends to be smaller compared with the case of $\hat{g} < \tilde{g}$. This implies that the direction of the estimation error $\hat{\theta} - \tilde{\theta}$ gives great influence to the decoding performance.

In Fig. 6 we show the decoding region of $(\hat{g}, \hat{b})$ for a (3,4)-regular LDPC code, where $\tilde{\theta} = (\tilde{g} = 0.02, \tilde{b} = 0.02, \tilde{\eta}_G = 0.04, \tilde{\eta}_B = 0.37)$, $\hat{\eta}_G = 0.04$ and $\hat{\eta}_B = 0.4$ are fixed. It is interesting to note that there is a large decoding region including the misestimation of the stationary state probability, that is, $\hat{g}/(\hat{g} + \hat{b}) \neq \tilde{g}/(\tilde{g} + \tilde{b})$. Figure 6

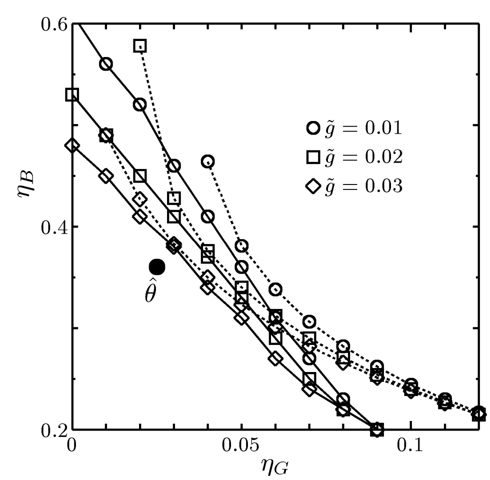

Fig. 7 Decoding region for the GE-LDPC decoder with $\hat{\theta}$ for a (3,4)-regular LDPC code, where $\hat{\theta} = (\hat{g} = 0.02, \hat{b} = 0.02, \hat{\eta}_G = 0.025, \hat{\eta}_B = 0.36)$ is fixed and $\tilde{g} = \tilde{b}$. Solid lines represent the upper boundaries of $\tilde{\eta}_B$ as the decoding region. Dotted lines represent the true thresholds.

also shows that the direction of the estimation error $\hat{\theta} - \tilde{\theta}$ is very important since the decoding region for the case where $\hat{g}/(\hat{g} + \hat{b}) > \tilde{g}/(\tilde{g} + \tilde{b})$ is much larger than that for the case where $\hat{g}/(\hat{g} + \hat{b}) < \tilde{g}/(\tilde{g} + \tilde{b})$.

Next, we consider the region of the true channel parameter vectors, where the decoding error probability approaches asymptotically to zero for a fixed estimated channel parameter vector $\hat{\theta}$.

We set $\hat{\theta} = (\hat{g} = 0.02, \hat{b} = 0.02, \hat{\eta}_G = 0.025, \hat{\eta}_B = 0.36)$ for a (3,4)-regular LDPC code, since this is the estimated channel parameter vector such that $P_{err}(l_{\max}) < \epsilon$ in Figs. 2 and 3. This result is presented in Fig. 7. Similarly, in Fig. 8 we show the result for a (λ, ρ)-irregular LDPC code, where λ and ρ are given by (43) and (44), respectively. Then we set $\hat{\theta} = (\hat{g} = 0.02, \hat{b} = 0.02, \hat{\eta}_G = 0.01, \hat{\eta}_B = 0.21)$ and $\tilde{g} = \tilde{b}$.

Surprisingly Figs. 7 and 8 show that the decoding region of $\tilde{\theta}$ is very large, although the decoder used a fixed $\hat{\theta}$.

Furthermore there exist some decodable points near the true threshold in these figures even when $\tilde{g} \neq \hat{g}$.

Here, we consider the case of the finite code length n since the assumption of the DE in Sect. 3.1 requires that the code length is infinite. Experimental results for a (λ, ρ)-irregular rate-1/2 LDPC code are presented in Fig. 9, where λ and ρ are given by (43) and (44), respectively. In this figure, we show the results for the following two decoders: (i) the ideal GE-LDPC decoder which uses the true channel parameter vector $\tilde{\theta}$, represented with dashed lines and (ii) the estimated GE-LEPC decoder which uses the estimated channel parameter vector $\hat{\theta}$, represented with solid lines, where $\hat{\theta} = (\hat{g} = 0.02, \hat{b} = 0.02, \hat{\eta}_G = 0.01, \hat{\eta}_B = 0.21)$ and $\tilde{\eta}_G = 0.02$ are fixed. In Figs. 9(a) and (b) for $\hat{g} \geq \tilde{g}$, there is no big difference between the BER performance of two decoders even when $\tilde{\eta}_B$ is small. These support the results of our DE analysis. In Fig. 9(c) for $\hat{g} < \tilde{g}$, however, the estimated GE-LDPC decoder has degradation of the BER performance compared with the ideal one. In Fig. 5, the upper boundary of $\hat{\eta}_B$ for $\hat{g} = 0.01$ and $\hat{\eta}_G = 0.01$ is smaller than the cases for $\hat{g} \geq \tilde{g} = 0.02$ and $\hat{\eta}_G = 0.01$. Thus it seems that this degradation was caused by the direction of the estimation error $\hat{\theta} - \tilde{\theta}$.

From the above results we can conclude that the GE-LDPC decoder over the GE channel has the robustness. These results can be used to design the channel estimater and the frequency of the channel estimation. Furthermore it may be possible to simplify the decoder.

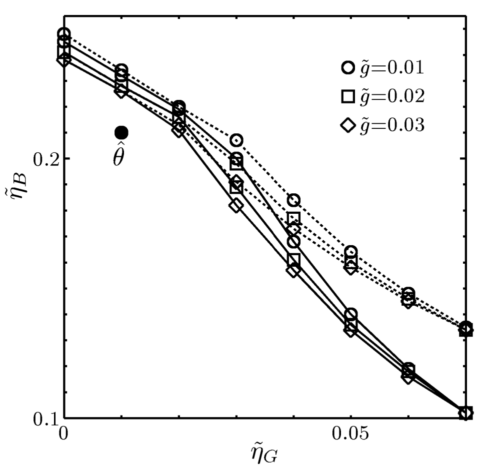

Fig. 8 Decoding region for the GE-LDPC decoder with $\hat{\theta}$ for a (λ, ρ)-irregular LDPC code, where $\hat{\theta} = (\hat{g} = 0.02, \hat{b} = 0.02, \hat{\eta}_G = 0.01, \hat{\eta}_B = 0.21)$ is fixed and $\tilde{g} = \tilde{b}$. Solid lines represent the upper boundaries of $\tilde{\eta}_B$ as the decoding region. Dotted lines represent the true thresholds.

6. Conclusion

In this paper, we proposed a DE algorithm for the case where the GE-LDPC decoder uses an estimated channel parameter vector over the GE channel. Then we introduced the joint PDF $f(\tilde{A}_1, \hat{A}_1)$ of the forward messages $\tilde{A}_1$ and $\hat{A}_1$ corresponding to $\tilde{\theta}$ and $\hat{\theta}$ for the GE-LDPC decoder and described a method to update this joint PDF. This DE algorithm is an extension of the DE algorithm proposed by Eckford et al. [8]. Furthermore, using the proposed DE alogrithm, we showed some examples of the region of the true and the estimated channel parameter vector achieving $P_{err}(l_{\max}) < \epsilon$. Consequently, we showed that the GE-LDEPC decoder has the robustness for LDPC codes over the GE channel and an estimation error of channel parameters does not give a big influence to the decoding performance asymptotically as $\tilde{\theta}$ goes from the true threshold. It may be possible to simplify the decoder by using the robustness.

On the other hand, the computational complexity of the proposed DE is larger than that of the conventional DE since the proposed DE needs to update the joint PDF $f(\tilde{A}_1, \hat{A}_1)$. If

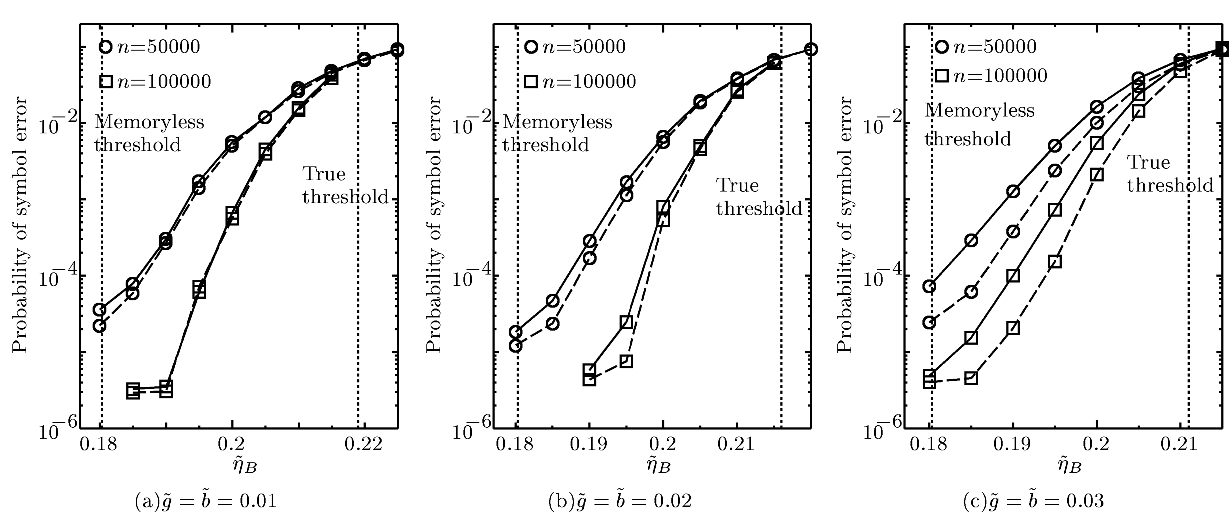

Fig. 9 Experimental results for a (λ, ρ)-irregular LDPC code, where $\hat{\theta} = (\hat{g} = 0.02, \hat{b} = 0.02, \hat{\eta}_G = 0.01, \hat{\eta}_B = 0.21)$ and $\tilde{\eta}_G = 0.02$ are fixed. Solid lines represent the results for the GE-LEPC decoder which uses $\hat{\theta}$, and dashed lines represent the results for the ideal GE-LDPC decoder which uses $\tilde{\theta}$. Dotted lines represent the thresholds.

it is possible to reduce the complexity of the proposed DE, we can consider to analyze a Markov channel, in which the cannel noise is characterized by a hidden Markov chain.

Acknowledgments

The authors wish to thank the reviewers for their invaluable comments which helped us to improve the quality of this paper.

M. Kobayashi wishes to thank Mr. G. Hosoya at Waseda University for his suggestions.

References

[1] R.G. Gallager, Low Density Parity Check Codes, MIT Press, Cambridge, MA, 1963.

[2] T.J. Richardson and R.L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," IEEE Trans. Inf. Theory, vol.47, no.2, pp.599–618, Feb. 2001.

[3] 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.

[4] S.Y. Chung, G.D. Forney, Jr., T.J. Richardson, and R. Urbanke, "On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit," IEEE Commun. Lett., vol.5, pp.58–60, Feb. 2001.

[5] T. Wadayama, "An iterative decoding algorithm of low density parity check codes for hidden Markov noise channels," International Conference on Information Theory and Its Applications, Hawaii, 2000.

[6] J. Garcia-Frias, "Decoding of low-density parity-check codes over finite-state binary Markov channels," IEEE Trans. Commun., vol.52, no.11, pp.1840–1843, Nov. 2004.

[7] Y. Kagyo and H. Ogiwara, "Decoding of LDPC codes for burst noise channels," Proc. 28th Symposium on Information Theory and Its Applications, pp.239–242, Nov. 2005.

[8] A.W. Eckford, F.R. Kschischang, and S. Pasupathy, "Analysis of low-density parity-check codes for the Gilbert-Elliott channel," IEEE Trans. Inf. Theory, vol.51, no.11, pp.3872–3889, Nov. 2005.

[9] A.W. Eckford, F.R. Kschischang, and S. Pasupathy, "On designing good LDPC codes for Markov channels," IEEE Trans. Inf. Theory, vol.53, no.1, pp.5–21, Jan. 2007.

[10] T.A. Summers and S.G. Wilson, "SNR mismatch and online estimation in turbo decoding," IEEE Trans. Commun., vol.46, no.4, pp.421–423, April 1998.

[11] C.R. Jones, T. Tian, J. Villasenor, and R.D. Wesel, "The universal operation of LDPC codes over scalar fading channels," IEEE Trans. Commun., vol.55, no.1, pp.122–132, Jan. 2007.

[12] M. Ardakani and F.R. Kschischang, "A more accurate one-dimensional analysis and design of irregular LDPC codes," IEEE Trans. Commun., vol.52, no.12, pp.2106–2114, Dec. 2004.

Appendix: Proof of Symmetry Condition

In this appendix, we show that the symmetry condition is fulfilled for the case where GE-LDPC decoder uses estimated channel parameters. Note that the notations used in this appendix are given in Sects. 3.2 and 3.3.

An estimated channel message $\hat{C}$ is symmetric if it can be written as $\hat{C} = \phi(X)T$, where $\phi(X) = (-1)^X$, X is the transmitted codeword symbol corresponding to the message $\hat{C}$, and T is a random variable independent of the transmitted codeword [8].

From (2), we have $\phi(Y) = \phi(X)\phi(Z)$, where Y is the received symbol, and Z is the noise symbol independent of a codeword. From (24) and (25),

$$\hat{C} = \phi(X)\phi(Z)\log\frac{\hat{N}_1^C + \hat{N}_2^C\hat{A}_1}{\hat{D}_1^C + \hat{D}_2^C\hat{A}_1}, \tag{A·1}$$

is symmetric if $\hat{A}_1$ and $\hat{B}_1$ are independent of a codeword.

Now we assume that the extrinsic information D is symmetric and can be written as $D = \phi(X)U$, where U is a random variable independent of the transmitted codeword. From $\tanh(-x) = -\tanh(x)$, $x \in \mathcal{R}$, and (5), we have

$$\begin{aligned}\gamma(D,Y) &= \frac{1}{2}\left[1 + \phi(X)\phi(Z)\tanh\left(\frac{\phi(X)U}{2}\right)\right]\\ &= \frac{1}{2}\left[1 + \phi(X)^2\phi(Z)\tanh\left(\frac{U}{2}\right)\right]\\ &= \gamma(U,Z).\end{aligned} \tag{A·2}$$

Therefore, $\gamma(D,Y)(=\gamma(U,Z))$ is independent of the transmitted codeword. In the initial step of the induction, we use $\hat{p}(S=G)$ as the initial value of $\hat{A}_1$. From (10) and (11), if $\hat{A}_1^-$ is independent of the transmitted codeword, $\hat{A}_1 = \mathcal{A}_1(\hat{\theta}, \hat{A}_1^-)$ is also independent of the transmitted codeword. Thus by induction, we can conclude that the messages $\hat{A}_1$ are independent of the transmitted codeword as long as the extrinsic informations D are symmetric. The same argument can be used for $\hat{B}_1$.

At last, we need to show that if $\hat{C}$ are symmetric, D are also symmetric. However, if channel messages to the SPA are symmetric, all messages passed in the SPA are also symmetric [2]. Therefore D are also symmetric.

As a result, the symmetry condition is fulfilled.

Manabu Kobayashi was born in Yokohama, Japan, on Oct. 30, 1971. He received the B.E. degree, M.E. degree and Dr.E. degree in Industrial and Management Systems Engineering form Waseda University, Tokyo, Japan, in 1994, 1996 and 2000, respectively. From 1998 to 2001, he was a research associate in Industrial and Management Systems Engineering at Waseda University. He is currently an assistant professor of the Department of Information Science at Shonan Institute of Technology, Kanagawa, Japan. His research interests are coding and information theory and data mining. He is a member of the Society of Information Theory and Its Applications, Information Processing Society of Japan and IEEE.

Hideki Yagi was born in Yokohama, Japan, on Oct. 14, 1975. He received the B.E. degree, M.E. degree, and Dr.E. degree in Industrial and Management Systems Engineering from Waseda University, Tokyo, Japan, in 2001, 2003 and 2005, respectively. From 2005 to 2007, he was a Research Associate and since 2007, has been an assistant professor at Media Network Center, Waseda University, Tokyo, Japan. His research interests are coding theory and information security. He is a member of the Society of Information Theory and its Applications and IEEE.

Toshiyasu Matsushima was born in Tokyo, Japan, on Nov. 26, 1955. He received the B.E. degree, M.E. degree and Dr.E. degree in Industrial and Management Systems Engineering from Waseda University, Tokyo, Japan, in 1978, 1980 and 1991, respectively. From 1980 to 1986, he was with Nippon Electric Corporation, Kanagawa, Japan. From 1986 to 1992, he was a lecturer at Department of Management Information, Yokohama College of Commerce. From 1993, he was an associate professor and from 1996 to 2007 was a professor of School of Science and Engineering, Waseda University, Tokyo, Japan. Since 2007, he has been a professor of School of Fundamental Science and Engineering, Waseda University. His research interests are information theory and its application, statistics and artificial intelligence. He is a member of the Society of Information Theory and Its Applications, the Japan Society for Quality Control, the Japan Industrial Management Association, the Japan Society for Artificial Intelligence and IEEE.

Shigeichi Hirasawa was born in Kobe, Japan, on Oct. 2, 1938. He received the B.S. degree in mathematics and the B.E. degree in electrical communication engineering from Waseda University, Tokyo, Japan, in 1961 and 1963, respectively, and the Dr.E. degree in electrical communication engineering from Osaka University, Osaka, Japan, in 1975. From 1963 to 1981, he was with the Mitsubishi Electric Corporation, Hyogo, Japan. From 1981 to 2007, he was a professor of School of Science and Engineering and since 2007, has been a professor of School of Creative Science and Engineering, Waseda University, Tokyo, Japan. In 1979, he was a Visiting Scholar in the Computer Science Department at the University of California, Los Angels (CSD, UCLA), CA. He was a Visiting Researcher at the Hungarian Academy of Science, Hungary, in 1985, and at the University of Trieste, Italy, in 1986. In 2002, he was also a Visiting Faculty at CSD, UCLA. From 1987 to 1989, he was the Chairman of Technical Group on Information Theory of IEICE. He received the 1993 Achievement Award, and the 1993 Kobayashi-Memorial Achievement Award from IEICE. In 1996, he was the President of the Society of Information Theory and Its Applications (Soc. of ITA). His research interests are information theory and its applications, and information processing systems. He is an IEEE Life Fellow, and a member of Soc. of ITA, the Operations Research Society of Japan, the Information Processing Society of Japan, the Japan Industrial Management Association, and Informs.