Effects of Single-Cycle Structure on Iterative Decoding of Low-Density Parity-Check Codes

Ryuhei Mori, *Student Member, IEEE*, Toshiyuki Tanaka, *Member, IEEE*, Kenta Kasai, *Member, IEEE*, and Kohichi Sakaniwa, *Senior Member, IEEE*

***Abstract*—We consider communication over the binary erasure channel (BEC) using low-density parity-check (LDPC) codes and belief propagation (BP) decoding. For fixed numbers of BP iterations, the bit error probability approaches a limit as the blocklength tends to infinity, and the limit is obtained via density evolution. The finite-blocklength correction behaves like $\alpha(\epsilon, t)/n + \Theta(n^{-2})$ as the blocklength n tends to infinity where $\alpha(\epsilon, t)$ denotes a specific constant determined by the code ensemble considered, the number t of iterations, and the erasure probability ϵ of the BEC. In this paper, we derive a set of recursive formulas which allows the evaluation of the constant $\alpha(\epsilon, t)$ for standard irregular ensembles. The dominant difference $\alpha(\epsilon, t)/n$ can be considered as effects of cycle-free and single-cycle structures of local graphs. Furthermore, it is confirmed via numerical simulations that estimation of the bit error probability using $\alpha(\epsilon, t)$ is accurate even for small blocklengths.**

***Index Terms*—low-density parity-check codes, belief propagation, binary erasure channel, density evolution, finite-length analysis.**

I. Introduction

It is well known that low-density parity-check (LDPC) codes for transmission over binary memoryless symmetric channels approach channel capacity with a low-complexity iterative decoder called *belief propagation* (BP) decoder. Especially for the binary erasure channels (BEC), suitably designed LDPC codes with BP decoder provably achieve channel capacity [4]. The large-blocklength limit of the bit error probability of the BP decoder with a fixed number of iterations can be calculated by the method called *density evolution* [5]. In this paper, we consider how fast the bit error probability approaches its limit as the blocklength tends to infinity. Although the performance analysis of LDPC codes is often developed for general memoryless binary-input output-symmetric channels [5], [6], [7], [8], [9], we restrict our attention in this paper to the case where the channel is the BEC, since performance analysis on the BEC [4], [10], [11], [12], [13] is generally simpler than that for general channels. In density evolution, the bit error probability is calculated recursively by considering tree neighborhoods whose depth is equal to the number of iterations. In the analysis of this paper, we consider not only tree neighborhood graphs but also single-cycle neighborhood graphs in order to derive the most dominant term in the bit error probability which vanishes in the large-blocklength limit. We would like to mention that it might be possible to generalize our analysis to other channels and iterative decoders since the approach taken in our analysis is based on density evolution which is applicable to any combination of a channel and an iterative decoder.

In this paper, we deal with one of the most popular LDPC ensembles, namely, standard irregular ensemble [4], [14]. Let $\mathrm{P_b}(n, \epsilon, t)$ denote the average bit error probability of a standard irregular ensemble of codes of blocklength n over the BEC(ϵ) after t BP iterations. The large-blocklength limit of the bit error probability after t iterations is denoted by $\mathrm{P_b}(\infty, \epsilon, t)$. Evaluation of $\mathrm{P_b}(\infty, \epsilon, t)$ using density evolution has revealed that there exists a threshold erasure probability ϵ_{BP} such that the bit error probability $\mathrm{P_b}(\infty, \epsilon, t)$ after a sufficient number of BP iterations tends to 0 if $\epsilon < \epsilon_{\mathrm{BP}}$ and to a strictly positive value if $\epsilon > \epsilon_{\mathrm{BP}}$.

From a practical point of view, it is desirable to evaluate $\mathrm{P_b}(n, \epsilon, t)$ for a finite n, which, however, is much more complicated than the evaluation of $\mathrm{P_b}(\infty, \epsilon, t)$. The bit and block error probabilities for finite blocklength and for an infinite number of iterations have been calculated exactly via stopping-set analysis for regular ensembles [10] and also for left-irregular right-regular ensembles [15]. Furthermore, the bit and block error probabilities of expurgated ensembles for finite blocklength and for finite numbers of iterations have also been calculated exactly in a combinatorial way for regular ensembles and left-regular Poisson ensembles [16]. However, these analyses are not applicable for the general standard irregular ensemble.

An approach to a finite-length analysis for irregular ensembles with low computational complexity would be to consider large-n asymptotics. There are two efficient methods to derive large-n asymptotics for the bit error probability for blocklength n and for an infinite number of iterations, which is denoted by $\mathrm{P_b}(n, \epsilon, \infty)$. The method proposed by Di, Richardson, and Urbanke [11] has shown that the bit error probability below the threshold after an infinite number of iterations is expressed as

$$\mathrm{P_b}(n, \epsilon, \infty) = \frac{1}{2}\frac{\epsilon\lambda'(0)\rho'(1)}{1-\epsilon\lambda'(0)\rho'(1)}\frac{1}{n} + o\left(\frac{1}{n}\right) \qquad (1)$$

where $\lambda'(x)$ and $\rho'(x)$ are the derivatives of polynomials $\lambda(x)$ and $\rho(x)$ which specify the degree distributions of variable nodes and check nodes, respectively, and are defined in Section II. One may thus obtain an approximation formula

Manuscript received October 1, 2010; revised April 26, 2012; accepted August 17, 2012. This paper was presented in part at [1], [2], [3].

*R. Mori and T. Tanaka are with the Department of System Science, Graduate School of Informatics, Kyoto University, Kyoto, 606-8501, Japan (e-mail: rmori@sys.i.kyoto-u.ac.jp, tt@i.kyoto-u.ac.jp)

†K. Kasai and K. Sakaniwa are with the Department of Communications and Integrated Systems, Graduate School of Science and Engineering, Tokyo Institute of Technology, Tokyo, 152-8552, Japan (e-mail: {kenta, sakaniwa}@comm.ss.titech.ac.jp)

for $\mathrm{P_b}(n, \epsilon, \infty)$ by ignoring the term $o(n^{-1})$ in (1). However, the approximation is not accurate near the threshold for any irregular ensembles due to the following reasons. If the limit $\lim_{n\to\infty} \mathrm{P_b}(n, \epsilon, \infty)$ is discontinuous at ϵ_{BP} as a function of ϵ (i.e., $\lim_{n\to\infty} \mathrm{P_b}(n, \epsilon_{\mathrm{BP}}, \infty) > 0$), convergence to the limit is not uniform since $\mathrm{P_b}(n, \epsilon, \infty)$ for any finite n is continuous with respect to ϵ. Hence, an arbitrarily large blocklength is required near the threshold so that the above approximation formula is expected to be accurate. On the other hand, the convergence is uniform for $\epsilon \in [0, \epsilon_{\mathrm{BP}}]$ if the limit $\lim_{n\to\infty} \mathrm{P_b}(n, \epsilon, \infty)$ is continuous at ϵ_{BP} as a function of ϵ (i.e., $\lim_{n\to\infty} \mathrm{P_b}(n, \epsilon_{\mathrm{BP}}, \infty) = 0$). In such cases, however, the coefficient of n^{-1} in (1) diverges as ϵ approaches the threshold ϵ_{BP} from below, since the threshold is given as $\epsilon_{\mathrm{BP}} = (\lambda'(0)\rho'(1))^{-1}$ in this case. Hence, an arbitrarily large blocklength is again required near the threshold so that the above-mentioned approximation formula is expected to be accurate. From the above facts, the approximation (1) is accurate only for a small-ϵ region which is often called an error floor.

As an alternative approach, a method that is based on the scaling law has been proposed [13], [17], which requires only a constant cost and is useful for estimating the bit and block error probabilities near the threshold where the error probabilities behave like what is called a waterfall curve. This analysis permits finite-length optimization which maximizes the rate of a code under a given blocklength, erasure probability and allowable error probability.

Both of these two methods are, however, applicable only for an infinite number of iterations, whereas the number of iterations is often constrained in practical applications due to limitation of resources, e.g., time, energy, etc., so that results for a finite number of iterations should be more significant than those for an infinite number of iterations. We therefore focus in this paper on an asymptotic bit error probability with respect to the blocklength when the number t of iterations is finite and fixed. The basic idea underlying our approach is to consider a large-n asymptotic expansion of the bit error probability and to evaluate the second dominant term in the asymptotic expansion. There exists a coefficient $\alpha(\epsilon, t)$ of n^{-1} on the basis of which the asymptotic expansion of $\mathrm{P_b}(n, \epsilon, t)$ is expressed as

$$\mathrm{P_b}(n, \epsilon, t) = \mathrm{P_b}(\infty, \epsilon, t) + \alpha(\epsilon, t)\frac{1}{n} + o\left(\frac{1}{n}\right). \tag{2}$$

The second term $\alpha(\epsilon, t)/n$ on the right-hand side of (2) is determined by tree and single-cycle structures of local graphs, while the first term $\mathrm{P_b}(\infty, \epsilon, t)$ is due to only tree local graphs. An important consequence of considering a finite-t asymptotic expansion is that the approximation formula derived by ignoring the term $o\left(n^{-1}\right)$ in (2) is expected to be accurate for all ϵ uniformly if the blocklength is sufficiently large, since the convergence $\lim_{n\to\infty} \mathrm{P_b}(n, \epsilon, t)$ is uniform for $\epsilon \in [0, 1]$, as we will see in later sections. Our main result is to derive a set of recursive formulas which allows the evaluation of the coefficient $\alpha(\epsilon, t)$ for irregular ensembles.

This paper is organized as follows. In Section II, we define the random ensembles of graphs used in this paper. In Section III-A, we see how the coefficient $\alpha(\epsilon, t)$ is decomposed into two components, one representing contributions of cycle-free neighborhood graphs and the other representing contributions of single-cycle neighborhood graphs. In Section III-B, we obtain the component for cycle-free neighborhood graphs in $\alpha(\epsilon, t)$ by developing a generating function method. In Section III-C, we see how to enumerate the coefficient of n^{-1} in the asymptotic expansion of the probability for single-cycle neighborhood graphs. The technique developed in Section III-C is then used in the calculation of the contributions of single-cycle neighborhood graphs in Section III-E via the single-cycle neighborhood graph ensemble defined in Section III-D. In Section IV, we study the limit $\lim_{t\to\infty} \alpha(\epsilon, t)$ for regular ensembles. In Section V, we show that when the number of iterations is fixed, the large-blocklength convergence is uniform with respect to ϵ. It implies that, for a sufficiently large blocklength, the approximation (2) is uniformly accurate for all ϵ. Furthermore, in Section VI, it is confirmed via numerical simulations that the approximations for several ensembles are accurate even for small blocklength. Finally, we conclude this paper in Section VII.

II. Preliminaries

A. Tanner graphs

A Tanner graph $H = (\mathcal{V}, \mathcal{C}, \mathcal{S}_{\mathcal{V}}, \mathcal{S}_{\mathcal{C}}, \mathcal{E})$ is a bipartite graph which is represented by a set $\mathcal{V}$ of variable nodes, a set $\mathcal{C}$ of check nodes, a set $\mathcal{S}_{\mathcal{V}}$ of variable-node sockets, a set $\mathcal{S}_{\mathcal{C}}$ of check-node sockets, and a set $\mathcal{E}$ of edges connecting variable-node and check-node sockets. A node u is identified as a tuple of sockets associated with u. The sets $\mathcal{V}$ and $\mathcal{C}$ satisfy the following conditions.

$$\mathcal{V} \subseteq \bigcup_{k=1}^{\infty} \{(s_1, \dots, s_k) \mid s_1 \in \mathcal{S}_{\mathcal{V}}, \dots, s_k \in \mathcal{S}_{\mathcal{V}}\}$$
$$\mathcal{C} \subseteq \bigcup_{k=1}^{\infty} \{(s'_1, \dots, s'_k) \mid s'_1 \in \mathcal{S}_{\mathcal{C}}, \dots, s'_k \in \mathcal{S}_{\mathcal{C}}\}.$$

Moreover, sockets associated with the same node are all distinct. Each variable-node socket $s \in \mathcal{S}_{\mathcal{V}}$ is associated with one and only one variable node in $\mathcal{V}$. Similarly, each check-node socket $s' \in \mathcal{S}_{\mathcal{C}}$ is associated with one and only one check node in $\mathcal{C}$. An edge e is identified as a pair of sockets which connect to e. To be precise,

$$\mathcal{E} \subseteq \{(s, s') \mid s \in \mathcal{S}_{\mathcal{V}}, s' \in \mathcal{S}_{\mathcal{C}}\}.$$

Each socket connects to one and only one edge. The number of variable-node sockets, the number of check-node sockets and the number of edges are equal, i.e., $|\mathcal{S}_{\mathcal{V}}| = |\mathcal{S}_{\mathcal{C}}| = |\mathcal{E}|$. Here, and in what follows, $|\mathcal{A}|$ denotes the cardinality of a set $\mathcal{A}$. The same notation is also applied to a node in order to denote its degree. For example, for $v \in \mathcal{V}$, $|v|$ denotes the degree of the variable node v.

B. Irregular LDPC code ensembles

An $(n, L(x), R(x))$-*irregular ensemble* is a random ensemble of LDPC codes of blocklength n which are represented by

Tanner graphs with variable-node degree distribution polynomial $L(x)$ from the node perspective and check-node degree distribution polynomial $R(x)$ from the node perspective [4]. These two degree distribution polynomials are expressed as

$$L(x) := \sum_i L_i x^i, \qquad R(x) := \sum_j R_j x^j.$$

Each Tanner graph in the $(n, L(x), R(x))$-irregular ensemble has n variable nodes, a fraction L_i of variable nodes of degree i, and a fraction R_j of check nodes of degree j. The sets $\mathcal{V}$, $\mathcal{C}$, $\mathcal{S}_{\mathcal{V}}$ and $\mathcal{S}_{\mathcal{C}}$ defining nodes and sockets in the Tanner graphs are arbitrarily fixed in an ensemble. Each instance of the edge set $\mathcal{E}$ is chosen randomly from all $E!$ possible realizations with uniform probability, where $E := |\mathcal{E}| = nL'(1)$ is the number of edges of the Tanner graphs.

We also use the degree distribution polynomials $\lambda(x)$ and $\rho(x)$ from the edge perspective, which are defined as

$$\lambda(x) = \sum_i \lambda_i x^{i-1} := \frac{L'(x)}{L'(1)}$$
$$\rho(x) = \sum_j \rho_j x^{j-1} := \frac{R'(x)}{R'(1)}.$$

An $(n, L(x), R(x))$-irregular ensemble has a fraction λ_i of edges incident to variable nodes of degree i and a fraction ρ_j of edges incident to check nodes of degree j. An $(n, L(x), R(x))$-irregular ensemble is also referred to as an $(n, \lambda(x), \rho(x))$-irregular ensemble.

The bit error probability of the $(n, \lambda(x), \rho(x))$*-irregular ensemble* is defined as the average bit error probability of the $(n, \lambda(x), \rho(x))$-irregular ensemble. In this paper, we deal with an asymptotic bit error probability with respect to blocklength, whereby the degree distribution $(\lambda(x), \rho(x))$ is fixed. In the following, we will also use the notation $(\lambda(x), \rho(x))$ to specify an irregular ensemble when the blocklength is not relevant.

C. Neighborhood graph ensembles

In a Tanner graph, we define the distance of a variable node v from a variable node v' as the number of check nodes in the shortest path from v' to v. Similarly, we define the distance of a check node c from a variable node v' as the number of check nodes minus 1 in the shortest path from v' to c. A *neighborhood graph* G of depth t of a variable node v_0 is a subgraph which consists of the variable nodes and the check nodes with distance from v_0 not greater than t and $(t-1)$, respectively. Each neighborhood graph is expressed as $G = (\mathcal{V}_{\mathcal{N}}, \mathcal{C}_{\mathcal{N}}, v_0, \mathcal{S}_{\mathcal{V}_{\mathcal{N}}}, \mathcal{S}_{\mathcal{C}_{\mathcal{N}}}, \mathcal{E}_{\mathcal{N}})$. Sets $\mathcal{V}_{\mathcal{N}}$, $\mathcal{C}_{\mathcal{N}}$, $\mathcal{S}_{\mathcal{V}_{\mathcal{N}}}$, $\mathcal{S}_{\mathcal{C}_{\mathcal{N}}}$ and $\mathcal{E}_{\mathcal{N}}$ are, respectively, a set of variable nodes, a set of check nodes, a set of variable-node sockets, a set of check-node sockets and a set of edges. Expressions and roles of $\mathcal{V}_{\mathcal{N}}$, $\mathcal{C}_{\mathcal{N}}$, $\mathcal{S}_{\mathcal{V}_{\mathcal{N}}}$, $\mathcal{S}_{\mathcal{C}_{\mathcal{N}}}$ and $\mathcal{E}_{\mathcal{N}}$ are the same as those of $\mathcal{V}$, $\mathcal{C}$, $\mathcal{S}_{\mathcal{V}}$, $\mathcal{S}_{\mathcal{C}}$ and $\mathcal{E}$ for a Tanner graph, respectively. The variable node $v_0 \in \mathcal{V}_{\mathcal{N}}$ is called the *root node*. The *depth* of a node in G is the distance of the node from the root node v_0. Variable nodes of depth t may have sockets which do not connect to any edges. With an abuse of notations, we will also write $\mathcal{V}(G) := \mathcal{V}_{\mathcal{N}}$ and $\mathcal{C}(G) := \mathcal{C}_{\mathcal{N}}$.

A *neighborhood graph ensemble* $\mathcal{N}_t(n, \lambda(x), \rho(x))$ induced by an $(n, \lambda(x), \rho(x))$-irregular ensemble is an ensemble of neighborhood graphs of depth t. Each neighborhood graph G is associated with the probability $\mathbb{P}_n(G)$ which is defined by the following steps. First, a Tanner graph $H = (\mathcal{V}, \mathcal{C}, \mathcal{S}_{\mathcal{V}}, \mathcal{S}_{\mathcal{C}}, \mathcal{E})$ is generated from the $(n, \lambda(x), \rho(x))$-irregular ensemble. A neighborhood graph $G = (\mathcal{V}_{\mathcal{N}}, \mathcal{C}_{\mathcal{N}}, v_0, \mathcal{S}_{\mathcal{V}_{\mathcal{N}}}, \mathcal{S}_{\mathcal{C}_{\mathcal{N}}}, \mathcal{E}_{\mathcal{N}})$ satisfying $\mathcal{V}_{\mathcal{N}} \subseteq \mathcal{V}$, $\mathcal{C}_{\mathcal{N}} \subseteq \mathcal{C}$, $\mathcal{S}_{\mathcal{V}_{\mathcal{N}}} \subseteq \mathcal{S}_{\mathcal{V}}$, $\mathcal{S}_{\mathcal{C}_{\mathcal{N}}} \subseteq \mathcal{S}_{\mathcal{C}}$ and $\mathcal{E}_{\mathcal{N}} \subseteq \mathcal{E}$ is generated from H by choosing the root node v_0 uniformly from $\mathcal{V}$ and the following rules. A variable node $v \in \mathcal{V}$ and a check node $c \in \mathcal{C}$ are members of $\mathcal{V}_{\mathcal{N}}$ and $\mathcal{C}_{\mathcal{N}}$ if and only if their distances from v_0 are not greater than t and $(t-1)$, respectively. Similarly, a variable-node socket $s \in \mathcal{S}_{\mathcal{V}}$ and a check-node socket $s' \in \mathcal{S}_{\mathcal{C}}$ are members of $\mathcal{S}_{\mathcal{V}_{\mathcal{N}}}$ and $\mathcal{S}_{\mathcal{C}_{\mathcal{N}}}$ if and only if s and s' are associated with nodes in $\mathcal{V}_{\mathcal{N}}$ and $\mathcal{C}_{\mathcal{N}}$, respectively. An edge $(s, s') \in \mathcal{E}$ is a member of $\mathcal{E}_{\mathcal{N}}$ if and only if $s \in \mathcal{S}_{\mathcal{V}_{\mathcal{N}}}$ and $s' \in \mathcal{S}_{\mathcal{C}_{\mathcal{N}}}$.

The random choice of the edge set $\mathcal{E}$ in the original irregular ensemble induces a probability distribution over the set of neighborhood graphs, under which each possible neighborhood graph G of depth t has a probability

$$\mathbb{Q}_n(G) = \frac{1}{nE(E-1)\cdots(E-(k-1))}$$

where $E = nL'(1)$ is the number of edges in the whole Tanner graph, as defined in Section II-B, and where k denotes the number of edges in the neighborhood graph G. The above probability $\mathbb{Q}_n(G)$ is obtained from the uniform choices of the root node and edge connections. For convenience, we will use a marginalized probability $\mathbb{P}_n(\cdot)$ which is induced from $\mathbb{Q}_n(\cdot)$ via the equivalence relation defined as follows: $G = (\mathcal{V}_{\mathcal{N}}, \mathcal{C}_{\mathcal{N}}, v_0, \mathcal{S}_{\mathcal{V}_{\mathcal{N}}}, \mathcal{S}_{\mathcal{C}_{\mathcal{N}}}, \mathcal{E}_{\mathcal{N}})$ and $G' = ({\mathcal{V}_{\mathcal{N}}}', {\mathcal{C}_{\mathcal{N}}}', v_0', \mathcal{S}_{\mathcal{V}'_{\mathcal{N}}}, \mathcal{S}_{\mathcal{C}'_{\mathcal{N}}}, {\mathcal{E}_{\mathcal{N}}}')$ are equivalent if and only if there exist bijections $\sigma_V : \mathcal{S}_{\mathcal{V}_{\mathcal{N}}} \to \mathcal{S}_{\mathcal{V}'_{\mathcal{N}}}$ and $\sigma_C : \mathcal{S}_{\mathcal{C}_{\mathcal{N}}} \to \mathcal{S}_{\mathcal{C}'_{\mathcal{N}}}$ such that

- c1. $(\sigma_V(s), \sigma_C(s')) \in {\mathcal{E}_{\mathcal{N}}}'$ for all $(s, s') \in \mathcal{E}_{\mathcal{N}}$
- c2. $\sigma_V(v_0) = v_0'$
- c3. $\forall v \in \mathcal{V}_{\mathcal{N}} \backslash v_0, \exists v' \in {\mathcal{V}_{\mathcal{N}}}' \backslash v_0'$ s.t. $\sigma_V(v) \overset{\mathrm{c}}{=} v'$
- c4. $\forall c \in \mathcal{C}_{\mathcal{N}}, \exists c' \in {\mathcal{C}_{\mathcal{N}}}'$ s.t. $\sigma_C(c) \overset{\mathrm{c}}{=} c'$

where $\sigma_V(v)$ (respectively $\sigma_C(c)$) are tuples whose i-th element is the image of the i-th element of v (respectively c) under σ_V (respectively σ_C), and where $u \overset{\mathrm{c}}{=} u'$ if and only if they are equal under a cyclic shift for tuples u and u' of sockets. This equivalence relation is weaker than what is used in $\mathbb{Q}_n(\cdot)$ and stronger than the conventional equivalence relation in graph theory which does not distinguish sockets.

Under this equivalence relation, the number of neighborhood graphs equivalent to G is

$$nL_{|v_0|} \left(\prod_i \prod_{a=0}^{n_i - 1} (nL_i - a) i \right) \left(\prod_j \prod_{b=0}^{m_j - 1} (mR_j - b) j \right)$$
$$= nL_{|v_0|} \left(\prod_i \prod_{a=0}^{n_i - 1} (E\lambda_i - ai) \right) \left(\prod_j \prod_{b=0}^{m_j - 1} (E\rho_j - bj) \right)$$

where n_i denotes the number of variable nodes of degree i in G, where m_j denotes the number of check nodes of degree

j in G, and where m denotes the number of check nodes in the whole Tanner graph, i.e., $m = nL'(1)/R'(1)$. Hence, the probability $\mathbb{P}_n(\cdot)$ which marginalizes equivalent neighborhood graphs is given as

$$\mathbb{P}_n(G) = L_{|v_0|} \frac{\prod_i \prod_{a=0}^{n_i-1}(E\lambda_i - ai) \prod_j \prod_{b=0}^{m_j-1}(E\rho_j - bj)}{\prod_{i=0}^{k-1}(E-i)}. \tag{3}$$

This defines the probability associated with a neighborhood graph G in the neighborhood graph ensemble.

Since $E = \Theta(n)$, the denominator and the numerator are $\Theta(n^k)$ and $\Theta(n^h)$, respectively, where h denotes the number of nodes in G except the root node. One therefore has $\mathbb{P}_n(G) = \Theta(n^{h-k})$. Here, $k-h$ is equal to the quantity called the circuit rank (also known as the cyclomatic number or the first Betti number) of G, which is the minimum number of edges to be removed from the graph G for obtaining a cycle-free graph. A graph is cycle-free or single-cycle if and only if $k-h=0$ or $k-h=1$, respectively.

Lemma 1. *For a neighborhood graph G with circuit rank C,*

$$\mathbb{P}_n(G) = \Theta(n^{-C}).$$

This lemma plays a key role in this paper. Classification of neighborhood graphs according to the circuit rank is also considered in [7].

D. Tree ensembles

From Lemma 1, neighborhood graphs of a fixed depth with cycles are not generated in the large-blocklength limit. To be precise,

$$\mathbb{P}_\infty(G) := \lim_{n\to\infty} \mathbb{P}_n(G) = L_{|v_0|} \prod_{v\in\mathcal{V}(G)\setminus v_0} \lambda_{|v|} \prod_{c\in\mathcal{C}(G)} \rho_{|c|} \tag{4}$$

for a tree graph G and $\mathbb{P}_\infty(G) = 0$ for any graph G with cycles. The ensemble of tree neighborhood graphs with probability $\mathbb{P}_\infty(G)$ is called the *tree ensemble from the node perspective*, and is denoted by $\mathring{\mathcal{T}}_t(\lambda(x),\rho(x))$.

We also define two other tree neighborhood graph ensembles $\vec{\mathcal{T}}_t^v(\lambda(x),\rho(x))$ and $\vec{\mathcal{T}}_t^c(\lambda(x),\rho(x))$ from the edge perspective. The tree neighborhood graphs from the edge perspective generated by $\vec{\mathcal{T}}_t^v(\lambda(x),\rho(x))$ and $\vec{\mathcal{T}}_t^c(\lambda(x),\rho(x))$ are rooted at an edge incident to a variable node and a check node, respectively. The number of check nodes in the path from the root edge to any node is not greater than t. Only variable nodes for which the path from the root edge includes t check nodes have sockets which do not connect to any edges. The probability of a neighborhood graph G rooted at an edge in both ensembles is

$$\prod_{v\in\mathcal{V}(G)} \lambda_{|v|} \prod_{c\in\mathcal{C}(G)} \rho_{|c|}.$$

The ensembles $\mathring{\mathcal{T}}_t(\lambda(x),\rho(x))$ and $\vec{\mathcal{T}}_t^v(\lambda(x),\rho(x))$ can be found in [14].

III. Main result

A. The decomposition of the coefficient of n^{-1}

In this subsection, it is explained how the coefficient $\alpha(\epsilon,t)$ of n^{-1} in the bit error probability is decomposed into two components, which are the contribution of cycle-free neighborhood graphs and the contribution of single-cycle neighborhood graphs. For each variable node, an error occurrence after t BP iterations depends only on the realization of the neighborhood graph G of depth t and realizations of channel outputs corresponding to variable nodes in G. Since the distribution of neighborhood graph obeys $\mathcal{N}_t(n,\lambda(x),\rho(x))$, and since the channel outputs are erased independently with probability ϵ, the bit error probability of the irregular ensemble is

$$\mathrm{P_b}(n,\epsilon,t) = \sum_{G\in\mathcal{G}_t} \mathbb{P}_n(G)\mathrm{P_b}(\epsilon,G) \tag{5}$$

where $\mathcal{G}_t$ denotes the set of all neighborhood graphs of depth t, and where $\mathrm{P_b}(\epsilon,G)$ denotes the erasure probability of the root node of G after t iterations when the erasure probability of each variable node in G is initialized with ϵ. From Lemma 1, it holds that

$$\mathrm{P_b}(\infty,\epsilon,t) = \lim_{n\to\infty}\mathrm{P_b}(n,\epsilon,t) = \sum_{G\in\mathcal{T}_t} \mathbb{P}_\infty(G)\mathrm{P_b}(\epsilon,G)$$

where $\mathcal{T}_t \subseteq \mathcal{G}_t$ denotes the set of all cycle-free neighborhood graphs of depth t. This fact allows us to calculate the limit of the bit error probability $\mathrm{P_b}(\infty,\epsilon,t)$ in a recursive manner, leading to the idea of density evolution.

Lemma 2 (Density evolution [5])**.** *For the limit of infinite blocklength, let $Q_\epsilon(t)$ denote the erasure probability of messages into check nodes at the t-th iteration, and let $P_\epsilon(t)$ denote the erasure probability of messages into variable nodes at the t-th iteration. Then*

$$\begin{aligned}
\mathrm{P_b}(\infty,\epsilon,t) &= \epsilon L(P_\epsilon(t)) \\
Q_\epsilon(t) &= \epsilon\lambda(P_\epsilon(t-1)) \\
P_\epsilon(t) &= \begin{cases} 1, & \textit{if } t=0 \\ 1-\rho(1-Q_\epsilon(t)), & \textit{otherwise.} \end{cases}
\end{aligned}$$

On the other hand, one observes from Lemma 1 that the second and the third dominant terms are $\Theta(n^{-1})$ and $\Theta(n^{-2})$, respectively. In other words, one has the following large-n asymptotic expansion of $\mathrm{P_b}(n,\epsilon,t)$:

$$\mathrm{P_b}(n,\epsilon,t) = \mathrm{P_b}(\infty,\epsilon,t) + \alpha(\epsilon,t)\frac{1}{n} + \Theta\left(\frac{1}{n^2}\right)$$

where the coefficient $\alpha(\epsilon,t)$ of n^{-1} is defined as

$$\alpha(\epsilon,t) := \lim_{n\to\infty} n(\mathrm{P_b}(n,\epsilon,t) - \mathrm{P_b}(\infty,\epsilon,t)).$$

Moreover, Lemma 1 tells us that $\alpha(\epsilon,t)$ can be decomposed into two components as follows:

$$\begin{aligned}
\alpha(\epsilon,t) = \lim_{n\to\infty} n &\left(\sum_{G\in\mathcal{T}_t}\mathbb{P}_n(G)\mathrm{P_b}(\epsilon,G) - \mathrm{P_b}(\infty,\epsilon,t)\right) \\
&\qquad + \lim_{n\to\infty} n \sum_{G\in\mathcal{U}_t}\mathbb{P}_n(G)\mathrm{P_b}(\epsilon,G) \\
=:&\ \beta(\epsilon,t) + \gamma(\epsilon,t)
\end{aligned}$$

where $\mathcal{U}_t \subseteq \mathcal{G}_t$ denotes the set of all single-cycle neighborhood graphs of depth t, and where the components $\beta(\epsilon, t)$ and $\gamma(\epsilon, t)$ represent contributions of cycle-free and single-cycle neighborhood graphs, respectively. In Section III-B and Section III-E, recursive formulas to evaluate $\beta(\epsilon, t)$ and $\gamma(\epsilon, t)$ for $(\lambda(x), \rho(x))$-irregular ensembles are derived, respectively.

B. The contribution of cycle-free neighborhood graphs

The contribution $\beta(\epsilon, t)$ of cycle-free neighborhood graphs is calculated as

$$\begin{aligned}\beta(\epsilon,t) &= \lim_{n\to\infty} n\left(\sum_{G\in\mathcal{T}_t}\mathbb{P}_n(G)\mathrm{P_b}(\epsilon,G) - \mathrm{P_b}(\infty,\epsilon,t)\right)\\ &= \sum_{G\in\mathcal{T}_t}\left[\lim_{n\to\infty} n\left(\mathbb{P}_n(G)-\mathbb{P}_\infty(G)\right)\right]\mathrm{P_b}(\epsilon,G).\end{aligned}$$

From (3) and (4), the contribution of a cycle-free neighborhood graph G to $\beta(\epsilon, t)$ is obtained as

$$\begin{aligned}&L_{|v_0|}\prod_{v\in\mathcal{V}(G)\setminus v_0}\lambda_{|v|}\prod_{c\in\mathcal{C}(G)}\rho_{|c|}\mathrm{P_b}(\epsilon,G)\lim_{n\to\infty}\\ &n\left(\frac{\prod_i\prod_{a=0}^{n_i-1}\left(E-a\frac{i}{\lambda_i}\right)\prod_j\prod_{b=0}^{m_j-1}\left(E-b\frac{j}{\rho_j}\right)}{\prod_{i=0}^{k-1}(E-i)}-1\right)\\ &= L_{|v_0|}\prod_{v\in\mathcal{V}(G)\setminus v_0}\lambda_{|v|}\prod_{c\in\mathcal{C}(G)}\rho_{|c|}\mathrm{P_b}(\epsilon,G)\frac{1}{2L'(1)}\\ &\times\left(k(k-1)-\sum_i\frac{i}{\lambda_i}n_i(n_i-1)-\sum_j\frac{j}{\rho_j}m_j(m_j-1)\right)\end{aligned}$$

where k, n_i and m_j are defined in Section II-C. Hence, $\beta(\epsilon, t)$ is obtained via expectation, denoted by $\mathbb{E}_t[\cdot]$, on the tree ensemble $\mathring{\mathcal{T}}_t(\lambda(x), \rho(x))$ of depth t from the node perspective as

$$\begin{aligned}\beta(\epsilon,t) = \frac{1}{2L'(1)}\Bigg[&\mathbb{E}_t[K(K-1)P] - \sum_i\frac{i}{\lambda_i}\mathbb{E}_t[N_i(N_i-1)P]\\ &-\sum_j\frac{j}{\rho_j}\mathbb{E}_t[M_j(M_j-1)P]\Bigg] \qquad (6)\end{aligned}$$

where K, N_i and M_j denote random variables representing the number of edges, the number of variable nodes of degree i and the number of check nodes of degree j, respectively, and where P denotes the erasure probability of the root node after t BP iterations.

The three expectations in (6) are obtained using generating functions as

$$\mathbb{E}_t[K(K-1)P] = \left.\frac{\partial^2\mathbb{E}_t[x^K P]}{\partial x^2}\right|_{x=1} \qquad (7)$$

$$\mathbb{E}_t[N_i(N_i-1)P] = \left.\frac{\partial^2\mathbb{E}_t[x^{N_i} P]}{\partial x^2}\right|_{x=1} \qquad (8)$$

$$\mathbb{E}_t[M_j(M_j-1)P] = \left.\frac{\partial^2\mathbb{E}_t[x^{M_j} P]}{\partial x^2}\right|_{x=1}. \qquad (9)$$

In order to deal with these generating functions, we now define the following "canonical" generating function:

$$\Phi(t;\,\{y_k\},\,\{z_l\}) = \mathbb{E}_t\left[\prod_k y_k^{N_k}\prod_l z_l^{M_l}P\right]. \qquad (10)$$

The three generating functions that appear on the right-hand sides of (7)–(9) are obtained from $\Phi(t;\,\{y_k\},\,\{z_l\})$ as

$$\begin{aligned}\mathbb{E}_t[x^K P] &= \frac{1}{x}\,\Phi(t;\,\{y_k\},\,\{z_l\})|_{y_k=x,z_l=x,\forall k,l}\\ \mathbb{E}_t[x^{N_i} P] &= \Phi(t;\,\{y_k\},\,\{z_l\})|_{y_i=x;\;y_k=1,\forall k\neq i;\;z_l=1,\forall l}\\ \mathbb{E}_t[x^{M_j} P] &= \Phi(t;\,\{y_k\},\,\{z_l\})|_{z_j=x;\;y_k=1,\forall k;\;z_l=1,\forall l\neq j}.\end{aligned}$$

The key idea here is that one can evaluate the canonical generating function $\Phi(t;\,\{y_k\},\,\{z_l\})$ via extending density evolution in such a way that "densities" to be updated in density evolution incorporate the auxiliary variables $\{y_k\}$ and $\{z_l\}$. We call our extension the augmented density evolution. In the conventional density evolution, expectation of density of messages over a tree ensemble is calculated in a recursive way. In the augmented density evolution, on the other hand, one considers, for each tree, a product of the density of messages at the root node and a monomial reflecting the degree histogram of the tree, and calculates its expectation over the tree ensemble, which can be performed recursively in a similar way to density evolution. The canonical generating function $\Phi(t;\,\{y_k\},\,\{z_l\})$ in the general case is thus a polynomial whose coefficients are conical combinations of densities. Since we are assuming $\mathrm{BEC}(\epsilon)$, we only have to deal with erasure probabilities of messages instead of densities of messages, as shown in Lemma 2. Hence, the canonical generating function $\Phi(t;\,\{y_k\},\,\{z_l\})$ is obtained by a recursive calculation of polynomials in $\{y_k\}$ and $\{z_l\}$ with real-valued coefficients. The next lemma provides a set of recursive formulas to evaluate the canonical generating function $\Phi(t;\,\{y_k\},\,\{z_l\})$.

Lemma 3. *The canonical generating function $\Phi(t;\,\{y_k\},\,\{z_l\})$ is given by*

$$\Phi(t;\,\{y_k\},\,\{z_l\}) = \epsilon\mathfrak{L}(F(t))$$

where

$$\begin{aligned}F(t) &:= \begin{cases}1, & \text{if } t=0\\ \mathcal{P}(g(t))-\mathcal{P}(G(t)), & \text{otherwise}\end{cases}\\ G(t) &:= \mathcal{L}(f(t-1))-\epsilon\mathcal{L}(F(t-1))\\ f(t) &:= \begin{cases}1, & \text{if } t=0\\ \mathcal{P}(g(t)), & \text{otherwise}\end{cases}\\ g(t) &:= \mathcal{L}(f(t-1))\end{aligned}$$

and where

$$\begin{aligned}\mathfrak{L}(x) &:= \sum_i L_i y_i x^i\\ \mathcal{L}(x) &:= \sum_i \lambda_i y_i x^{i-1}\\ \mathcal{P}(x) &:= \sum_j \rho_j z_j x^{j-1}.\end{aligned}$$

Proof: The generating function is calculated as

$$\begin{aligned}\mathbb{E}_t\left[\prod_k y_k^{N_k}\prod_l z_l^{M_l}P\right] &= \sum_i L_i y_i \epsilon\left(\mathbb{E}_t^c\left[\prod_k y_k^{M_k}\prod_l z_l^{M_l}Q\right]\right)^i \\ &= \epsilon\mathfrak{L}\left(\mathbb{E}_t^c\left[\prod_k y_k^{N_k}\prod_l z_l^{M_l}Q\right]\right)\end{aligned}$$

where $\mathbb{E}_t^c[\cdot]$ denotes expectation on $\vec{\mathcal{T}}_t^c$, and where Q denotes a random variable corresponding to the erasure probability of messages transmitted to the root edge at t-th iteration. Now define

$$\begin{aligned}f(t) &= \mathbb{E}_t^c\left[\prod_k y_k^{N_k}\prod_l z_l^{M_l}\right] \\ g(t) &= \mathbb{E}_{t-1}^v\left[\prod_k y_k^{N_k}\prod_l z_l^{M_l}\right] \\ F(t) &= \mathbb{E}_t^c\left[\prod_k y_k^{N_k}\prod_l z_l^{M_l}Q\right] \\ G(t) &= \mathbb{E}_{t-1}^v\left[\prod_k y_k^{N_k}\prod_l z_l^{M_l}(1-Q)\right]\end{aligned}$$

where $\mathbb{E}_t^v[\cdot]$ denotes expectation on $\vec{\mathcal{T}}_t^v$. The functions $f(t)$ and $g(t)$ are the generating functions of $\{N_k\}$ and $\{M_l\}$ on the ensembles $\vec{\mathcal{T}}_t^c$ and $\vec{\mathcal{T}}_{t-1}^v$, respectively. The functions $F(t)$ and $G(t)$ are reweighted versions of the generating functions, where reweighting is done on the basis of the erasure probability at the root node. It should be noted that dependence of these functions on the auxiliary variables $\{y_k\}$ and $\{z_l\}$ is implicit in the notation. The desired expectations are calculated recursively as

$$\begin{aligned}&f(0) = F(0) = 1 \\ &f(t) = \mathbb{E}_t^c\left[z_m g(t)^{m-1}\right] = \mathcal{P}(g(t)), \qquad \text{if } t \geq 1 \\ &g(t) = \mathbb{E}_{t-1}^v\left[y_m f(t-1)^{m-1}\right] = \mathcal{L}(f(t-1))\end{aligned}$$

$$\begin{aligned}F(t) &= \mathbb{E}_t^c\left[z_m\left(g(t)^{m-1} - G(t)^{m-1}\right)\right] \\ &= f(t) - \mathcal{P}(G(t)), \qquad \text{if } t \geq 1 \\ G(t) &= \mathbb{E}_{t-1}^v\left[y_m\left(f(t-1)^{m-1} - \epsilon F(t-1)^{m-1}\right)\right] \\ &= g(t) - \epsilon\mathcal{L}(F(t-1)).\end{aligned}$$

■

Considering appropriate derivatives of the recursive formulas given by Lemma 3, one obtains explicit formulas to evaluate the three expectations in (6) recursively, on the basis of which one can evaluate $\beta(\epsilon, t)$ explicitly. The derivation is elaborate but straightforward, so that we omit the details of the derivation and only show the end result. Let us define, for $d = 1$ and 2,

$$\begin{aligned}f^{(d)}(t) &:= \left.\frac{\partial^d f(t)|_{y_k=x, z_l=x \text{ for all } k,l}}{\partial x^d}\right|_{x=1} \\ f_v^{(d)}(t,i) &:= \left.\frac{\partial^d f(t)}{\partial y_i^d}\right|_{y_k=1,\, z_l=1 \text{ for all } k,l} \\ f_c^{(d)}(t,j) &:= \left.\frac{\partial^d f(t)}{\partial z_j^d}\right|_{y_k=1,\, z_l=1 \text{ for all } k,l}.\end{aligned}$$

Similar definitions are applied to $g(t)$, $F(t)$ and $G(t)$ to define $g^{(d)}(t)$, $g_v^{(d)}(t,i)$, etc. The resulting 24 functions are to be used to evaluate the relevant expectations in (6), and the recursive formulas of these functions used in the evaluation are summarized in the next theorem.

Theorem 1. *The value of $\beta(\epsilon, t)$ for $(\lambda(x), \rho(x))$-irregular ensembles is calculated as*

$$\begin{aligned}\beta(\epsilon,t) = \frac{1}{2L'(1)}\Bigg[&\mathbb{E}_t[K(K-1)P] \\ &- \sum_i \frac{i}{\lambda_i}\mathbb{E}_t[N_i(N_i-1)P] - \sum_j \frac{j}{\rho_j}\mathbb{E}_t[M_j(M_j-1)P]\Bigg]\end{aligned}$$

where $\mathbb{E}_t[K(K-1)P]$, $\mathbb{E}_t[N_i(N_i-1)P]$ and $\mathbb{E}_t[M_j(M_j-1)P]$ are calculated by (11), (12) and (13), respectively. The functions $P_\epsilon(t)$ and $Q_\epsilon(t)$ appearing in these formulas are to be evaluated recursively via the conventional density evolution (Lemma 2).

$$\begin{aligned}f^{(1)}(t) &= \begin{cases}0, & \text{if } t=0 \\ 1 + \rho'(1)g^{(1)}(t), & \text{otherwise}\end{cases} \\ g^{(1)}(t) &= 1 + \lambda'(1)f^{(1)}(t-1) \\ F^{(1)}(t) &= \begin{cases}0, & \text{if } t=0 \\ f^{(1)}(t) - \rho(1-Q_\epsilon(t)) & \\ \quad -\rho'(1-Q_\epsilon(t))G^{(1)}(t), & \text{otherwise}\end{cases} \\ G^{(1)}(t) &= g^{(1)}(t) - \epsilon\lambda(P_\epsilon(t-1)) \\ &\quad - \epsilon\lambda'(P_\epsilon(t-1))F^{(1)}(t-1)\end{aligned}$$

$$\begin{aligned}f^{(2)}(t) &= \begin{cases}0, & \text{if } t=0 \\ 2\rho'(1)g^{(1)}(t) & \\ \quad +\rho''(1)g^{(1)}(t)^2 + \rho'(1)g^{(2)}(t), & \text{otherwise}\end{cases} \\ g^{(2)}(t) &= 2\lambda'(1)f^{(1)}(t-1) + \lambda''(1)f^{(1)}(t-1)^2 \\ &\quad + \lambda'(1)f^{(2)}(t-1) \\ F^{(2)}(t) &= \begin{cases}0, & \text{if } t=0 \\ f^{(2)}(t) - 2\rho'(1-Q_\epsilon(t))G^{(1)}(t) & \\ \quad -\rho''(1-Q_\epsilon(t))G^{(1)}(t)^2 & \\ \quad -\rho'(1-Q_\epsilon(t))G^{(2)}(t), & \text{otherwise}\end{cases} \\ G^{(2)}(t) &= g^{(2)}(t) - 2\epsilon\lambda'(P_\epsilon(t-1))F^{(1)}(t-1) \\ &\quad - \epsilon\lambda''(P_\epsilon(t-1))F^{(1)}(t-1)^2 \\ &\quad - \epsilon\lambda'(P_\epsilon(t-1))F^{(2)}(t-1)\end{aligned}$$

$$\mathbb{E}_t[K(K-1)P] = \epsilon L''(P_\epsilon(t))F^{(1)}(t)^2 + \epsilon L'(P_\epsilon(t))F^{(2)}(t) \tag{11}$$

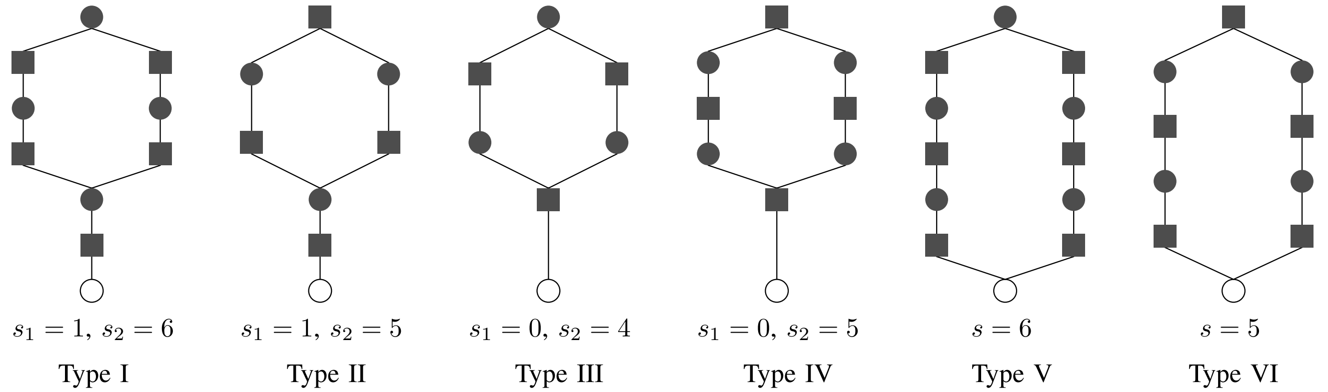

Fig. 1. Six types of single-cycle neighborhood graphs. All nodes which are not included in the two minimum path from the root node to the deepest node in the cycle are not shown in the above figure. The neighborhood graphs are classified according to whether the shallowest and the deepest nodes in the cycle are variable nodes, check nodes or the root node. The depth of the shallowest node in the cycle corresponds to s_1. The number of nodes in the shortest path from the root node to the deepest node in the cycle corresponds to $s_2 + 1$ and $s + 1$.

$$
\begin{aligned}
f_v^{(1)}(t,i) &= \begin{cases} 0, & \textit{if } t = 0 \\ \rho'(1) g_v^{(1)}(t,i), & \textit{otherwise} \end{cases} \\
g_v^{(1)}(t,i) &= \lambda'(1) f_v^{(1)}(t-1,i) + \lambda_i \\
F_v^{(1)}(t,i) &= \begin{cases} 0, & \textit{if } t = 0 \\ f_v^{(1)}(t,i) - \rho'(1 - Q_\epsilon(t)) G_v^{(1)}(t,i), & \textit{otherwise} \end{cases} \\
G_v^{(1)}(t,i) &= g_v^{(1)}(t,i) - \epsilon\lambda'(P_\epsilon(t-1)) F_v^{(1)}(t-1,i) \\
&\quad - \epsilon\lambda_i P_\epsilon(t-1)^{i-1}
\end{aligned}
$$

$$
\begin{aligned}
f_v^{(2)}(t,i) &= \begin{cases} 0, & \textit{if } t = 0 \\ \rho''(1) g_v^{(1)}(t,i)^2 + \rho'(1) g_v^{(2)}(t,i), & \textit{otherwise} \end{cases} \\
g_v^{(2)}(t,i) &= \lambda''(1) f_v^{(1)}(t-1,i)^2 + \lambda'(1) f_v^{(2)}(t-1,i) \\
&\quad + 2\lambda_i (i-1) f_v^{(1)}(t-1,i) \\
F_v^{(2)}(t,i) &= \begin{cases} 0, & \textit{if } t = 0 \\ f_v^{(2)}(t,i) - \rho''(1 - Q_\epsilon(t)) G_v^{(1)}(t,i)^2 & \\ \quad -\rho'(1 - Q_\epsilon(t)) G_v^{(2)}(t,i), & \textit{otherwise} \end{cases} \\
G_v^{(2)}(t,i) &= g_v^{(2)}(t,i) - \epsilon\lambda''(P_\epsilon(t-1)) F_v^{(1)}(t-1,i)^2 \\
&\quad - \epsilon\lambda'(P_\epsilon(t-1)) F_v^{(2)}(t-1,i) \\
&\quad - 2\epsilon\lambda_i (i-1) P_\epsilon(t-1)^{i-2} F_v^{(1)}(t-1,i)
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}_t[N_i(N_i - 1)P] &= \epsilon L''(P_\epsilon(t)) F_v^{(1)}(t,i)^2 \\
&\quad + \epsilon L'(P_\epsilon(t)) F_v^{(2)}(t,i) + 2\epsilon L_i i P_\epsilon(t)^{i-1} F_v^{(1)}(t,i)
\end{aligned} \tag{12}
$$

$$
\begin{aligned}
f_c^{(1)}(t,j) &= \begin{cases} 0, & \textit{if } t = 0 \\ \rho'(1) g_c^{(1)}(t,j) + \rho_j, & \textit{otherwise} \end{cases} \\
g_c^{(1)}(t,j) &= \lambda'(1) f_c^{(1)}(t-1,j) \\
F_c^{(1)}(t,j) &= \begin{cases} 0, & \textit{if } t = 0 \\ f_c^{(1)}(t,j) - \rho'(1 - Q_\epsilon(t)) G_c^{(1)}(t,j) & \\ \quad -\rho_j (1 - Q_\epsilon(t))^{j-1}, & \textit{otherwise} \end{cases} \\
G_c^{(1)}(t,j) &= g_c^{(1)}(t,j) - \epsilon\lambda'(P_\epsilon(t-1)) F_c^{(1)}(t-1,j)
\end{aligned}
$$

$$
\begin{aligned}
f_c^{(2)}(t,j) &= \begin{cases} 0, & \textit{if } t = 0 \\ \rho''(1) g_c^{(1)}(t,j)^2 + \rho'(1) g_c^{(2)}(t,j) & \\ \quad +2\rho_j (j-1) g_c^{(1)}(t,j), & \textit{otherwise} \end{cases} \\
g_c^{(2)}(t,j) &= \lambda''(1) f_c^{(1)}(t-1,j)^2 + \lambda'(1) f_c^{(2)}(t-1,j) \\
F_c^{(2)}(t,j) &= \begin{cases} 0, & \textit{if } t = 0 \\ f_c^{(2)}(t,j) - \rho''(1 - Q_\epsilon(t)) G_c^{(1)}(t,j)^2 & \\ \quad -\rho'(1 - Q_\epsilon(t)) G_c^{(2)}(t,j) & \\ \quad -2\rho_j (j-1)(1 - Q_\epsilon(t))^{j-2} & \\ \quad \times G_c^{(1)}(t,j), & \textit{otherwise} \end{cases} \\
G_c^{(2)}(t,j) &= g_c^{(2)}(t,j) - \epsilon\lambda''(P_\epsilon(t-1)) F_c^{(1)}(t-1,j)^2 \\
&\quad - \epsilon\lambda'(P_\epsilon(t-1)) F_c^{(2)}(t-1,j)
\end{aligned}
$$

$$
\mathbb{E}_t[M_j(M_j - 1)P] = \epsilon L''(P_\epsilon(t)) F_c^{(1)}(t,j)^2 + \epsilon L'(P_\epsilon(t)) F_c^{(2)}(t,j). \tag{13}
$$

C. Method of enumeration

In order to calculate the coefficient $\alpha(\epsilon, t)$ of n^{-1}, it is necessary to evaluate the contribution of single-cycle neighborhood graphs, i.e.,

$$
\gamma(\epsilon, t) := \lim_{n \to \infty} n \sum_{G \in \mathcal{U}_t} \mathbb{P}_n(G) \mathrm{P_b}(\epsilon, G).
$$

For ease of the explanation of how to evaluate $\gamma(\epsilon, t)$, which is deferred to Section III-E, we consider in this subsection a relevant quantity, namely the coefficient of n^{-1} in the probability of single-cycle neighborhood graphs:

$$
\xi(t) := \lim_{n \to \infty} n \sum_{G \in \mathcal{U}_t} \mathbb{P}_n(G).
$$

Methods for enumeration of $\xi(t)$ introduced in this subsection will be extended to those for calculation of $\gamma(\epsilon, t)$ in Section III-E. In both calculations, we consider the subgraph $S(G)$ of a single-cycle neighborhood graph G consisting of the nodes which are included by the two shortest paths from the root node to the deepest node in the cycle. We classify single-cycle neighborhood graphs into six types of subgraphs $S(G)$ as shown in Fig. 1. They are classified according to whether the shallowest node in the cycle is a non-root variable, a check, or the root node, as well as whether the deepest node in the cycle is a variable or check node. Types I to IV of the neighborhood graphs have two parameters: s_1 corresponding to the depth of the shallowest node in the cycle, and s_2 for which s_2+1 equals to the number of nodes (including the root node) in the shortest path from the root node to the deepest node in the cycle. Types V and VI of the neighborhood graphs have a parameter s which plays the same role as s_2 in Types I to IV. The set of single-cycle neighborhood graphs of Types I and II with the parameters s_1 and s_2 is denoted by $\mathcal{U}_v(t, s_1, s_2)$. The sets $\mathcal{U}_c(t, s_1, s_2)$ and $\mathcal{U}_r(t, s)$, corresponding to the neighborhood graphs of Types III and IV and to the neighborhood graphs of Types V and VI, respectively, are defined in a similar manner.

We consider marginalization of the probability using the classification of neighborhood graphs. The probability $\mathbb{P}_n(G)$ of a single-cycle neighborhood graph G is

$$L_{|v_0|}\frac{\prod_i \prod_{a=0}^{n_i-1}(\lambda_i E - ai)\prod_j \prod_{b=0}^{m_j-1}(\rho_j E - bj)}{\prod_{i=0}^{k-1}(E-i)}$$

where v_0 is the root node of G. Since $E = nL'(1)$, we obtain the coefficient of n^{-1} as

$$\lim_{n\to\infty} n\mathbb{P}_n(G) = \frac{1}{L'(1)} L_{|v_0|} \prod_{v\in\mathcal{V}(G)\setminus v_0} \lambda_{|v|} \prod_{c\in\mathcal{C}(G)} \rho_{|c|}.$$

In order to enumerate the coefficient of n^{-1} in the probability of single-cycle neighborhood graphs, we consider an equivalence relation in which positions of sockets connected by an edge to a socket associated with a node in $S(G)$ are not distinguished, which is weaker than what is used in defining $\mathbb{P}_n(\cdot)$. The sets of representatives of the resulting equivalence classes in $\mathcal{U}_v(t, s_1, s_2)$, $\mathcal{U}_c(t, s_1, s_2)$ and $\mathcal{U}_r(t, s)$ are denoted by $\bar{\mathcal{U}}_v(t, s_1, s_2)$, $\bar{\mathcal{U}}_c(t, s_1, s_2)$ and $\bar{\mathcal{U}}_r(t, s)$, respectively. The coefficients of n^{-1} in the probability of single-cycle neighborhood graphs of Types I and II with parameters s_1 and s_2 are evaluated in a unified way (s_2 is even for Type I and odd for Type II), and are obtained as

$$\begin{aligned}
&\lim_{n\to\infty} \sum_{G\in\mathcal{U}_v(t,s_1,s_2)} n\mathbb{P}_n(G) = \frac{1}{L'(1)} \sum_{G\in\bar{\mathcal{U}}_v(t,s_1,s_2)} \\
&L_{|v_0|}|v_0| \prod_{v\in\mathcal{V}(S(G))\setminus\{v_0,w\}} \lambda_{|v|}(|v|-1) \prod_{c\in\mathcal{C}(S(G))} \rho_{|c|}(|c|-1) \\
&\quad\times \lambda_{|w|}\binom{|w|-1}{2} \prod_{v\in\mathcal{V}(G)\setminus\mathcal{V}(S(G))} \lambda_{|v|} \prod_{c\in\mathcal{C}(G)\setminus\mathcal{C}(S(G))} \rho_{|c|} \\
&= \frac{1}{2}\lambda''(1)\rho'(1)^2(\lambda'(1)\rho'(1))^{s_2-s_1-2} \qquad (14)
\end{aligned}$$

where w denotes the shallowest variable node in the cycle. In the first equality in (14), single-cycle neighborhood graphs of Type I or II are marginalized according to the equivalence relation. In the second equality, by the marginalizations, quantities corresponding to nodes not included in $S(G)$ become 1, and quantities corresponding to the root node, the shallowest node in the cycle, other variable nodes in $S(G)$, and check nodes in $S(G)$ become $L'(1)$, $\lambda''(1)/2$, $\lambda'(1)$ and $\rho'(1)$, respectively. The concept of the equivalence classes $\bar{\mathcal{U}}_v(t, s_1, s_2)$, $\bar{\mathcal{U}}_c(t, s_1, s_2)$ and $\bar{\mathcal{U}}_r(t, s)$ is useful not only for the calculation (14) but also for the calculation of $\gamma(\epsilon, t)$ in Section III-E.

In the same way, the coefficients of n^{-1} in the probability of single-cycle neighborhood graphs of Types III and IV with parameters s_1 and s_2 are calculated as

$$\frac{1}{2}\rho''(1)\lambda'(1)(\lambda'(1)\rho'(1))^{s_2-s_1-2}$$

and those for Types V and VI with the parameter s are calculated as

$$\frac{1}{2}(\lambda'(1)\rho'(1))^s.$$

Similar calculations are also used in [7]. The classification of single-cycle neighborhood graphs in this subsection is finer than that in [7] for the purpose of calculation of $\gamma(\epsilon, t)$ in Section III-E. Summing up the above contributions of all types of single-cycle neighborhood graphs, we obtain

$$\begin{aligned}
\xi(t) &= \sum_{s_1=1}^{t-1} \sum_{s_2=2s_1+1}^{2t} \frac{1}{2}\lambda''(1)\rho'(1)^2(\lambda'(1)\rho'(1))^{s_2-s_1-2} \\
&\quad+ \sum_{s_1=0}^{t-1} \sum_{s_2=2s_1+2}^{2t} \frac{1}{2}\rho''(1)\lambda'(1)(\lambda'(1)\rho'(1))^{s_2-s_1-2} \\
&\quad+ \sum_{s=1}^{2t} \frac{1}{2}(\lambda'(1)\rho'(1))^s \\
&= \frac{1}{2}\Bigg[\lambda''(1)\rho'(1)^2\frac{(1-(\lambda'(1)\rho'(1))^{t-1})(1-(\lambda'(1)\rho'(1))^t)}{(1-\lambda'(1)\rho'(1))^2} \\
&\quad+ \rho''(1)\lambda'(1)\frac{(1-(\lambda'(1)\rho'(1))^t)^2}{(1-\lambda'(1)\rho'(1))^2} \\
&\quad+ \lambda'(1)\rho'(1)\frac{1-(\lambda'(1)\rho'(1))^{2t}}{1-\lambda'(1)\rho'(1)}\Bigg].
\end{aligned}$$

It should be noted that the above result can alternatively be obtained via the generating function method described in the previous subsection. Indeed, since the probability of all neighborhood graphs is exactly 1 and since the probability of neighborhood graphs which contain more than one cycle is $\Theta(n^{-2})$, the coefficient of n^{-1} in the probability of cycle-free neighborhood graphs is $-\xi(t)$, i.e., the probability of tree neighborhood graphs is $1-\xi(t)/n+\Theta(n^{-2})$. Hence, the above result for the quantity $\xi(t)$ is also obtained by the enumeration of the coefficient $-\beta(1,t)$ of n^{-1} in the probability of cycle-free neighborhood graphs using the generating function method in the previous subsection.

D. Single-cycle neighborhood graph ensembles

Single-cycle neighborhood graph ensembles are defined in this subsection in order to make the description of the calculation of $\gamma(\epsilon, t)$ in the next subsection more tractable. A single-cycle neighborhood graph ensemble for an arbitrary fixed type and parameters is defined not in terms of single-cycle neighborhood graphs themselves but in terms of representatives of their equivalence classes, with the specified type and parameters. The single-cycle neighborhood graph ensembles are defined by normalizing (14). The probability, to be defined in this subsection, of a representative, denoted as G by a slight abuse of notation, can be considered as the large blocklength limit of the conditional probability, measured by the neighborhood graph ensemble, of the single-cycle neighborhood graphs in the equivalence class represented by G conditioned on that a single-cycle neighborhood graph has a particular type and parameters. The probability of a representative G of an equivalence class of single-cycle neighborhood graphs in $\bar{\mathcal{U}}_v(t,s_1,s_2)$ is

$$\begin{aligned}\mathbb{P}_v^{(t,s_1,s_2)}(G) :=& \frac{L_{|v_0|}|v_0|}{L'(1)}\frac{\lambda_{|w|}(|w|-1)(|w|-2)}{\lambda''(1)}\\ &\times \prod_{v\in\mathcal{V}(S(G))\setminus\{v_0,w\}}\frac{\lambda_{|v|}(|v|-1)}{\lambda'(1)}\prod_{c\in\mathcal{C}(S(G))}\frac{\rho_{|c|}(|c|-1)}{\rho'(1)}\\ &\times \prod_{v\in\mathcal{V}(G)\setminus\mathcal{V}(S(G))}\lambda_{|v|}\prod_{c\in\mathcal{C}(G)\setminus\mathcal{C}(S(G))}\rho_{|c|}\end{aligned} \tag{15}$$

where w denotes the shallowest variable node in the cycle. Similarly, the probability of a representative G of an equivalence class of single-cycle neighborhood graphs in $\bar{\mathcal{U}}_c(t,s_1,s_2)$ is

$$\begin{aligned}\mathbb{P}_c^{(t,s_1,s_2)}(G) :=& \frac{L_{|v_0|}|v_0|}{L'(1)}\frac{\rho_{|y|}(|y|-1)(|y|-2)}{\rho''(1)}\\ &\times \prod_{v\in\mathcal{V}(S(G))\setminus v_0}\frac{\lambda_{|v|}(|v|-1)}{\lambda'(1)}\prod_{c\in\mathcal{C}(S(G))\setminus y}\frac{\rho_{|c|}(|c|-1)}{\rho'(1)}\\ &\times \prod_{v\in\mathcal{V}(G)\setminus\mathcal{V}(S(G))}\lambda_{|v|}\prod_{c\in\mathcal{C}(G)\setminus\mathcal{C}(S(G))}\rho_{|c|}\end{aligned}$$

where y denotes the shallowest check node in the cycle, and the probability of $G\in\bar{\mathcal{U}}_r(t,s)$ is

$$\begin{aligned}\mathbb{P}_r^{(t,s)}(G) :=& \frac{L_{|v_0|}|v_0|(|v_0|-1)}{L''(1)}\prod_{v\in\mathcal{V}(G)\setminus\mathcal{V}(S(G))}\lambda_{|v|}\\ &\times \prod_{c\in\mathcal{C}(G)\setminus\mathcal{C}(S(G))}\rho_{|c|}\prod_{v\in S(G)\setminus v_0}\frac{\lambda_{|v|}(|v|-1)}{\lambda'(1)}\prod_{c\in S(G)}\frac{\rho_{|c|}(|c|-1)}{\rho'(1)}.\end{aligned}$$

These ensembles are used in Section III-E for calculating $\gamma(\epsilon,t)$.

E. The contribution of single-cycle neighborhood graphs

The contribution $\gamma(\epsilon,t)$ of single-cycle neighborhood graphs can be decomposed according to the types and parameters of single-cycle neighborhood graphs.

$$\begin{aligned}\gamma(\epsilon,t) =& \sum_{G\in\mathcal{U}_t}(\lim_{n\to\infty} n\mathbb{P}_n(G))\mathrm{P_b}(\epsilon,G)\\ =& \sum_{s_1=1}^{t-1}\sum_{s_2=2s_1+1}^{2t}\sum_{G\in\mathcal{U}_v(t,s_1,s_2)}(\lim_{n\to\infty} n\mathbb{P}_n(G))\mathrm{P_b}(\epsilon,G)\\ &+\sum_{s_1=0}^{t-1}\sum_{s_2=2s_1+2}^{2t}\sum_{G\in\mathcal{U}_c(t,s_1,s_2)}(\lim_{n\to\infty} n\mathbb{P}_n(G))\mathrm{P_b}(\epsilon,G)\\ &+\sum_{s=1}^{2t}\sum_{G\in\mathcal{U}_r(t,s)}(\lim_{n\to\infty} n\mathbb{P}_n(G))\mathrm{P_b}(\epsilon,G)\\ =:& \sum_{s_1=1}^{t-1}\sum_{s_2=2s_1+1}^{2t}F_v(t,s_1,s_2)\end{aligned}$$

$$f(t,s,p) := \begin{cases}\epsilon, & \text{if } t=0\\ \epsilon\frac{\lambda'(P_\epsilon(t))}{\lambda'(1)}g(t,s-1,p), & \text{otherwise}\end{cases},\qquad g(t,s,p) := \begin{cases}p, & \text{if } s=0\\ 1-\frac{\rho'(1-Q_\epsilon(t))}{\rho'(1)}(1-f(t-1,s,p)), & \text{otherwise}\end{cases}$$

$$G_1(t,s) := \begin{cases}1, & \text{if } s=0\\ \left(1-\frac{\rho'(1-Q_\epsilon(t))}{\rho'(1)}\right)^2+2\frac{\rho'(1-Q_\epsilon(t))}{\rho'(1)}\left(1-\frac{\rho'(1-Q_\epsilon(t))}{\rho'(1)}\right)f(t-1,s,1) & \\ \qquad +\left(\frac{\rho'(1-Q_\epsilon(t))}{\rho'(1)}\right)^2 G_2(t-1,s-1), & \text{otherwise}\end{cases}$$

$$G_2(t,s) := \begin{cases}\epsilon\frac{\lambda'(P_\epsilon(t))}{\lambda'(1)}, & \text{if } s=0\\ \left(\epsilon\frac{\lambda'(P_\epsilon(t))}{\lambda'(1)}\right)^2 G_1(t,s-1), & \text{otherwise}\end{cases}$$

$$G_3(t,s) := \begin{cases}1-\epsilon\frac{\lambda'(P_\epsilon(t))}{\lambda'(1)}, & \text{if } s=0\\ 1-2f(t,s+1,1)+\left(\epsilon\frac{\lambda'(P_\epsilon(t))}{\lambda'(1)}\right)^2 G_1(t,s-1), & \text{otherwise}\end{cases}$$

$$\begin{aligned}F_v(t,s_1,s_2) =& \frac{1}{2}\lambda''(1)\rho'(1)^2(\lambda'(1)\rho'(1))^{s_2-s_1-2}Q_\epsilon(t+1)\\ &\times g\left(t,s_1-1,1-\frac{\rho'(1-Q_\epsilon(t-s_1+1))}{\rho'(1)}\left(1-\epsilon\frac{\lambda''(P_\epsilon(t-s_1))}{\lambda''(1)}G_1(t-s_1,s_2-2s_1-1)\right)\right)\end{aligned} \tag{16}$$

$$F_c(t,s_1,s_2) = \frac{1}{2}\rho''(1)\lambda'(1)(\lambda'(1)\rho'(1))^{s_2-s_1-2}Q_\epsilon(t+1)g\left(t,s_1,1-\frac{\rho''(1-Q_\epsilon(t-s_1))}{\rho''(1)}G_3(t-s_1-1,s_2-2s_1-2)\right) \tag{17}$$

$$F_r(t,s) = \frac{1}{2}(\lambda'(1)\rho'(1))^s\epsilon\frac{\lambda'(P_\epsilon(t))}{\lambda'(1)}G_1(t,s-1) \tag{18}$$

$$+\sum_{s_1=0}^{t-1}\sum_{s_2=2s_1+2}^{2t}F_c(t,s_1,s_2)+\sum_{s=1}^{2t}F_r(t,s)$$

where $F_v(t,s_1,s_2)$, $F_c(t,s_1,s_2)$ and $F_r(t,s)$ are the contributions of single-cycle neighborhood graphs in $\mathcal{U}_v(t,s_1,s_2)$, $\mathcal{U}_c(t,s_1,s_2)$ and $\mathcal{U}_r(t,s)$, respectively. A set of formulas for calculating these quantities are shown in the next theorem.

Theorem 2. *The value of $\gamma(\epsilon,t)$ for $(\lambda(x),\rho(x))$-irregular ensembles is calculated as*

$$\gamma(\epsilon,t)=\sum_{s_1=1}^{t-1}\sum_{s_2=2s_1+1}^{2t}F_v(t,s_1,s_2)+\sum_{s_1=0}^{t-1}\sum_{s_2=2s_1+2}^{2t}F_c(t,s_1,s_2)+\sum_{s=1}^{2t}F_r(t,s)$$

where $F_v(t,s_1,s_2)$, $F_c(t,s_1,s_2)$ and $F_r(t,s)$ are shown in (16), (17) and (18), respectively. If $\lambda''(1)=0$, $F_v(t,s_1,s_2)$ is defined as 0.

Proof: A derivation of $F_v(t,s_1,s_2)$ is described in the following. Similarly to (14), the contribution $F_v(t,s_1,s_2)$ of neighborhood graphs in $\mathcal{U}_v(t,s_1,s_2)$ to $\gamma(\epsilon,t)$ is obtained as

$$\begin{aligned}
F_v(t,s_1,s_2)&=\sum_{G\in\mathcal{U}_v(t,s_1,s_2)}\left(\lim_{n\to\infty}n\mathbb{P}_n(G)\right)\mathrm{P_b}(\epsilon,G)\\
&=\frac{1}{L'(1)}\sum_{G\in\bar{\mathcal{U}}_v(t,s_1,s_2)}L_{|v_0|}|v_0|\\
&\quad\times\prod_{v\in\mathcal{V}(S(G))\setminus\{v_0,w\}}\lambda_{|v|}(|v|-1)\prod_{c\in\mathcal{C}(S(G))}\rho_{|c|}(|c|-1)\\
&\quad\times\lambda_{|w|}\binom{|w|-1}{2}\prod_{v\in\mathcal{V}(G)\setminus\mathcal{V}(S(G))}\lambda_{|v|}\prod_{c\in\mathcal{C}(G)\setminus\mathcal{C}(S(G))}\rho_{|c|}\\
&\quad\times\mathrm{P_b}(\epsilon,G)\\
&=\frac{1}{2}\lambda''(1)\rho'(1)^2(\lambda'(1)\rho'(1))^{s_2-s_1-2}\\
&\quad\times\sum_{G\in\bar{\mathcal{U}}_v(t,s_1,s_2)}\frac{L_{|v_0|}|v_0|}{L'(1)}\prod_{v\in\mathcal{V}(S(G))\setminus\{v_0,w\}}\frac{\lambda_{|v|}(|v|-1)}{\lambda'(1)}\\
&\quad\times\left(\prod_{c\in\mathcal{C}(S(G))}\frac{\rho_{|c|}(|c|-1)}{\rho'(1)}\right)\frac{\lambda_{|w|}(|w|-1)(|w|-2)}{\lambda''(1)}\\
&\quad\times\prod_{v\in\mathcal{V}(G)\setminus\mathcal{V}(S(G))}\lambda_{|v|}\prod_{c\in\mathcal{C}(G)\setminus\mathcal{C}(S(G))}\rho_{|c|}\mathrm{P_b}(\epsilon,G)\\
&=\frac{1}{2}\lambda''(1)\rho'(1)^2(\lambda'(1)\rho'(1))^{s_2-s_1-2}\\
&\quad\times\sum_{G\in\bar{\mathcal{U}}_v(t,s_1,s_2)}\mathbb{P}_v^{(t,s_1,s_2)}(G)\mathrm{P_b}(\epsilon,G)
\end{aligned}$$

Hence, we have to calculate the expected error probability over the single-cycle neighborhood graph ensemble. Marginalizing the non-cycle part of $S(G)$ and trees incident to them, if any, we obtain

$$\begin{aligned}
F_v(t,s_1,s_2)&=\frac{1}{2}\lambda''(1)\rho'(1)^2(\lambda'(1)\rho'(1))^{s_2-s_1-2}\\
&\times\sum_{Y,Z}\prod_{v\in\mathcal{V}(Y)\setminus w}\frac{\lambda_{|v|}(|v|-1)}{\lambda'(1)}\prod_{c\in\mathcal{C}(Y)}\frac{\rho_{|c|}(|c|-1)}{\rho'(1)}\\
&\times\frac{\lambda_{|w|}(|w|-1)(|w|-2)}{\lambda''(1)}\prod_{v\in\mathcal{V}(Z)}\lambda_{|v|}\prod_{c\in\mathcal{C}(Z)}\rho_{|c|}\\
&\times\epsilon\frac{L'(P_\epsilon(t))}{L'(1)}\left(1-\frac{\rho'(1-Q_\epsilon(t))}{\rho'(1)}\left(1-\epsilon\frac{\lambda'(P_\epsilon(t-1))}{\rho'(1)}\right.\right.\\
&\left.\left.\cdots(1-p)\right)\right)
\end{aligned}\tag{19}$$

where Y denotes the subgraph which consists of nodes in the cycle, where Z denotes trees incident to Y, and where p denotes the erasure probability of the message from w to the shallow check node connected to w. In the summation $\sum_{Y,Z}$, all subgraphs consisting of w and its descendants are summed. The calculation of the non-cycle part in (19) is similar to the derivation of density evolution in Lemma 2. Equation (19) is calculated as

$$\begin{aligned}
&\frac{1}{2}\lambda''(1)\rho'(1)^2(\lambda'(1)\rho'(1))^{s_2-s_1-2}Q_\epsilon(t+1)\\
&\quad\times g\left(t,s_1-1,1-\frac{\rho'(1-Q_\epsilon(t-s_1+1))}{\rho'(1)}\left(1\right.\right.\\
&-\sum_{Y,Z}\prod_{v\in\mathcal{V}(Y)\setminus w}\frac{\lambda_{|v|}(|v|-1)}{\lambda'(1)}\prod_{c\in\mathcal{C}(Y)}\frac{\rho_{|c|}(|c|-1)}{\rho'(1)}\\
&\quad\left.\left.\times\frac{\lambda_{|w|}(|w|-1)(|w|-2)}{\lambda''(1)}\prod_{v\in\mathcal{V}(Z)}\lambda_{|v|}\prod_{c\in\mathcal{C}(Z)}\rho_{|c|}p\right)\right).
\end{aligned}$$

Hence, if one can prove the equality

$$\begin{aligned}
&\sum_{Y,Z}\prod_{v\in\mathcal{V}(Y)\setminus w}\frac{\lambda_{|v|}(|v|-1)}{\lambda'(1)}\prod_{c\in\mathcal{C}(Y)}\frac{\rho_{|c|}(|c|-1)}{\rho'(1)}\\
&\quad\times\frac{\lambda_{|w|}(|w|-1)(|w|-2)}{\lambda''(1)}\prod_{v\in\mathcal{V}(Z)}\lambda_{|v|}\prod_{c\in\mathcal{C}(Z)}\rho_{|c|}p\\
&\qquad=\epsilon\frac{\lambda''(P_\epsilon(t-s_1))}{\lambda''(1)}G_1(t-s_1,s_2-2s_1-1)
\end{aligned}\tag{20}$$

then (16) will immediately be obtained.

Now we prove (20). First, marginalizing w and trees incident to w, denoted by Z_w, the left-hand side of (20) is calculated as

$$\begin{aligned}
&\sum_{Y\setminus w,Z\setminus Z_w}\prod_{v\in\mathcal{V}(Y\setminus w)}\frac{\lambda_{|v|}(|v|-1)}{\lambda'(1)}\prod_{c\in\mathcal{C}(Y\setminus w)}\frac{\rho_{|c|}(|c|-1)}{\rho'(1)}\\
&\quad\times\prod_{v\in\mathcal{V}(Z\setminus Z_w)}\lambda_{|v|}\prod_{c\in\mathcal{C}(Z\setminus Z_w)}\rho_{|c|}\epsilon\frac{\lambda''(P_\epsilon(t-s_1))}{\lambda''(1)}q
\end{aligned}\tag{21}$$

where q denotes the probability that two messages into w from the check nodes connected to w in the cycle are both erased. Let c_1 and c_2 denote the check nodes in the cycle incident to w. If c_1 and c_2 are the same, i.e., if $s_2-2s_1-1=0$ holds, then $q=1$. Otherwise, q is decomposed into four components as

$$\begin{aligned}
q=P(e_1,e_2)&=P(A_1,A_2,e_1,e_2)+P(\bar{A}_1,A_2,e_1,e_2)\\
&\quad+P(A_1,\bar{A}_2,e_1,e_2)+P(\bar{A}_1,\bar{A}_2,e_1,e_2)\\
&=P(A_1)P(A_2)+P(e_1\mid\bar{A}_1,A_2)P(\bar{A}_1)P(A_2)\\
&\quad+P(e_2\mid A_1,\bar{A}_2)P(A_1)P(\bar{A}_2)+P(\bar{A}_1,\bar{A}_2,e_1,e_2)
\end{aligned}\tag{22}$$

where e_1 and e_2 denote the events that the messages from c_1 and c_2 to w are erased, respectively, and where A_1 and A_2 denote the events that at least one message from outside the cycle into c_1 and c_2 is erased, respectively. Calculating the marginalization in (21), the first term in (22) becomes

$$\left(1 - \frac{\rho'(1 - Q_\epsilon(t - s_1))}{\rho'(1)}\right)^2. \tag{23}$$

Each of the second and third terms becomes

$$\frac{\rho'(1 - Q_\epsilon(t - s_1))}{\rho'(1)}\left(1 - \frac{\rho'(1 - Q_\epsilon(t - s_1))}{\rho'(1)}\right) \times f(t - s_1 - 1, s_2 - 2s_1 - 1, 1). \tag{24}$$

At last, the fourth term becomes

$$\left(\frac{\rho'(1 - Q_\epsilon(t - s_1))}{\rho'(1)}\right)^2 r \tag{25}$$

where r denotes the probability that both of the messages to c_1 and c_2 from variable nodes v_1 and v_2 are erased. Here, v_1 (v_2) is a variable node in the cycle, is not w, and connects to c_1 (c_2), respectively. If v_1 and v_2 are the same, i.e., if $s_2 - 2s_2 - 1 = 1$ holds, then $r = \epsilon\lambda'(P_\epsilon(t - s_1 - 1))/\lambda'(1)$. Otherwise,

$$r = \left(\epsilon\frac{\lambda'(P_\epsilon(t - s_1 - 1))}{\lambda'(1)}\right)^2 q'$$

where q' denotes the probability that both of the messages to v_1 and v_2 from check nodes c_3 and c_4 are erased. Here, c_3 (c_4) is a check node in the cycle, is not c_1 (c_2) and connects to v_1 (v_2), respectively. The probability q' is obtained in the same way as q. Summing (23), (24) and (25), we obtain $G_1(t-s_1, s_2-2s_1-1)$. Hence, we obtain (20) and the proof showing that the contribution of neighborhood graphs of Types I and II with the parameters s_1 and s_2 is $F_v(t, s_1, s_2)$ is done.

In almost the same way, the contributions of neighborhood graphs of Types III and IV are obtained as

$$\frac{1}{2}\rho''(1)\lambda'(1)(\lambda'(1)\rho'(1))^{s_2-s_1-2}Q_\epsilon(t+1) \times g\Bigg(t, s_1, 1 - \frac{\rho''(1 - Q_\epsilon(t - s_1))}{\rho''(1)} \times G_3(t - s_1 - 1, s_2 - 2s_1 - 2)\Bigg)$$

and those of Types V and VI are obtained as

$$\frac{1}{2}(\lambda'(1)\rho'(1))^s\epsilon\frac{\lambda'(P_\epsilon(t))}{\lambda'(1)}G_1(t, s - 1).$$

Since the derivation is similar, the proof is omitted. ∎

IV. The limit of $\alpha(\epsilon, t)$

In this section, the limit value $\alpha(\epsilon,\infty) := \lim_{t\to\infty}\alpha(\epsilon,t)$ is shown for the (l,r)-regular ensembles, which is defined as the $(\lambda(x) = x^{l-1}, \rho(x) = x^{r-1})$-irregular ensemble. The limit $\alpha(\epsilon,\infty)$ has a simple expression while the expression of $\alpha(\epsilon,t)$ is complicated and recursive. Empirically, the approximation using $\alpha(\epsilon,\infty)$ instead of $\alpha(\epsilon,t)$ is accurate even for small blocklength if ϵ is close to 0 or 1, as will be observed in Section VI. The proof of the following theorem is in Appendix A.

Theorem 3. *For the (l,r)-regular ensembles, let*

$$\begin{aligned}
P_\epsilon(\infty) &:= \lim_{t\to\infty} P_\epsilon(t)\\
Q_\epsilon(\infty) &:= \lim_{t\to\infty} Q_\epsilon(t)\\
\mathsf{p} &:= \epsilon(l-1)P_\epsilon(\infty)^{l-2}\\
\mathsf{q} &:= (r-1)(1 - Q_\epsilon(\infty))^{r-2}\\
\mathsf{v} &:= \epsilon(l-1)(l-2)P_\epsilon(\infty)^{l-3}\\
\mathsf{w} &:= (r-1)(r-2)(1 - Q_\epsilon(\infty))^{r-3}
\end{aligned}$$

where $P_\epsilon(t)$ and $Q_\epsilon(t)$ are defined in Lemma 2. If $\mathsf{pq} < 1$, the limit is

$$\begin{aligned}
\alpha(\epsilon,\infty) = &\frac{1}{2}\frac{1}{1-\mathsf{pq}}\left(\mathsf{pq} + Q_\epsilon(\infty)\frac{1}{1-\mathsf{pq}}\mathsf{q}^2\mathsf{v}\right)\\
&\times\left[\frac{1}{1-\mathsf{pq}}(P_\epsilon(\infty) - Q_\epsilon(\infty)) + 1 - P_\epsilon(\infty)Q_\epsilon(\infty)\right]\\
&+ \frac{1}{2}Q_\epsilon(\infty)\frac{1}{(1-\mathsf{pq})^2}\mathsf{wp}\\
&\times\left[\frac{1}{1-\mathsf{pq}}(Q_\epsilon(\infty) - P_\epsilon(\infty)) + (1 - P_\epsilon(\infty))(1 - Q_\epsilon(\infty))\right].
\end{aligned}$$

The quantity pq which appears in the condition of the theorem is the slope of the function of density evolution $f_{\text{de}}(x) = \epsilon\lambda(1 - \rho(1 - x))$, which described the evolution of $Q_\epsilon(t)$ in Lemma 2, at the largest fixed point $x = Q_\epsilon(\infty) \in [0,1]$. Hence, $\mathsf{pq} \le 1$ is always satisfied. The condition $\mathsf{pq} = 1$ holds if and only if $y = f_{\text{de}}(x)$ touches $y = x$ at the largest fixed point. Such points of ϵ include the threshold ϵ_{BP} and the points where the largest fixed point changes discontinuously with respect to ϵ.

Especially below the threshold, i.e., $\epsilon < \epsilon_{\text{BP}}$, one obtains

$$\alpha(\epsilon,\infty) = \frac{1}{2}\frac{\epsilon\lambda'(0)\rho'(1)}{1 - \epsilon\lambda'(0)\rho'(1)}.$$

This quantity also appears in (1). This fact implies that the following two limits are equal below the threshold for regular ensembles

$$\lim_{t\to\infty}\lim_{n\to\infty} n(\mathrm{P_b}(n,\epsilon,t) - \mathrm{P_b}(\infty,\epsilon,t)) = \lim_{n\to\infty}\lim_{t\to\infty} n(\mathrm{P_b}(n,\epsilon,t) - \mathrm{P_b}(\infty,\epsilon,t)).$$

Here, the right-hand side is obtained in [11] for any irregular ensemble below the threshold while the left-hand side $\alpha(\epsilon,\infty)$ for regular ensembles is obtained in Theorem 3 for any erasure probability satisfying $\mathsf{pq} < 1$.

The limit $\alpha(\epsilon,\infty)$ for irregular ensembles is an open problem.

V. Uniform convergence under fixed number of iterations

As mentioned in the introduction, the bit error probability after an infinite number of iterations converges to a discontinuous curve with respect to the erasure probability of a channel as the blocklength tends to infinity if $\lambda'(0)\rho'(1)\epsilon_{\text{BP}} < 1$ i.e., if the BP threshold is not given by the stability condition [18].

Since the bit error probability for a finite blocklength is continuous, the convergence is not uniform. Due to the lack of uniform convergence, an approximation using the asymptotic expansion (1) with respect to blocklength is not accurate near the discontinuous points. Hence, for an accurate approximation near discontinuous points, other approximations should be considered. The scaling-law-based approximation method was introduced by Amraoui et al. [13], [17] for this purpose.

In this section, we will show that the bit error probability after a fixed number of iterations converges to a limit uniformly in contrast to the case of an infinite number of iterations, which immediately implies that the approximation (2) is accurate for all ϵ uniformly when the blocklength is sufficiently large. We have to show

$$|\mathrm{P_b}(n,\epsilon,t) - \mathrm{P_b}(\infty,\epsilon,t)| \le \delta \tag{26}$$

where $\delta = o(1)$ as $n \to \infty$ and δ does not depend on ϵ. The left-hand side of (26) is bounded as

$$\begin{aligned}
&|\mathrm{P_b}(n,\epsilon,t) - \mathrm{P_b}(\infty,\epsilon,t)| \\
=&\Bigg| \sum_{G\in\mathcal{T}_t} \mathbb{P}_n(G)\mathrm{P_b}(\epsilon,G) + \sum_{G\in\mathcal{U}_t} \mathbb{P}_n(G)\mathrm{P_b}(\epsilon,G) \\
&+ \sum_{G\in\mathcal{G}_t\setminus(\mathcal{T}_t\cup\mathcal{U}_t)} \mathbb{P}_n(G)\mathrm{P_b}(\epsilon,G) - \mathrm{P_b}(\infty,\epsilon,t)\Bigg| \\
\le&\left| \sum_{G\in\mathcal{T}_t} \mathbb{P}_n(G)\mathrm{P_b}(\epsilon,G) - \mathrm{P_b}(\infty,\epsilon,t) - \beta(\epsilon,t)\frac{1}{n}\right| \\
&+ \left| \sum_{G\in\mathcal{U}_t} \mathbb{P}_n(G)\mathrm{P_b}(\epsilon,G) - \gamma(\epsilon,t)\frac{1}{n}\right| \\
&+ |\alpha(\epsilon,t)|\frac{1}{n} + \sum_{G\in\mathcal{G}_t\setminus(\mathcal{T}_t\cup\mathcal{U}_t)} \mathbb{P}_n(G).
\end{aligned} \tag{27}$$

From Lemma 1, the last term on the rightmost side of (27), which depends on t but not on ϵ, is $\Theta(n^{-2})$. The first term on the rightmost side of (27) is bounded as

$$\begin{aligned}
&\left| \sum_{G\in\mathcal{T}_t} \mathbb{P}_n(G)\mathrm{P_b}(\epsilon,G) - \mathrm{P_b}(\infty,\epsilon,t) - \beta(\epsilon,t)\frac{1}{n}\right| \\
=&\Bigg| \sum_{G\in\mathcal{T}_t} \bigg(\mathbb{P}_n(G) - \mathbb{P}_\infty(G) \\
&- \frac{1}{n}\left[\lim_{n'\to\infty} n'\left(\mathbb{P}_{n'}(G) - \mathbb{P}_\infty(G)\right)\right]\bigg)\mathrm{P_b}(\epsilon,G)\Bigg| \\
\le& \sum_{G\in\mathcal{T}_t} \left| \mathbb{P}_n(G) - \mathbb{P}_\infty(G) - \frac{1}{n}\left[\lim_{n'\to\infty} n'\left(\mathbb{P}_{n'}(G) - \mathbb{P}_\infty(G)\right)\right]\right|.
\end{aligned}$$

Similarly, the second term on the rightmost side of (27) is also bounded as

$$\begin{aligned}
&\left| \sum_{G\in\mathcal{U}_t} \mathbb{P}_n(G)\mathrm{P_b}(\epsilon,G) - \gamma(\epsilon,t)\frac{1}{n}\right| \\
=&\left| \sum_{G\in\mathcal{U}_t} \left(\mathbb{P}_n(G) - \frac{1}{n}\left[\lim_{n'\to\infty} n'\mathbb{P}_{n'}(G)\right]\right)\mathrm{P_b}(\epsilon,G)\right| \\
\le& \sum_{G\in\mathcal{U}_t} \left| \mathbb{P}_n(G) - \frac{1}{n}\left[\lim_{n'\to\infty} n'\mathbb{P}_{n'}(G)\right]\right|.
\end{aligned}$$

The above two bounds are $\Theta(n^{-2})$ and are independent of ϵ. Hence, (27) is upper bounded by

$$|\alpha(\epsilon,t)|\frac{1}{n} + D$$

where $D = \Theta(n^{-2})$ depends on t but not on ϵ.

Since $|\alpha(\epsilon,t)|$ is continuous on $\epsilon \in [0,1]$ and so bounded, we conclude that the bit error probability under a finite number of iterations converges to the limit uniformly as the blocklength tends to infinity. More accurately, we obtain

$$\left|\mathrm{P_b}(n,\epsilon,t) - \mathrm{P_b}(\infty,\epsilon,t) - \alpha(\epsilon,t)\frac{1}{n}\right| \le D \tag{28}$$

from the above discussion. Equation (28) bounds the error of the approximation. However, this bound is available only under the assumption that the blocklength is sufficiently large so that all possible neighborhood graphs could be generated. In the next section, we observe via numerical calculations and simulations that the approximation is also accurate even if the assumption is not satisfied.

VI. Numerical calculations and simulations

In this section, we show calculation results of $\alpha(\epsilon,t)$ and $\alpha(\epsilon,\infty)$ and show simulation results of $n|\mathrm{P_b}(n,\epsilon,t) - \mathrm{P_b}(\infty,\epsilon,t)|$, the quantity which tends to $|\alpha(\epsilon,t)|$ as n tends to infinity.

The results of calculating $\alpha(\epsilon,t)$ for the $(2,3)$-regular ensemble, the $(3,6)$-regular ensemble and an irregular ensemble are shown in Fig. 2, Fig. 3 and Fig. 4, respectively. The coefficient $\alpha(\epsilon,t)$ seems to approach the limit $\alpha(\epsilon,\infty)$ quickly if ϵ is close to 0 or 1. For the $(3,6)$-regular ensemble, $\max_{\epsilon\in[0,\epsilon_{\mathrm{BP}})}\alpha(\epsilon,t)$ seems to diverge so that the convergence $\lim_{t\to\infty}\alpha(\epsilon,t)$ seems not to be uniform.

If ϵ satisfies the two conditions $\lambda'(0)\rho'(1)\epsilon < (\lambda'(1)\rho'(1))^{-1}$ and $\epsilon < \epsilon_{\mathrm{BP}}$, then $\beta(\epsilon,t)$ tends to zero and $\gamma(\epsilon,t)$ tends to the limit $\alpha(\epsilon,\infty)$ as t tends to infinity. In this case, we can understand intuitively that the dominant events of decoding error are events of errors of channel outputs in single-cycle neighborhood graphs consisting of variable nodes of degree 2 and check nodes, i.e., single-cycle stopping sets equivalent to single-cycle codewords, which were also discussed in [7]. However, if $\lambda'(0)\rho'(1)\epsilon > (\lambda'(1)\rho'(1))^{-1}$, which is the case when $\epsilon > 0.25$ in Fig. 2 and when $\epsilon > 0.113\,48$ in Fig. 4, even below the threshold, $\beta(\epsilon,t)$ grows to $-\infty$ and $\gamma(\epsilon,t)$ grows to $+\infty$ exponentially in t. The reason of this large cancellation between $\beta(\epsilon,t)$ and $\gamma(\epsilon,t)$ is not sufficiently understood.

Because of the large cancellation, multiprecision arithmetic was necessary in our calculations to avoid cancellation errors in computation of $\alpha(\epsilon,t)$ with large t.

Simulation results for the above ensembles are shown in Fig. 5, Fig. 6 and Fig. 7, respectively. For the $(2,3)$-regular ensemble, the simulation results with $n = 801$ almost converge to $\alpha(\epsilon,t)$ for all ϵ, as shown in Fig. 5. This is also the case with the irregular ensemble which has variable nodes of

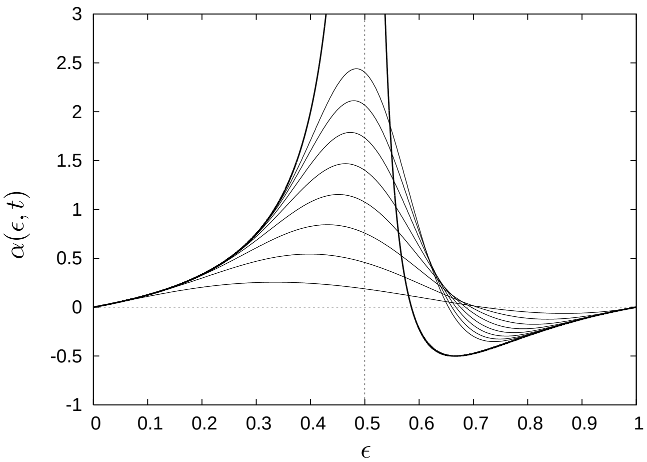

Fig. 2. Calculation results for the $(2,3)$-regular ensemble. Thin curves show $\alpha(\epsilon, t)$ for $t = 1, 2, \ldots, 8$. Thick curve shows the limit $\alpha(\epsilon, \infty)$. The threshold ϵ_{BP} is 0.5.

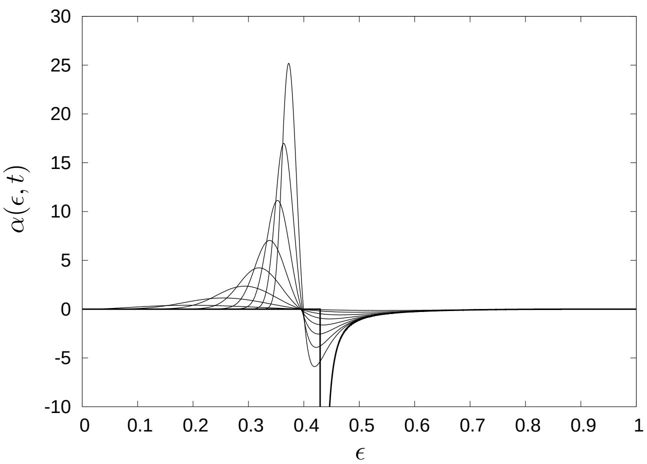

Fig. 3. Calculation results for the $(3,6)$-regular ensemble. Thin curves show $\alpha(\epsilon, t)$ for $t = 1, 2, \ldots, 8$. Thick curve shows the limit $\alpha(\epsilon, \infty)$. The threshold ϵ_{BP} is about 0.42944.

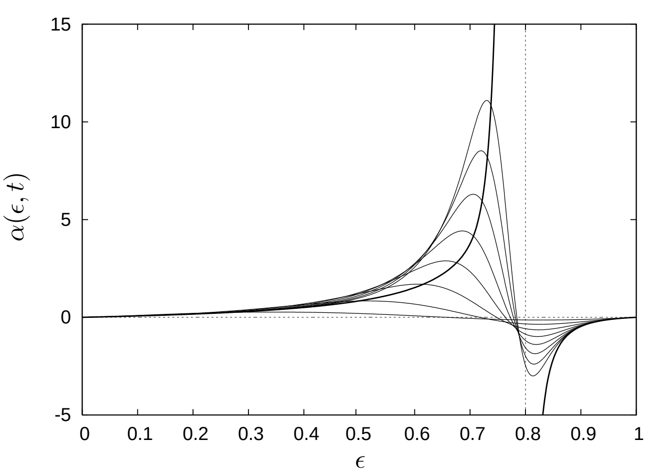

Fig. 4. Calculation results for an irregular ensemble. $\lambda(x) = 0.500x + 0.153x^2 + 0.112x^3 + 0.055x^4 + 0.180x^8$, $\rho(x) = 0.492x^2 + 0.508x^3$. Thin curves show $\alpha(\epsilon, t)$ for $t = 1, 2, \ldots, 8$. Thick curve shows the result for $t = 50$. The threshold ϵ_{BP} is about 0.8.

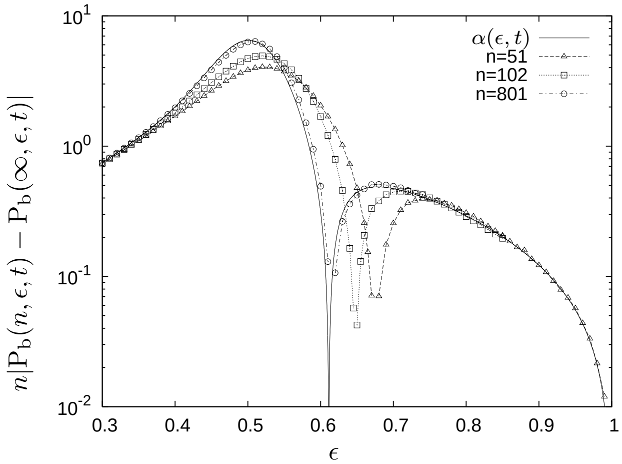

Fig. 5. Simulation results for the $(2,3)$-regular ensemble. Blocklengths are 51, 102 and 801. Number of iterations is 20.

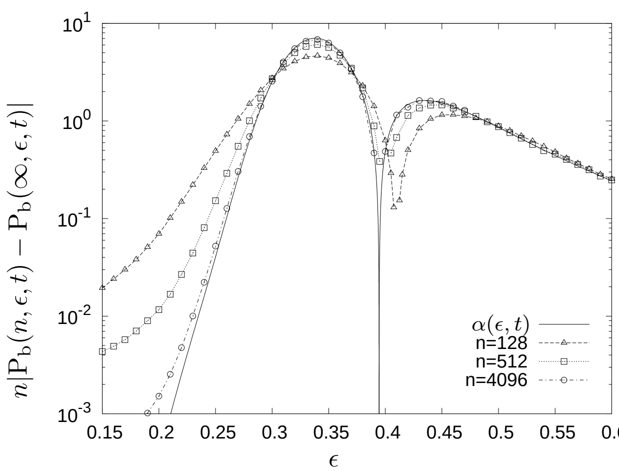

Fig. 6. Simulation results for the $(3,6)$-regular ensemble. Blocklengths are 128, 512 and 4096. Number of iterations is 5.

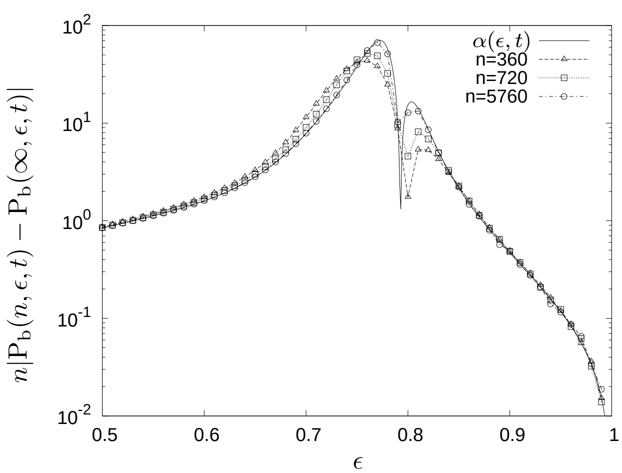

Fig. 7. Simulation results for an irregular ensemble. $\lambda(x) = 0.500x + 0.153x^2 + 0.112x^3 + 0.055x^4 + 0.180x^8$, $\rho(x) = 0.492x^2 + 0.508x^3$. Blocklengths are 360, 720 and 5760. Number of iterations is 20.

degree 2 (Fig. 7), where the simulation results with $n = 5760$ are observed to converge well to $\alpha(\epsilon, t)$ for all ϵ. For the $(3, 6)$-regular ensemble (Fig. 6), the simulation results almost converge to $\alpha(\epsilon, t)$ for $\epsilon > 0.25$ at $n = 4096$. The agreement between simulation results and theoretical results is rather surprising, since the pairs of the blocklength and the number of iterations are not suitable for density evolution technique in which one assumes that neighborhood graphs are trees with high probability. Indeed, the number of variable nodes in tree graphs is well above the total number of variable nodes in each of the three cases considered: The number of variable nodes in a tree graph of depth 20 in the $(2, 3)$-regular ensemble is $4\,194\,302$ which is much greater than the blocklength 801. The number of variable nodes in a tree graph of depth 5 in the $(3, 6)$-regular ensemble is $166\,666$ which is much greater than the blocklength 4096. The number of variable nodes in the minimum tree graph of depth 20 in the irregular ensemble is $4\,194\,302$ which is much greater than the blocklength 5760. We have not succeeded in finding an appropriate explanation to the observed quick convergence.

For the $(3, 6)$-regular ensemble, the convergence to $\alpha(\epsilon, t)$ is not fast for $\epsilon < 0.25$. In the low-ϵ region, the dominant error events after an infinite number of iterations are those induced by small stopping sets. The $(3, 6)$-regular ensemble does not contain single-cycle stopping sets but contains three double-cycle stopping sets. When ϵ is close to 0, unless the blocklength is sufficiently large, the bit error probability after a small number of iterations is almost the same as that after an infinite number of iterations, since decoding will succeed after a few number of iterations with high probability. This is also the case when ϵ is close to 1, in which case decoding will fail after a few number of iterations with high probability. Hence, in the low-ϵ region, the bit error probability decays like $\Theta(n^{-2})$ rather than $\Theta(n^{-1})$ unless the blocklength is sufficiently large.

The well-established fact that the bit error probability at an error floor is well approximated by (1) [7] is interpreted as the statement that the bit error probability $\mathrm{P_b}(n, \epsilon, t)$ when ϵ is close to 0 and $\lambda'(0) > 0$ is well approximated by $\mathrm{P_b}(\infty, \epsilon, t) + \alpha(\epsilon, \infty)/n$ for large n. From the observed quick convergence of $\alpha(\epsilon, t)$ to $\alpha(\epsilon, \infty)$ and that of $n(\mathrm{P_b}(n, \epsilon, t) - \mathrm{P_b}(\infty, \epsilon, t))$ to $\alpha(\epsilon, t)$ for ϵ close to 1, the same statement is empirically valid when ϵ is close to 1 as well.

VII. Conclusion

We have obtained the coefficient $\alpha(\epsilon, t)$ of the second dominant term in the asymptotic expansion of the bit error probability after a fixed number of iterations for irregular ensembles. Furthermore, we have obtained the limit $\alpha(\epsilon, \infty)$ for regular ensembles. At last, we have confirmed that approximations using $\alpha(\epsilon, t)$ are accurate even for small blocklength.

There are two important open problems. The first one is the large cancellation problem between $\beta(\epsilon, t)$ and $\gamma(\epsilon, t)$. The underlying mechanism of this cancellation has not been understood sufficiently, so that, for example, we do not know whether similar cancellations occur in higher-order terms. The second one is the fast convergence problem of $\alpha(\epsilon, t)$. Simulation results show that the convergence to $\alpha(\epsilon, t)$ is very fast. This fact is surprising since neighborhood graphs must include many cycles in moderate blocklengths.

Some other open problems need to be addressed as well. First, the limit $\alpha(\epsilon, \infty)$ for the irregular ensemble has not been derived. Second, optimization of finite-length irregular and expurgated ensembles given the number of iterations, blocklength, erasure probability, and allowable error probability, similar to the finite-blocklength optimization by Amraoui et al. [13], [17] for an infinite number of iterations, is practically important. Third, derivation of the coefficients of higher-order terms $n^{-2}, n^{-3}, \dots$ is an interesting problem. Fourth, other limits may also be important in practice. An example is the limit of the blocklength and the number of iterations tending to infinity simultaneously. Assume $t = c \log n / \log(\lambda'(1)\rho'(1))$ for some constant $c > 0$. Then the probability of cycle-free neighborhood graphs tends to 1 for $c < 1/2$ and tends to 0 for $c > 1/2$ [19]. It means that the cycle-free assumption is applicable only for $c < 1/2$, so that methods like density evolution under $c > 1/2$ are not available. At last, generalization of the methods to general channels and BP or other message passing decoders is important. There is a technically difficult problem due to the reuse of messages from the same edges for calculation of the contributions of single-cycle neighborhood graphs.

Acknowledgment

The authors acknowledge Dr. Pascal Vontobel, the associate editor, for many detailed comments.

Appendix A
Proof of Theorem 3

First, we show an alternative expression of $\alpha(\epsilon, t)$ for regular ensembles. The new expression is useful for proving Theorem 3. Furthermore, the expression does not require multiprecision arithmetic, which the previous expression requires in order to avoid the cancellation errors in the calculation of the sum $\beta(\epsilon, t) + \gamma(\epsilon, t)$.

Lemma 4. *For the (l, r)-regular ensemble, $\alpha(\epsilon, t)$ is calculated as*

$$\alpha(\epsilon, t) = \sum_{s_1=1}^{t-1} \sum_{s_2=2s_1+1}^{2t} T_v(t, s_1, s_2) + \sum_{s_1=0}^{t-1} \sum_{s_2=2s_1+2}^{2t} T_c(t, s_1, s_2) + \sum_{s=1}^{2t} T_r(t, s)$$

where

$$T_v(t, s_1, s_2) := \frac{1}{2} Q_\epsilon(t+1) \rho'(1 - Q_\epsilon(t)) \epsilon \lambda''(P_\epsilon(t - s_1)) \times \left(\prod_{k=1}^{s_1-1} \epsilon \lambda'(P_\epsilon(t-k)) \rho'(1 - Q_\epsilon(t-k)) \right) \times H_1(t - s_1, s_2 - 2s_1 - 1)$$

$$T_c(t,s_1,s_2) := \frac{1}{2}Q_\epsilon(t+1)\rho''(1-Q_\epsilon(t-s_1)) \times \left(\prod_{k=0}^{s_1-1}\epsilon\lambda'(P_\epsilon(t-k-1))\rho'(1-Q_\epsilon(t-k))\right) \times H_3(t-s_1-1, s_2-2s_1-2)$$

$$T_r(t,s) := \frac{1}{2}\epsilon\lambda'(P_\epsilon(t))H_1(t,s-1)$$

$$H_1(t,s) := \begin{cases} \rho'(1)(1-P_\epsilon(t)^2), & \textit{if } s=0 \\ \rho'(1-Q_\epsilon(t))^2 H_2(t-1,s-1), & \textit{if } s \ge t \\ 2(\rho'(1)-\rho'(1-Q_\epsilon(t)))(1-P_\epsilon(t-s)) & \\ \times \prod_{k=0}^{s-1}\rho'(1-Q_\epsilon(t-k))\epsilon\lambda'(P_\epsilon(t-k-1)) & \\ +\rho'(1-Q_\epsilon(t))^2 H_2(t-1,s-1), & \textit{otherwise} \end{cases}$$

$$H_2(t,s) := \begin{cases} \epsilon\lambda'(P_\epsilon(t)) - \lambda'(1)Q_\epsilon(t+1)^2, & \textit{if } s=0 \\ (\epsilon\lambda'(P_\epsilon(t)))^2 H_1(t,s-1), & \end{cases}$$

$$H_3(t,s) := \begin{cases} \epsilon\lambda'(P_\epsilon(t)) - \lambda'(1)Q_\epsilon(t+1)(2-Q_\epsilon(t+1)), & \textit{if } s=0 \\ -(\epsilon\lambda'(P_\epsilon(t)))^2 H_1(t,s-1), & \textit{if } s \ge t \\ 2\epsilon\lambda'(P_\epsilon(t))(1-P_\epsilon(t-s)) & \\ \times \prod_{k=0}^{s-1}\rho'(1-Q_\epsilon(t-k))\epsilon\lambda'(P_\epsilon(t-k-1)) & \\ -(\epsilon\lambda'(P_\epsilon(t)))^2 H_1(t,s-1), & \textit{otherwise.} \end{cases}$$

and where $\lambda(x) = x^{l-1}$ *and* $\rho(x) = x^{r-1}$.

Outline of proof of Lemma 4: For the (l,r)-regular ensemble, the cycle-free neighborhood graph is unique. The coefficient of n^{-1} in the probability of the unique cycle-free neighborhood graph is

$$-\frac{1}{2}l(r-1)\frac{1-\{(l-1)(r-1)\}^t}{1-(l-1)(r-1)}\{(l-1)(r-1)\}^t.$$

Hence, $\beta(\epsilon,t)$ for the (l,r)-regular ensemble is obtained as

$$\beta(\epsilon,t) = -\frac{1}{2}l(r-1)\frac{1-\{(l-1)(r-1)\}^t}{1-(l-1)(r-1)}\{(l-1)(r-1)\}^t \times \epsilon P_\epsilon(t)^l.$$

It is decomposed as follows.

$$\beta(\epsilon,t) = -\epsilon P_\epsilon(t)^l \times \frac{1}{2}\Bigg[\sum_{s_1=1}^{t-1}\sum_{2s_1+1}^{2t}\lambda''(1)\rho'(1)(\lambda'(1)\rho'(1))^{s_2-s_1-2} + \sum_{s_1=0}^{t-1}\sum_{s_2=2s_1+2}^{2t}\rho''(1)\lambda'(1)(\lambda'(1)\rho'(1))^{s_2-s_1-2} + \sum_{s=1}^{2t}(\lambda'(1)\rho'(1))^s\Bigg].$$

Hence, $\alpha(\epsilon,t)$ is calculated as

$$\begin{aligned}\alpha(\epsilon,t) &= \sum_{s_1=1}^{t-1}\sum_{2s_1+1}^{2t} \left(F_v(t,s_1,s_2) - \frac{1}{2}\lambda''(1)\rho'(1)(\lambda'(1)\rho'(1))^{s_2-s_1-2}\epsilon P_\epsilon(t)^l\right) \\ &+ \sum_{s_1=0}^{t-1}\sum_{s_2=2s_1+2}^{2t}\left(F_c(t,s_1,s_2) - \frac{1}{2}\rho''(1)\lambda'(1)(\lambda'(1)\rho'(1))^{s_2-s_1-2}\epsilon P_\epsilon(t)^l\right) \\ &+ \sum_{s=1}^{2t}\left(F_r(t,s) - \frac{1}{2}(\lambda'(1)\rho'(1))^s\epsilon P_\epsilon(t)^l\right) \\ &= \sum_{s_1=1}^{t-1}\sum_{s_2=2s_1+1}^{2t}T_v(t,s_1,s_2) + \sum_{s_1=0}^{t-1}\sum_{s_2=2s_1+2}^{2t}T_c(t,s_1,s_2) \\ &+ \sum_{s=1}^{2t}T_r(t,s).\end{aligned}$$

We omit calculations of $T_v(t,s_1,s_2)$, $T_c(t,s_1,s_2)$ and $T_r(t,s)$. ■

Proof of Theorem 3: After some calculations, we obtain

$$\begin{aligned}&\lim_{t\to\infty}\sum_{s_1=1}^{t-1}\sum_{s_2=2s_1+1}^{2t}\lim_{t'\to\infty}T_v(t',s_1,s_2) \\ &= \frac{1}{2}Q_\epsilon(\infty)\frac{1}{(1-\mathsf{pq})^2}\mathsf{q}^2\mathsf{v} \\ &\times\left[\frac{1}{1-\mathsf{pq}}(P_\epsilon(\infty)-Q_\epsilon(\infty)) + 1 - P_\epsilon(\infty)Q_\epsilon(\infty)\right]\end{aligned}$$

$$\begin{aligned}&\lim_{t\to\infty}\sum_{s_1=0}^{t-1}\sum_{s_2=2s_1+2}^{2t}\lim_{t'\to\infty}T_c(t',s_1,s_2) \\ &= \frac{1}{2}Q_\epsilon(\infty)\frac{1}{(1-\mathsf{pq})^2}\mathsf{wp} \\ &\times\left[\frac{1}{1-\mathsf{pq}}(Q_\epsilon(\infty)-P_\epsilon(\infty)) + (1-P_\epsilon(\infty))(1-Q_\epsilon(\infty))\right]\end{aligned}$$

$$\begin{aligned}&\lim_{t\to\infty}\sum_{s=1}^{2t}\lim_{t'\to\infty}T_r(t',s) = \frac{1}{2}\frac{1}{1-\mathsf{pq}}\mathsf{pq} \\ &\times\left[\frac{1}{1-\mathsf{pq}}(P_\epsilon(\infty)-Q_\epsilon(\infty)) + 1 - P_\epsilon(\infty)Q_\epsilon(\infty)\right].\end{aligned}$$

If there exist $\bar{T}_v(s_1,s_2)$, $\bar{T}_c(s_1,s_2)$ and $\bar{T}_r(s)$ such that

$$\begin{aligned}|T_v(t,s_1,s_2)| &\le \bar{T}_v(s_1,s_2), && \text{for all } t \\ |T_c(t,s_1,s_2)| &\le \bar{T}_c(s_1,s_2), && \text{for all } t \\ |T_r(t,s)| &\le \bar{T}_r(s), && \text{for all } t\end{aligned}$$

and such that

$$\lim_{t\to\infty}\sum_{s_1=1}^{t-1}\sum_{s_2=2s_1+1}^{2t}\bar{T}_v(s_1,s_2)<\infty$$
$$\lim_{t\to\infty}\sum_{s_1=0}^{t-1}\sum_{s_2=2s_1+2}^{2t}\bar{T}_c(s_1,s_2)<\infty$$
$$\lim_{t\to\infty}\sum_{s=1}^{2t}\bar{T}_r(s)<\infty$$

then Theorem 3 is a consequence of Lebesgue's dominated convergence theorem. If $\epsilon\lambda'(P_\epsilon(\infty))\rho'(1-Q_\epsilon(\infty))<1$, there exists $\delta>0$ such that

$$\epsilon(\lambda'(P_\epsilon(\infty))+\delta)(\rho'(1-Q_\epsilon(\infty))+\delta)<1.$$

On the other hand,

$$|\lambda'(P_\epsilon(t))-\lambda'(P_\epsilon(\infty))|<\delta \tag{29}$$
$$|\rho'(1-Q_\epsilon(t))-\rho'(1-Q_\epsilon(\infty))|<\delta \tag{30}$$

for all but finite t. One can therefore take $\bar{T}_v(s_1,s_2)$, $\bar{T}_c(s_1,s_2)$ and $\bar{T}_r(s)$ satisfying the above conditions by replacing $\lambda'(P_\epsilon(t))$ and $\rho'(1-Q_\epsilon(t))$ in $T_v(t,s_1,s_2)$, $T_c(t,s_1,s_2)$ and $T_r(t,s)$ with $\lambda'(P_\epsilon(\infty))+\delta$ and $\rho'(1-Q_\epsilon(\infty))+\delta$, respectively, and multiplying them with an appropriate constant in order to take into account the fact that a finite number of $\lambda'(P_\epsilon(t))$ and $\rho'(1-Q_\epsilon(t))$ in $T_v(t,s_1,s_2)$, $T_c(t,s_1,s_2)$ and $T_r(t,s)$ do not satisfy (29) and (30). ■

REFERENCES

[1] R. Mori, K. Kasai, T. Shibuya, and K. Sakaniwa, "Asymptotic bit error probability of LDPC codes for the binary erasure channel with finite number of iterations," in *Proc. 2008 IEEE Int. Symp. Inf. Theory, Toronto, Canada*, Jul. 6–11, 2008, pp. 449–453.

[2] ——, "Asymptotic gaps between BP decoding and local-MAP decoding for low-density parity-check codes," in *Proc. 2008 IEEE 5th Int. Symp. on Turbo Codes and Related Topics, Lausanne, Switzerland*, Sep. 1–5, 2008, pp. 162–167.

[3] R. Mori, K. Kasai, T. Tanaka, and K. Sakaniwa, "Finite-length analysis of irregular expurgated LDPC codes under finite number of iterations," in *Proc. 2009 IEEE Int. Symp. Inf. Theory, Seoul, South Korea*, Jun. 28–Jul. 3, 2009, pp. 2497–2501.

[4] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, D. A. Spielman, and V. Stemann, "Practical loss-resilient codes," in *STOC '97: Proceedings of the twenty-ninth annual ACM Symposium on Theory of Computing*. New York, NY, USA: ACM, 1997, pp. 150–159.

[5] T. Richardson and R. 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.

[6] T. Richardson, "Error floors of LDPC codes," in *41st Allerton Conference on Communications, Control and Computing, Monticello*, Oct. 2003, pp. 1426–1435.

[7] A. Montanari, "The asymptotic error floor of LDPC ensembles under BP decoding," in *44th Allerton Conference on Communications, Control and Computing, Monticello*, Oct. 2006, pp. 1168–1172.

[8] C. Méasson, A. Montanari, T. Richardson, and R. Urbanke, "The generalized area theorem and some of its consequences," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 4793–4821, Nov. 2009.

[9] J. Ezri, A. Montanari, and R. Urbanke, "A generalization of the finite-length scaling approach beyond the BEC," in *Proc. 2007 IEEE Int. Symposium on Inform. Theory, Nice, France*, Jun. 24–29, 2007, pp. 1011–1015.

[10] C. Di, D. Proietti, I. Telatar, T. Richardson, and R. Urbanke, "Finite-length analysis of low-density parity-check codes on the binary erasure channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1570–1579, Jun. 2002.

[11] C. Di, T. Richardson, and R. Urbanke, "Weight distribution of low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 4839–4855, Nov. 2006.

[12] C. Méasson, A. Montanari, and R. Urbanke, "Maxwell construction: The hidden bridge between iterative and maximum a posteriori decoding," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5277–5307, Dec. 2008.

[13] A. Amraoui, A. Montanari, T. Richardson, and R. Urbanke, "Finite-length scaling for iteratively decoded LDPC ensembles," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 473–498, Feb. 2009.

[14] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge University Press, 2008.

[15] J. Zhang and A. Orlitsky, "Finite-length analysis of LDPC codes with large left degrees," in *Proc. 2002 IEEE Int. Symp. Inf. Theory, Lausanne, Switzerland*, Jun. 30–Jul. 5, 2002, p. 3.

[16] T. Richardson and R. Urbanke. (2003) Finite-length density evolution and the distribution of the number of iterations for the binary erasure channel. unpublished.

[17] A. Amraoui, A. Montanari, and R. Urbanke, "How to find good finite-length codes: from art towards science," *European Transactions on Telecommunications*, vol. 18, no. 5, pp. 491–508, 2007.

[18] T. Richardson, M. Shokrollahi, and R. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[19] M. Mézard and A. Montanari, *Information, Physics and Computation*. Oxford University Press, 2009.

Ryuhei Mori received the B.E. degree from Tokyo Institute of Technology, Tokyo, Japan in 2008, and the M.Inf. degree from Kyoto University, Kyoto, Japan in 2010. Since 2010, he has been a Ph.D. student in Kyoto University. His research interests include information theory, coding theory and statistical physics.

Toshiyuki Tanaka received the B.E., M.E., and D.E. degrees in electronics engineering from the University of Tokyo, Tokyo, Japan, in 1988, 1990, and 1993, respectively. From 1993 to 2005, he was with the Department of Electronics and Information Engineering at Tokyo Metropolitan University, Tokyo, Japan. He is currently a professor at the Graduate School of Informatics, Kyoto University, Kyoto, Japan. He received DoCoMo Mobile Science Prize in 2003, and Young Scientist Award from the Minister of Education, Culture, Sports, Science and Technology, Japan, in 2005. His research interests are in the areas of information and communication theory, statistical mechanics of information processing, machine learning, and neural networks. He is a member of the IEEE, the Japanese Neural Network Society, the Acoustical Society of Japan, the Physical Society of Japan, and the Architectural Institute of Japan.

Kenta Kasai received B.E., M.E., and Ph.D. degrees from the Tokyo Institute of Technology, Tokyo Japan, in 2001, 2003 and 2006, respectively. He joined the faculty of Tokyo Institute of Technology in 2006 as an assistant processor. Since 2012, he has been an associate professor in the Department of Communications and Integrated Systems, Graduate School of Science and Engineering, Tokyo Institute of Technology. From 2008 to 2009, he was a visiting researcher at ETIS laboratory, Ecole Nationale Supérieure de l'Electronique et de ses Applications, Cergy-Pontoise, France. His current interests include coding theory.

Kohichi Sakaniwa received B.E., M.E., and Ph.D. degrees all in electronic engineering from the Tokyo Institute of Technology, Tokyo Japan, in 1972, 1974 and 1977, respectively.

He joined the Tokyo Institute of Technology in 1977 as a research associate and served as an associate professor from 1983 to 1991. Since 1991 he has been a professor in the Department of Electrical and Electronic Engineering, and since 2000 in the Department of Communications and Integrated Systems, Graduate School of Science and Engineering, both in the Tokyo Inst. of Tech. From November 1987 to July 1988, he stayed at the University of Southwestern Louisiana as a Visiting Professor. He received the Excellent Paper Award from the IEICE of Japan in 1982, 1990, 1992 and 1994. His research area includes Communication Theory, Error Correcting Coding, (Adaptive) Digital Signal Processing and so on.

Dr. Sakaniwa is a member of IEEE, IEICE Japan, Information Processing Society of Japan, and Institute of Image Information and Television Engineers of Japan.