Detailedly Represented Irregular Low-Density Parity-Check Codes*

Kenta KASAI††, Nonmember, Tomoharu SHIBUYA†, Regular Member, and Kohi SAKANIWA††, Fellow

1. Introduction

Richardson and Urbanke [1] established an elegant algorithm called the density evolution to determine the threshold of a channel parameter which characterizes a binary input memoryless channel. The threshold is uniquely determined for a given code ensemble and defined as the supremum of a parameter of the channel under consideration, at which most of codes in the ensemble realize reliable communication with belief propagation (BP) decoding.

It is widely accepted that an ensemble of bipartite graphs acts a fundamental role in an argument of the density evolution. We denote by $G := (V \cup C, E)$ a bipartite graph with a node set $V \cup C$ with $V \cap C = \phi$ and an edge set $E$ each of which has one end in $V$ and the other end in $C$. We write elements of $V$ and $C$ by $V = \{v_1, v_2, \ldots, v_n\}$ and $C = \{c_1, c_2, \ldots, c_m\}$, and associate a bipartite graph $G$ with a $|C| \times |V|$ parity check matrix $H_G = (h_{ij})$ in such a way that $h_{ij} = 1$ if and only if $c_i$ and $v_j$ have odd numbers of edges between them. Through this association, we can regard a bipartite graph as a code defined by $H_G$, and therefore elements of $V$ and $C$ are called variable nodes and check nodes, respectively.

Moreover, we introduce a pair of polynomials $\lambda(x) := \sum_{i=2}^{d_v} \lambda_i x^{i-1}$ and $\rho(x) := \sum_{i=2}^{d_c} \rho_i x^{i-1}$ with positive coefficients where $d_v \geq 2$, $d_c \geq 2$ and $\lambda(1) = \rho(1) = 1$, which are called a degree distribution pair. Then we can define, through the association described above, a code ensemble $\mathcal{C}(n, \lambda(x), \rho(x))$ as a collection of all bipartite graphs such that the fraction of all edges connecting variable nodes (resp. check nodes) of degree $i$ is equal to $\lambda_i$ (resp. $\rho_i$).

MacKay [3] has clarified experimentally that there exists variation for the performance in decoding among binary irregular LDPC codes in $\mathcal{C}(n, \lambda(x), \rho(x))$, which depends on construction. The construction called Poisson specifies the number of ones in disjointly divided square sub-matrices of the parity check matrix. In other words, the construction specifies the number of edges connecting between disjointly divided sets of variable nodes and check nodes. The variation of performance of these irregular codes is the motivation for this paper, in which we explore ways to find better code ensembles. The aim of this paper is to clarify the reason of the variation theoretically. For this purpose, we will introduce detailed representation of code ensembles.

In this paper, we introduce new code ensembles by specifying the fraction of edges connecting sets of variable nodes and check nodes with given degree at the same time. Then we present generalized density evolution for the newly proposed code ensembles. We also show that the generalized density evolution can treat the density evolution for a conventional code ensemble as a special case. Moreover by starting from known highly optimized conventional code ensembles used over a binary erasure channel (BEC) and a binary input additive white Gaussian noise channel (BIAWGNC), we find code ensembles which have the same degree distributions and better thresholds by simple try-and-check approach.

2. Preliminaries

2.1 Belief Propagation Decoder

In decoding of a low-density parity-check (LDPC) code associated with a bipartite graph $G = (V \cup C, E)$, a BP
decoder sends messages through each edge of $G$, which are set initially and updated iteratively in the following manner. Let $\mathbb{R}$ be the set of all real numbers and we denote by $X$ and $Y$ random variables describing transmitted and received symbols, respectively, associated with a variable node $v \in V$.

We define an initial message $m^{(0)}(y) \in \mathbb{R} \cup \{+\infty, -\infty\}$ by

$$m^{(0)}(y) = \begin{cases} +\infty & (p_+^1(y) = 0), \\ \ln \frac{p_+^1(y)}{p_-^1(y)} & (0 < p_+^1(y) < 1), \\ -\infty & (p_+^1(y) = 0), \\ \end{cases}$$

where $p_{\pm}^1(y) = \Pr\{Y = y \mid X = \pm 1\}$. Note that we employ the standard binary PAM map $GF(2) \rightarrow \{1, -1\}$ defined by 0 $\mapsto$ 1 and 1 $\mapsto$ -1.

For a variable node $v \in V$ and a check node $c \in C$ which have a common edge between them, let $m_{vc}^{(i)}$ and $m_{cv}^{(i)}$ be messages sent from $v$ to $c$ and from $c$ to $v$ at the $i$-th iteration of BP decoding. The update rule for these messages are given as:

$$m_{vc}^{(i)} = \begin{cases} m^{(0)}(y) & (i = 0), \\ m^{(0)}(y) + \sum_{c' \in C \setminus \{c\}} m_{c'v}^{(i)} & (i \geq 1), \\ \end{cases}$$

$$m_{cv}^{(i)} = \gamma^{-1} \left( \sum_{v' \in V \setminus \{v\}} \gamma \left( m_{v'c}^{(i-1)} \right) \right)$$

where $C_v := \{c \in C \mid (v, c) \in E\}$ and $V_c := \{v \in V \mid (v, c) \in E\}$. In the update rule of $m_{cv}^{(i)}$, $\gamma$ denotes a function

$$\gamma: \mathbb{R} \cup \{+\infty, -\infty\} \rightarrow GF(2) \times [0, +\infty],$$

$$x \mapsto (\text{sgn}_p x, -\ln \tanh \frac{x}{2})$$

where $\ln(0) := -\infty$ and $\text{sgn}_p x$ denotes

$$\text{sgn}_p x = \begin{cases} 0 & (x > 0), \\ 0 \text{ or } 1 \text{ with equiprobable} & (x = 0), \\ 1 & (x < 0). \\ \end{cases}$$

### 2.2 Density Evolution

In order to investigate the evolution of a probability density function of messages in accordance with the increment of iteration number in BP decoding, Richardson et al. [2] introduced a notion of a general class of distributions. In this subsection, we briefly review the notion and summarize the density evolution. In the remaining part of this paper, we abbreviate a probability density function as a density according to the terminology in [2].

Let $F$ denote a space of right-continuous nondecreasing functions $F$ on $\mathbb{R}$ satisfying $\lim_{x \rightarrow -\infty} F(x) = 0$ and $\lim_{x \rightarrow \infty} F(x) \leq 1$ for $F(x) \in F$. We can consider a random variable $z$ by regarding a function $F(x) \in F$ as $F(x) = \Pr\{z \in (-\infty, x]\}$, a distribution function of $z$. Then a density of $z$ can be defined by Random-Nikodym derivative of $F(x)$.

For $F$ and $G$ in $F$, their convolution is defined as $(F \otimes G)(x) := \int_{\mathbb{R}} F(x-y) dG(y) = \int_{\mathbb{R}} G(x-y) dF(y)$.

Let also $G$ denote a space of functions on $GF(2) \times [0, +\infty)$ whose element is expressed in the form of $G(l, x) = \delta_{G_0}(l, x) + \delta_{G_1}(l, x)$ where $G_0(l, x)$ and $G_1(l, x)$ represent right-continuous nondecreasing functions on $[0, +\infty)$ satisfying $G_0(0) \geq 0$ and $G_1(0) = 0$, and $\delta_l$ denotes Kronecker’s delta, that is, $\delta_l$ is 1 if $l = i$ and 0 if $l \neq i$.

For two elements of $G$ expressed as $G(l, x) := \delta_{G_0}(l, x) + \delta_{G_1}(l, x)$ and $H(l, x) := \delta_{H_0}(l, x) + \delta_{H_1}(l, x)$, their convolution is defined as $(G \otimes H)(l, x) := \delta_{G_0}(l) \cdot \delta_{H_1}(l) + \delta_{H_0}(l) \cdot \delta_{G_1}(l)$. Then it is known that for a random variable $z \in (-\infty, +\infty)$ with a distribution function $F_z(x) \in F$, $\gamma(z)$ has a distribution function $\Gamma(F_z)(l, x) = \delta_{G_0} \Gamma_0(F_z)(x) + \delta_{G_1} \Gamma_1(F_z)(x)$ where

$$\Gamma_0(F_z)(x) = 1 - F_z^{-1}(\ln \tanh \frac{x}{2}),$$

$$\Gamma_1(F_z)(x) = F_z^{-1}(\ln \tanh \frac{x}{2})$$

and $F_z^{-1}(x)$ is defined as the left limit of $F_z(x)$ at $x$.

It can be easily verified that $\Gamma$ and $\Gamma^{-1}$ are additive operators on $F$ and $G$, respectively. Therefore for $F_i \in F$ and $G_i \in G$ $(i = 1, 2, \ldots, n)$ and positive real numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$ with $\sum_1^n \alpha_i = 1$, we have

$$\begin{aligned}
\Gamma \left( \sum_{i=1}^n \alpha_i F_i \right) & = \sum_{i=1}^n \alpha_i \Gamma(F_i), \\
\Gamma^{-1} \left( \sum_{i=1}^n \alpha_i G_i \right) & = \sum_{i=1}^n \alpha_i \Gamma^{-1}(G_i).
\end{aligned}$$

For convenience, although it may be an abuse of notation, we also apply $\Gamma$ and $\Gamma^{-1}$ to densities. More precisely, for a density $f$ whose corresponding distribution is $F \in F$, $\Gamma(f)$ denotes the density whose corresponding distribution is $\Gamma(F) \in G$. Similarly, for a density $f$ whose corresponding distribution is $G \in G$, $\Gamma^{-1}(y)$ denotes the density whose corresponding distribution is $\Gamma^{-1}(G) \in F$.

By using notations described above and assuming that every graph in a given ensemble has no cycles of length less than or equal to $2T$, which is called local tree assumption of depth $2T$, Richardson et al. developed the density evolution which is summarized in the following theorem.

**Theorem 1:** [2] For $G(n, \lambda(x), \rho(x))$, let $P_0$ and $P_1$ denote densities of $m^{(0)}$ and $m^{(1)}$. Then under local tree assumption of depth $2T$, $P_1 \otimes \lambda(\Gamma^{-1}(\rho(\Gamma(P_1))))$
for \(1 \leq i < T\), where a multiplication of polynomials represents convolution.

Let \(\sigma\) be a channel parameter, which determines the density \(P_b\) of \(m^{(0)}\) over the channel. A crossover probability for binary symmetric channel, an erasure probability for BEC and a standard deviation for BI-AWGN are examples of the parameters. Threshold for an ensemble is determined as the supremum of the channel parameter \(\sigma\) such that \(\sum_{m=0}^{\infty} P_b(m)dm\) converges to 0 as \(i\) tends to infinity, assuming that all one codeword\(^{1}\) is sent. Note that since the fraction of graphs which satisfy local tree assumption out of all graphs in the ensemble converges to 1 as \(n\) tends to infinity [1], Theorem 1 is true for sufficiently large \(n\).

### 3. Density Evolution with Detailed Representation

In this section we shall introduce new code ensembles and investigate their density evolution.

#### 3.1 New Ensembles

##### 3.1.1 Definitions

Let \(B\) and \(D\) be sets of containers which are associated \(\pi : B \times D \rightarrow [0,1]\) is said to be a joint degree distribution of \((B,D)\) if

\[
\sum_{b \in B} \sum_{d \in D} \pi(b,d) = 1.
\]

Let \(G = (V \cup C,E)\) be a bipartite graph and assume that \(B\) and \(D\) are given.

We say that \(v \in V\) (resp. \(c \in C\)) belongs to \(b \in B\) (resp. \(d \in D\)), expressed as \(v \in b\) (resp. \(c \in d\)), if degree of \(v\) (resp. degree of \(c\)) is equal to the associated natural number of \(b\) (resp. \(d\)). Moreover, we say that \(V\) and \(C\) are divisible on \(B\) and \(D\) if all \(v \in V\) and \(c \in C\) have \(b \in B\) and \(d \in D\) with \(v \in b\) and \(c \in d\). From these definition, we regard elements of \(B\) and \(D\) as sets of variable and check nodes and call them variable blocks and check blocks, respectively. Furthermore, we call the associated numbers of \(b\) and \(d\) are their degrees and denoted by \(\deg(b)\) and \(\deg(d)\), respectively.

For \(\pi(b,d)\), we define two fractions

\[
\lambda(b,d) := \pi(b,d)/\hat{\rho}_d, \quad \rho(b,d) := \pi(b,d)/\hat{\lambda}_b
\]

where \(\hat{\lambda}_b := \sum_{d \in D} \pi(b,d)\) and \(\hat{\rho}_d := \sum_{b \in B} \pi(b,d)\). It can be easily verified that \(\rho(b,d)\) (resp. \(\lambda(b,d)\)) agree with the fraction of edges connecting nodes in \(b\) and \(d\) among all edges emanating from variable nodes in \(b\) (resp. check nodes in \(d\)). We also define marginal degree distributions of variable and check blocks with respect to \(\pi\) by \(\lambda(x) := \sum_{b \in B} \hat{\lambda}_b x^{\deg(b)-1}\) and \(\hat{\rho}(x) := \sum_{d \in D} \hat{\rho}_d x^{\deg(d)-1}\), respectively.

Moreover, we define a variable socket (resp. check socket) as what sticks to a variable node (resp. check node) and edges plugged into. We say a socket is unplugged if the socket has no edge stuck to it. And an unplugged variable socket (resp. check socket) is said to be bound with a check block \(d\) (resp. bound with a variable block \(b\)), if such a restriction that one has to plug an edge into the variable socket (resp. the check socket) at one end and a socket stuck to a check node in \(d\) (a variable node in \(b\)) at the other end is imposed.

#### 3.1.2 Ensemble \(C_1(n,\pi)\)

Let \(E_{b,d} := \{(v,c) \in E : v \in b, c \in d\}\). We define \(C_1(n,\pi)\) as a set of all bipartite graphs \(G = (V \cup C,E)\) such that \(|V| = n\), \(V\) and \(C\) are divisible on \(B\) and \(D\), and \(|E_{b,d}| = \pi(b,d)|E|\).

Counting both variable and check nodes of \(G = V \cup C \in C_1(n,\pi)\), we can easily verify that \(|V| = |E| \int_0^1 \lambda(x)dx\) and \(|C| = |E| \int_0^1 \rho(x)dx\). We may assume all parity check equations are linearly independent for given \(\pi\) and sufficiently large \(n\). Then all codes in \(C_1(n,\pi)\) have the identical rate given as \(R = 1 - \int_0^1 \rho(x)dx/\int_0^1 \lambda(x)dx\). This means that the rate of codes in \(C_1(n,\pi)\) is determined only by \(\pi\) through marginal degree distributions. On the other hand, all graphs in \(C_1(n,\pi)\) have the identical number of edges given as \(E := n/\int_0^1 \lambda(x)dx\), which is determined by both \(n\) and \(\pi\) through marginal degree distributions.

For given code length \(n\), the construction of codes in \(C_1(n,\pi)\) can be realized in the following two successive steps.

**Assigning sockets step:** Let all sockets be unplugged. First, for each \(b \in B\), divide a set of all unplugged \(\lambda_b E\) sockets stuck to variable nodes in \(b\) randomly into disjoint \(|D|\) sets \(s_b(d)\) for \(d \in D\) such that each \(s_b(d)\) is of size \(\pi(b,d)E\) and make every elements of \(s_b(d)\) be bound with \(d\). Similarly, for each \(d \in D\), divide a set of all unplugged \(\hat{\rho}_d E\) sockets stuck to check nodes in \(d\) randomly into disjoint \(|B|\) sets \(s_d(b)\) for \(b \in B\) such that each \(s_d(b)\) is of size \(\pi(b,d)E\) and make every elements of \(s_d(b)\) be bound with \(b\).

**Assigning edges step:** Secondly connect sockets of \(s_b(d)\) and \(s_d(b)\) randomly with \(\pi(b,d)E\) edges for every pair \(b \in B\) and \(d \in D\).

**Example 1:** A graph illustrated in Fig.1 is included in \(C_1(n,\pi)\). Round and square nodes are variable and check nodes, respectively. Variable blocks \(b_2\) and \(b_3\) contains variable nodes of degree 2 and 3 respectively. Check blocks \(d_3\) and \(d_4\) contains check nodes of degree 3 and 4 respectively. The joint degree distribution \(\pi(b_i,d_j)\) which are the fractions of edges connecting variable nodes of degree \(i\) and check nodes of degree \(j\) are \(\pi(b_2,d_3) = 4/17, \pi(b_2,d_4) = 4/17, \pi(b_3,d_3) = 5/17, \pi(b_3,d_4) = 4/17\).
3.1.3 Ensemble $\mathcal{C}_2(n, \pi)$

Let $B' \subset \mathcal{B}$ and $D' \subset \mathcal{D}$ be multisets\(^1\) of size $\deg(d)$ and $\deg(b)$. Define $p_d^{(G)}(B') := \left\{ \frac{\lambda(b', d')w(B', b')}{\rho_{d}(B')\deg(d)} \right\}$ and $q_d^{(G)}(D') := \left\{ \frac{\rho(b', d')w(D', d')}{\deg(d)} \right\}$ where $D_v$ is the multiset of check blocks which $v$’s neighbors belong to and $B_v$ is the multiset of variable blocks which $c$’s neighbors belong to. Then we define $\mathcal{C}_2(n, \pi)$ as an ensemble which consists of all bipartite graphs $\mathcal{G} = (V \cup C, E)$ such that $|V| = n$, $V$ and $C$ are divisible on $\mathcal{B}$ and $\mathcal{D}$, and

\[
\begin{align*}
\forall b \in \mathcal{B}, & \quad p_d^{(G)}(B') = \deg(d) \prod_{b' \in \mathcal{B}} \frac{\lambda(b', d')w(B', b')}{w(B', b')!}, \\
\forall d \in \mathcal{D}, & \quad q_d^{(G)}(D') = \deg(b) \prod_{d' \in \mathcal{D}} \frac{\rho(b', d')w(D', d')}{w(D', d')!}
\end{align*}
\]

for every $b \in \mathcal{B}, d \in \mathcal{D}$, and multisets $B' \subset \mathcal{B}$ and $D' \subset \mathcal{D}$ of size $\deg(d)$ and $\deg(b)$ respectively. In Eq. (3) $w(X, a)$ denotes the number of $a$ belonging to a multiset $X$. It can be verified $|E_{b, d}| = \pi(b, d)|E|$ holds for every $\mathcal{G} \in \mathcal{C}_2(n, \pi)$ and every pair of $b \in \mathcal{B}$ and $d \in \mathcal{D}$ in the following way. We have

\[
|E_{b, d}| = \sum_{B' \subset \mathcal{B}} |\{c \in d : B_c = B'\}|w(B', b) \quad (4)
\]

because any check node $c \in d$ such that $B_c = B'$ has $w(B', b)$ edges emanating from variable nodes in $b$. By substituting the definition of $p_d^{(G)}(B')$ to (4) and using

\[
1 = \left( \sum_{b' \in \mathcal{B}} \lambda(b', d) \right)^{\deg(d)-1} = \sum_{B \in \mathcal{B}} \left( \deg(d)-1 \right)! \prod_{b' \in \mathcal{B}} \frac{\lambda(b', d)w(B, b')}{w(B, b')!},
\]

and Eq. (2) and (3), we derive $|E_{b, d}| = \pi(b, d)|E|$, which implies $\mathcal{C}_2(n, \pi) \subset \mathcal{C}_1(n, \pi)$.

For given code length $n$, an element of an ensemble $\mathcal{C}_2(n, \pi)$ can be realized in the following two successive steps.

**Assigning sockets step:** Let all sockets be unplugged. For each block $b \in \mathcal{B}$ and for all multiset $D' \subset \mathcal{D}$ of size $\deg(b)$, prepare $\frac{\lambda_{b,D}(b)!}{\deg(b)!} \prod_{d' \in \mathcal{D}} \frac{\rho(b, d')w(b, d')}{w(b, d')!}$ variable nodes each $v$ of which has randomly permuted $\deg(b)$ variable sockets such that a multiset of variable blocks with which the variable sockets are bound agrees with $D'$. Similarly, for each block $d \in \mathcal{D}$ and all multiset $B' \subset \mathcal{B}$ of size $\deg(d)$, prepare $\frac{\rho_{b,D}(d)!}{\deg(d)!} \prod_{b \in \mathcal{B}} \frac{\lambda_{b,D}(b)'w(b, \pi')}{w(b, \pi')!}$ check nodes each $c$ of which has randomly permuted $\deg(d)$ check sockets such that a multiset of variable blocks with which the check sockets are bound agrees with $B'$. Now, all $\lambda_{b,D}$ sockets stuck to variable nodes in $b$ have been divided into disjoint $|D|$ sets $s_{b,d}(d) \in \mathcal{D}$ such that each $s_{b,d}(d)$ is of size $\pi(b, d)E$ and every elements of $s_{b,d}(d)$ is bound with $d$. Similarly, all $\rho_{b,D}$ sockets stuck to check nodes in $d$ has been divided into disjoint $|\mathcal{B}|$ sets $s_{d,b}(b) \in \mathcal{B}$ such that each $s_{d,b}(b)$ is of size $\pi(b, d)E$ and every elements of $s_{d,b}(b)$ is bound with $b$.

**Assigning edges step:** We connect sockets of $s_{b,d}(d)$ and $s_{d,b}(b)$ randomly with $\pi(b, d)E$ edges for every pair $b \in \mathcal{B}$ and $d \in \mathcal{D}$.

3.2 Density Evolution

In the analysis of the conventional ensembles, it is assumed that in any graph, messages from variable nodes to check nodes at the $i$-th iteration of BP decoding obey the identical distribution with the density $P_i^B$. On the other hand we assume that all messages from variable nodes in $b$ to check nodes in $d$ and in the opposite direction in a graph $\mathcal{G}$ at the $i$-th iteration obey the densities depending on $b$ and $d$, respectively, each of which is denoted by $P_i^b(b, d)$ and $Q_i^d(d, b)$, respectively. We suppose that $P_i^b(b, d)$ varies according to a pair of $b \in \mathcal{B}$ and $d \in \mathcal{D}$ and so does $Q_i^d(d, b)$. Further, we denote by $P_0(b)$ a density of the initial message of a variable node $v$ because we assume that the density depends on a block $b$ with $v \in b$.

Fix blocks $b \in \mathcal{B}$ and $d \in \mathcal{D}$ and a graph $\mathcal{G} \in \mathcal{C}_2(n, \pi)$. For a multiset $B \subset \mathcal{B}$ of size $\deg(d)-1$, we denote by $P_{i,B}^d(B)$ a probability that a randomly chosen edge $(v_0, c_0) \in E_{b,d}$ satisfies $B_{c_0} \setminus \{b\} = B$. Similarly, for a multiset $D \subset \mathcal{D}$ of size $\deg(b)-1$, we denote by $Q_{i,D}^b(D)$ the probability that a randomly chosen edges $(v_0, c_0) \in E_{b,d}$ satisfies $D_{v_0} \setminus \{d\} = D$. In other words, $P_{i,B}^d(B)$ and $Q_{i,D}^b(D)$ are fractions $|E_{b,d}|/|E_{b,d}|$ and $|E_{b,d}|/|E_{b,d}|$ of the cardinality of the following subset of

\(^1\)A multiset is a set-like object in which order is ignored, but multiplicity is explicitly significant, e.g. $\{a, a, b\}$ and $\{a, b, a\}$ are equivalent.
respectively, where
\[ E^B_{b,d} := \{(v_0, c_0) \in E_{b,d} \mid B_{c_0} \backslash \{b\} = B \}, \]
\[ E^D_{b,d} := \{(v_0, c_0) \in E_{b,d} \mid D_{c_0} \backslash \{d\} = D \}. \]

By analyzing density of messages over edges between variable and check blocks, we obtain the detailed density evolution described as the following theorem.

**Theorem 2:** For any graph \( G \in C_1(n, \pi) \), let \( p^\varnothing_{b,d}(B) \) and \( q^\varnothing_{b,d}(D) \) be probabilities, and \( P_0(b) \), \( P^i_t(b, d) \) and \( Q^i_t(b, d) \) be densities defined as above. Then under local tree assumption of depth \( 2T \),
\[ Q^i_t(b, d) = \Gamma^{-1} \left( \sum_{B \subseteq B_d} p^\varnothing_{b,d}(B) \right) \left( \bigotimes_{b' \in B} \Gamma \left( p^\varnothing_{b',d}(B') \right)^{\otimes(w(B,b'))} \right) \]
and
\[ P^i_t(b, d) = P_0(b) \otimes \left( \sum_{D \subseteq D_D} q^\varnothing_{b,d}(D) \right) \left( \bigotimes_{d' \in D} (Q^i_t(b, d'))^{\otimes(w(D,d'))} \right) \]
for \( 1 \leq i < T \).

Proof: Because of local tree assumption of depth \( 2T \), random variables describing each incoming message to \( c_0 \) are independent, and therefore the sum of random variables has a density given as convolution of densities of those random variables. Hence the density of the messages sent from \( c_0 \) to \( v_0 \) is given as
\[ \Gamma^{-1} \left( \bigotimes_{b' \in B} \Gamma \left( p^\varnothing_{b',d}(B') \right)^{\otimes(w(B,b'))} \right). \]

From the definition of \( p^\varnothing_{b,d}(B) \), we derive the density \( Q^i_t(b, d) \) by averaging Eq. (9) with probability \( p^\varnothing_{b,d}(B) \) on all \( B \subseteq B_d \) which is of size \( \deg(d) - 1 \) as
\[ \sum_{B \subseteq B_d} p^\varnothing_{b,d}(B) \Gamma^{-1} \left( \bigotimes_{b' \in B} \Gamma \left( p^\varnothing_{b',d}(B') \right)^{\otimes(w(B,b'))} \right) \]
which agrees with Eq. (7) for \( \sum_{|B|=\deg(d)-1} p^\varnothing_{b,d}(B) = 1 \) and Eq. (1). From the similar argument, we have Eq. (8).

3.2.1 Case of \( C_2(n, \pi) \)

Recall that \( C_2(n, \pi) \subset C_1(n, \pi) \). By limiting an ensemble \( C_1(n, \pi) \) to \( C_2(n, \pi) \), we can derive further theorem on its density evolution. We need the following lemma.

**Lemma 1:** For any \( G \in C_2(n, \pi) \), \( b \in B \), \( d \in D \), \( B \subseteq B \) of size \( \deg(d) - 1 \) and \( D \subseteq D \) of size \( \deg(b) - 1 \),
\[ p^\varnothing_{b,d}(B) = (\deg(d) - 1)! \prod_{b' \in B} \frac{\lambda(b', d) w(B, b')}{w(B, b')!}, \]
\[ q^\varnothing_{b,d}(D) = (\deg(b) - 1)! \prod_{d' \in D} \frac{\rho(b, d') w(D, d')}{w(D, d')!}. \]

Proof: When \( b \in B_z \) and \( B_z \backslash \{b\} = B \) we have \( B_z = B \cup \{b\} \). Then from the definition of \( E^B_{b,d} \), we have \( |E^B_{b,d}| \) is \( w(B \cup \{b\}, b) \) times of \(|\{c \in b \mid B_z = B \cup \{b\}\}| \). It follows that
\[ \pi(b, d) E p^\varnothing_{b,d}(B) = w(B \cup \{b\}, b) \frac{\rho_d E}{\deg(b)} q^\varnothing_{b,d}(B \cup \{b\}). \]

And similarly we have
\[ \pi(b, d) E q^\varnothing_{b,d}(D) = w(D \cup \{d\}, d) \frac{\lambda_d E}{\deg(b)} q^\varnothing_{b,d}(D \cup \{d\}). \]

By substituting these equations to (3) we have the lemma.

Using Lemma 1 and Theorem 2 and factorizing we can derive the following density evolution for the new ensembles.

**Theorem 3:** Under local tree assumption of depth \( 2T \), for any \( G \in C_2(n, \pi) \), \( b \in B \) and \( d \in D \), \( Q^i_t(b, d) \) and \( P^i_t(b, d) \) are equal to
\[ \tilde{Q}_i(d) = \Gamma^{-1} \left( \sum_{b \in B} \lambda(b, d) \tilde{P}_{i-1}(b) \right)^{(\otimes(\deg(d)-1))}, \]
\[ \tilde{P}_i(b) = \tilde{P}_0(b) \otimes \left( \sum_{d \in D} \rho(b, d) \tilde{Q}_i(d) \right)^{(\otimes(\deg(b)-1))} \]
respectively for \( 1 \leq i < T \) where \( \tilde{P}_0(b) := P_0(b) \). Note that \( \tilde{P}_i(b) \) and \( \tilde{Q}_i(d) \) do not depend on \( d \) and \( b \) respectively.

3.2.2 Asymptotic Analysis of \( C_1(n, \pi) \)

In this section we induce the property of the density evolution of graphs in \( C_1(n, \pi) \). We show that almost every graph in \( C_1(n, \pi) \) has the same form of density evolution as that of graphs in \( C_2(n, \pi) \) asymptotically as code length \( n \) tends to infinity. First we present the following lemma.

**Lemma 2:** Let \( P(a) \) is a map on a finite set \( \chi \) such that \( 0 \leq P(a) \leq 1 \) and \( \sum_{a \in \chi} P(a) = 1 \) and \( T^i_{P,k} \) be a set of all sequences \( X \in \chi^k \) such that the number of entries in \( X \) which is equal to \( a \) is \( P(a)k \). Let \( r \) be an integer such that \( r|k \). For given \( X = (x_1, x_2, \ldots, x_k) \), we define a multiset of multisets of \( \chi \) of size \( r \) as
Then for any $\epsilon > 0$ and $\delta > 0$ there exists $k_0 \in \mathbb{N}$ such that if $k > k_0$

$$
\left\{ X \in T_{P,k}^\chi \middle| \frac{w(X^*, X)}{|X^*|} - r! \prod_{a \in \mathcal{A}_X} P(a)^{w(X,a')}! \leq \epsilon \right\} \geq (1 - \delta)|T_{P,k}^\chi|
$$

(11)

for every multiset $X \subset \chi$ of size $r$.

Using this lemma we have the following lemma.

**Lemma 3:** For any $\epsilon > 0$ and $\delta > 0$, almost every graph as much as greater than or equal to $(1 - \delta)|C_1(n, \pi)|$ in $C_1(n, \pi)$ satisfies that

$$
\left| p_{b,d}(B) - (\deg(d) - 1)! \prod_{b' \in B} \lambda(b', d) w(B, b')! \right| \leq \epsilon,
$$

(12)

for all $b \in B$, $d \in D$. $B \subset B$ of size $\deg(d) - 1$ and $D \subset D$ of size $\deg(b) - 1$.

Proof: Recall two construction steps of codes in $C_1(n, \pi)$. They imply $p_{b,d}^\mathcal{G}(B)$ and $q_{b,d}^\mathcal{G}(D)$ are determined only by assignment of sockets step. And the step can be done by random choosing from $T_{P_e, \pi, E}^\mathcal{G}$ and $T_{Q_{\pi, E}}^\mathcal{G}$, where $P_{e}(b) = \lambda(b, d)$ and $Q_{b,d}(d) = \rho(b, d)$. Therefore, using Lemma 3 we have that for any $\epsilon' > 0$ and $\delta > 0$, almost every graph as much as greater than $(1 - \delta)|C_1(n, \pi)|$ in $C_1(n, \pi)$ satisfies that

$$
\left| p_{b,d}'(B') - \deg(d)! \prod_{b' \in B'} \lambda(b', d) w(B', b')! \right| \leq \epsilon',
$$

(13)

for all $b \in B$, $d \in D$, $B' \subset B$ of size $\deg(d)$ and $D' \subset D$ of size $\deg(b)$. Then using Eq. (13), we can derive Eq. (12) in the similar way of Lemma 1.

Using this lemma and the fact that convolution and $\Gamma$ are continuous as functions of densities, we have the following theorem which is the asymptotic version of density evolution for $C_1(n, \pi)$.

**Theorem 4:** Under local tree assumption of depth $2T$, for any $\epsilon > 0$, $\delta > 0$, there exist $n_0 \in \mathbb{N}$ such that if $n > n_0$ then for every $b \in B$ and $d \in D$, and for almost every $\mathcal{G}$ as much as equal or greater than $(1 - \delta)|C_1(n, \pi)|$ in $C_1(n, \pi)$, $|Q_{b,d}(d) - \hat{Q}_i(d)| \leq \epsilon$ and $|\hat{P}_i(b) - \hat{P}_i^0(b)| \leq \epsilon$ for $1 \leq i < T$ where $\hat{Q}_i(d)$ and $\hat{P}_i(b)$ are same as those in Eq. (10).

Note that $\hat{P}_i(b)$ and $\hat{Q}_i(d)$ do not depend on $d$ and $b$ respectively.

3.3 Connection with Conventional Ensembles

By averaging density over all blocks, an expectation of the density $\hat{P}_i$ of messages sent from variable nodes to check nodes at the $i$-th iteration is expressed as $\hat{P}_i := \sum_{b \in B} \lambda b \hat{P}_i(b)$. By a similar way, an expectation of the density $\hat{Q}_i$ of messages sent in the opposite direction at the $i$-th iteration is expressed as $\hat{Q}_i := \sum_{d \in D} \rho d \hat{Q}_i(d)$. The density evolution of $C(n, \lambda(x), \rho(x))$ can also be represented as the density evolution of a particular joint degree distribution constructed in the following proposition.

**Proposition 1:** For given degree distributions $\lambda(x)$ and $\rho(x)$, let $b_t$ and $d_t$ be blocks which consist of all variable and check nodes of degree $\ell$ and $\rho$ with $\lambda_\ell \neq 0$ and $\rho_\rho \neq 0$, respectively. We also define $B := \{b_t | \lambda_\ell \neq 0\}$ and $D := \{d_t | \rho_\rho \neq 0\}$. For these $B$ and $D$ we define a joint degree distribution $\pi$ as $\pi(b_t, d_t) := \lambda_\ell \rho_\rho$ for every $(\ell, r)$ with $\lambda_\ell \neq 0$ and $\rho_\rho \neq 0$. If $P_0(b_t) = P_0$ for all $\ell$ with $\lambda_\ell \neq 0$ then $\hat{P}_t = P_t$, where $P_0$ and $P_t$ are as given in Theorem 1.

Proof: From the definition of $\pi$, it holds that $\lambda b_t = \lambda(b_t, d_t) = \lambda_\ell$ and $\rho d_t = \rho(b_t, d_t) = \rho_\rho$ for every $(\ell, r)$ with $\lambda_\ell \neq 0$ and $\rho_\rho \neq 0$. Therefore we have from Eq. (10) that

$$
\hat{Q}_t(d_t) = \Gamma^{-1} \left( \left( \Gamma(\hat{P}_{t-1}) \right)^{\otimes(r-1)} \right), \quad \hat{P}_t(b_t) = P_0 \otimes \hat{Q}_t^{\otimes(t-1)}.
$$

These result in $\hat{Q}_t = \Gamma^{-1} \left( \rho(\Gamma(\hat{P}_{t-1})) \right)$ and $\hat{P}_t = P_0 \otimes \lambda(\hat{Q}_t)$. Finally, we see by induction that $\hat{P}_t$ coincides with $P_t$ for $i \geq 1$.

We see from Proposition 1 that density evolution of $C(n, \lambda(x), \rho(x))$ is equivalent to that of $C_2(n, \pi)$, when $\pi$ is so determined that the edges which emanate from any variable (resp. check) blocks have distribution $\rho(x)$ (resp. $\lambda(x)$). We utilize this result in the next section to find a new code ensemble whose joint degree distribution yield improved threshold.

4. Improvement of Threshold

So far, we expand the space of code ensembles by introducing detailed representation of code ensembles. Threshold for an ensemble $C_1(n, \pi)$ can be obtained by calculating the supremum of the channel parameter $\sigma$ such that $f_{\ell}^0 \hat{P}_i(m)dm$ converges to 0 as $i$ tends to infinity where $P_0$ is the density of $m_0$ over the channel with parameter $\sigma$, assuming that all one codeword was sent. Now we are interested in finding ensembles with better decoding performance. If we applied changes to the joint degree distribution obtained as Proposition 1, from good degree distribution which is optimized in the
space of conventional ensembles, would the threshold of the ensemble possibly get better? In this section we give such better code ensembles used over a BEC and an BIAWGNC compared with conventionally known ones. To find a joint degree distribution which determines a code ensemble with a better threshold, we use the following simple try-and-check approach. We first select an ensemble from highly optimized ensembles [2],[4],[5] as initial ensembles and construct a joint degree distribution from a degree distribution pair of the code ensemble by Proposition 1. Then we apply small changes to the derived joint degree distribution so that the rate and the marginal degree distributions may not change, and check if the ensemble has better threshold.

Since the rate of codes in an ensemble is determined by marginal degree distributions as shown in Sect.3.1.2, it is easily verified that increasing $\pi(b,d)$ and $\pi(b',d')$ by $\varepsilon$ and decreasing $\pi(b,d')$ and $\pi(b',d)$ by $\varepsilon$ respectively can change $\pi$ with small positive real number $\varepsilon$, keeping the marginal degree distribution and rate $R$ unchanged. If this change yields better threshold successfully, set the new degree distribution to be the currently best degree distribution, otherwise keep the original degree distribution. We repeat the change until we can no longer have any changes of threshold with any possible changes of $\pi$. Finally, we can have better degree distribution.

**Example 2:** From Eq. (10), the distribution for messages along edges connecting to block $b$ at the $i$-th iteration of BP decoding over a BEC is given as $x_i(b)\chi_{\{m\geq0\}}$ where

$$x_i(b) = x_0(b) \left( \sum_{d\in D} \rho(b,d) y_i(d) \right)^{\deg(b)-1},$$

$$y_i(d) = 1 - \left( 1 - \sum_{b\in B} \lambda(b,d) x_{i-1}(b) \right)^{\deg(d)-1}$$

and $\chi_{\{m\geq0\}}$ is 1 for $m \geq 0$ and 0 otherwise. A better degree distribution obtained by try-and-check approach is shown in Table 1. There listed $\lambda(b_i,d_j)$ and $\rho(b_i,d_j)$ for $i = 2,3,7,10$ and $j = 8,9$. An ensemble with the listed distribution has threshold 0.9549 and rate 0.5. It is obtained by starting with highly optimized conventional distributions $\lambda_2 = 0.26328$, $\lambda_3 = 0.18020$, $\lambda_7 = 0.27000$, $\lambda_{10} = 0.28649$, $\rho_8 = 0.63407$, $\rho_9 = 0.36593$ of threshold 0.49553 [4].

**Example 3:** We also present an example of codes of a BIAWGNC. The initial message density of BP over a BIAWGNC with standard deviation $\sigma$ is

$$P_0(m) = \sqrt{\frac{\sigma^2}{8\pi}} \exp \left( -\frac{(m - \frac{1}{2})^2 \sigma^2}{8} \right).$$

Amraoui presents a web site named LdpcOpt [5] which optimizes LDPC codes degree distribution and many good ensembles specified by a degree distribution for various channels can be obtained at the site. We pick an optimized degree distribution pair under constraints that the rate is 0.5 and the maximum degree of variable nodes is 10. A joint degree distribution which gives better threshold obtained by starting from this code is shown in Table 2. There listed $\lambda(b_i,d_j)$ and $\rho(b_i,d_j)$ for $i = 2,3,7,10$ and $j = 7,8$. An ensemble with the listed distribution has threshold 0.9622 and rate 0.5. It is obtained by starting with highly optimized conventional distributions $\lambda_2 = 0.27254$, $\lambda_3 = 0.23755$, $\lambda_4 = 0.07038$, $\lambda_{10} = 0.41953$, $\rho_7 = 0.70000$, $\rho_8 = 0.30000$ of threshold 0.9549. The input bit-error probability $p$ of hard-decision decoder and the gap to Shannon limit are also listed. Surprisingly we have such large improvement of thresholds for almost code ensembles listed in [2] for a BIAWGNC.

### Table 1 A degree distribution of a good detailedly represented LDPC code ensemble for BEC.

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$d_8$</th>
<th>$d_9$</th>
<th>$\rho$</th>
<th>$d_8$</th>
<th>$d_9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_2$</td>
<td>0.26288</td>
<td>0.26400</td>
<td>$b_2$</td>
<td>0.63308</td>
<td>0.36692</td>
</tr>
<tr>
<td>$b_3$</td>
<td>0.18062</td>
<td>0.17949</td>
<td>$b_3$</td>
<td>0.63551</td>
<td>0.36449</td>
</tr>
<tr>
<td>$b_7$</td>
<td>0.26972</td>
<td>0.27056</td>
<td>$b_7$</td>
<td>0.63340</td>
<td>0.36660</td>
</tr>
<tr>
<td>$b_{10}$</td>
<td>0.28678</td>
<td>0.28601</td>
<td>$b_{10}$</td>
<td>0.63470</td>
<td>0.36530</td>
</tr>
</tbody>
</table>

### Table 2 A degree distribution of a good detailedly represented LDPC code ensemble for a BIAWGNC.

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$d_7$</th>
<th>$d_8$</th>
<th>$\rho$</th>
<th>$d_7$</th>
<th>$d_8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_2$</td>
<td>0.25682</td>
<td>0.30920</td>
<td>$b_2$</td>
<td>0.65964</td>
<td>0.34036</td>
</tr>
<tr>
<td>$b_3$</td>
<td>0.23612</td>
<td>0.24089</td>
<td>$b_3$</td>
<td>0.69579</td>
<td>0.30421</td>
</tr>
<tr>
<td>$b_4$</td>
<td>0.06181</td>
<td>0.11371</td>
<td>$b_4$</td>
<td>0.51529</td>
<td>0.48741</td>
</tr>
<tr>
<td>$b_{10}$</td>
<td>0.45325</td>
<td>0.33620</td>
<td>$b_{10}$</td>
<td>0.75059</td>
<td>0.24041</td>
</tr>
</tbody>
</table>

### References


ldpcopt/}

Appendix: Proof of Lemma 2

Define a map $Y_X : \mathcal{X}_r \to \{0,1,\ldots,k_r\}$, by $Y_X(X) = w(X^r,X)$, where $\mathcal{X}_r$ is a set of all multisets $X \subset \chi$ of size $r$, i.e. $\mathcal{X}_r := \{X : X \subset \chi, |X| = r\}$. By regarding $Y_X$ as a random variable, we have that

$$\Pr \left\{ \left| \frac{Y_X(X)}{k_r} - r! \prod_{a \in \chi} \frac{P(a)^{w(X,a)}}{w(X,a)!} \right| > \varepsilon \right\} < \delta$$

are equivalent to (11), where $k_r := \frac{k}{r}$. If we can show that for every $X \in \mathcal{X}_r$

$$\mu_{k,X} := \mathbb{E} \left[ \left( \frac{Y_X(X)}{k_r} - r! \prod_{a \in \chi} \frac{P(a)^{w(X,a)}}{w(X,a)!} \right)^2 \right]$$

$$= 0 \quad (k \to \infty) \quad \text{(A-1)}$$

then using Markov’s inequality, we easily obtain that for any $\varepsilon, \delta > 0$, there exists $k_0$ such that if $k > k_0$ then

$$\Pr \left\{ \left| \frac{Y_X(X)}{k_r} - r! \prod_{a \in \chi} \frac{P(a)^{w(X,a)}}{w(X,a)!} \right| > \varepsilon \right\}$$

$$= \Pr \left\{ \left( \frac{Y_X(X)}{k_r} - r! \prod_{a \in \chi} \frac{P(a)^{w(X,a)}}{w(X,a)!} \right)^2 > \varepsilon^2 \right\}$$

$$< \frac{\mu_{k,X}}{\varepsilon^2} < \delta$$

for every $X \in \mathcal{X}_r$. The remaining part of this appendix is devoted to show that

$$\mathbb{E} \left[ \frac{Y_X(X)}{k_r} \right] = r! \prod_{a \in \chi} \frac{P(a)^{w(X,a)}}{w(X,a)!} \quad (k \to \infty) \quad \text{(A-2)}$$

and

$$\mathbb{E} \left[ \left( \frac{Y_X(X)}{k_r} \right)^2 \right] = \left( r! \prod_{a \in \chi} \frac{P(a)^{w(X,a)}}{w(X,a)!} \right)^2 \quad (k \to \infty) \quad \text{(A-3)}$$

for every $X \in \mathcal{X}_r$. If so then we have that

$$\mu_{k,X} = \mathbb{E} \left[ \left( \frac{Y_X(X)}{k_r} \right)^2 \right] + \mathbb{E} \left[ \frac{Y_X(X)}{k_r} \right]^2$$

$$- 2r! \prod_{a \in \chi} \frac{P(a)^{w(X,a)}}{w(X,a)!} \mathbb{E} \left[ \frac{Y_X(X)}{k_r} \right] = 0 \quad (k \to \infty),$$

for every $X \in \mathcal{X}_r$ which means (A-1).

First, we show (A-2). We have that the probability $\Pr \{ Y_X = y \}$, i.e. the fraction of $X \in \mathcal{T}_{P,k}$ such that $Y_X = y$, denote by $Q_{P,k}(y)$ is given by

$$\frac{k_r!}{\prod_{X' \in \mathcal{X}_r} Y(X')}! \prod_{X' \in \mathcal{X}_r} \left( \frac{r!}{\prod_{a \in \chi} w(X',a)!} \right)^{y(X')} |\mathcal{T}_{P,k}|^{-1},$$

because there are $\frac{k_r!}{\prod_{X' \in \mathcal{X}_r} Y(X')}!$ ways to arrange $k_r$ elements in $\mathcal{X}_r$ so that the number of $X'$ in the sequence is $y(X')$ for each $X' \in \mathcal{X}_r$ and there are $\prod_{a \in \chi} w(X',a)!$ elements in $\chi^r$ which are equivalent to $X$ as a multiset in $\mathcal{X}_r$. From the definition of $Q_{P,k}(y)$, it is straightforward that $\sum_{y \in \mathcal{Y}_{P,k}} Q_{P,k}(y) = 1$, where $\mathcal{Y}_{P,k}$ is a set of all maps $y : \mathcal{X}_r \to \{0,1,\ldots,k_r\}$ such that $\sum_{X \in \mathcal{X}_r} w(X,a)Y(X) = P(a)_k \forall a \in \chi$. Therefore we have that the expected fraction of $X \in \mathcal{X}_r$ which defined as

$$\mathbb{E} \left[ \frac{Y_X(X)}{k_r} \right] = \sum_{y \in \mathcal{Y}_{P,k}} \frac{y(X)}{k_r} Q_{P,k}(y).$$

is equal to

$$\frac{1}{k_r} \sum_{y \in \mathcal{Y}_{P,k}} \frac{y(X)}{\prod_{X' \in \mathcal{X}_r} Y(X')}! \prod_{X' \in \mathcal{X}_r} \left( \frac{r!}{\prod_{a \in \chi} w(X',a)!} \right)^{y(X')} |\mathcal{T}_{P,k}|^{-1}$$

$$= \prod_{a \in \chi} \frac{r!}{w(X,a)!} |\mathcal{T}_{P,k}|^{-1} \cdot \sum_{y \in \mathcal{Y}_{P,k}} \frac{(k_r - 1)!}{\prod_{X' \in \mathcal{X}_r \setminus \{X\} \neq 0} Y(X')!} \cdot$$

$$\cdot \prod_{X' \in \mathcal{X}_r \setminus \{X\}} \left( \frac{r!}{\prod_{a \in \chi} w(X',a)!} \right)^{y(X') - 1}.$$
By substituting (A.5) to (A.6) and then (A.5) to (A.4), we finally obtain
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 Communication 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, Information Processing Society of Japan, Institute of Image Information and Television Engineers of Japan, and Society of Information Theory and its Applications.