On Undetected Error Probability of Binary Matrix Ensembles

Tadashi Wadayama
Nagoya Institute of Technology
Nagoya, Aichi, Japan

Email: wadayama@nitech.ac.jp

***Abstract*— In this paper, analysis on undetected error probability of ensembles of $m \times n$ binary matricies is presented. Two ensembles are considered: One is an ensemble of dense matrices and another is an ensemble of sparse matrices. The main contributions of this work are (i) derivation of the error exponent of average undetected error probability and (ii) closed form expressions of the variance of undetected error probability. It is shown that the behavior of the exponent for a sparse ensemble is fairly different from that for a dense ensemble. The analysis for the sparse ensemble indicates the error detection performance achievable with time complexity $O(n)$. The variance of undetected error probability leads to a concentration result. Furthermore, as a byproduct of the proof of the variance formulas, simple covariance formulas of the weight distribution have been derived.**

I. Introduction

The *random coding* is extremely powerful technique to show the existence of a code satisfying certain properties. It has been used for proving direct part (achievability) of many types of coding theorems. In recent days, the idea of random coding is becoming more important as well from the practical point of view. An LDPC (Low-density parity-check) code can be constructed by choosing a parity check matrix from a sparse matrix ensemble. Thus, there are growing interests in randomly generated codes.

One of the largest problems for such a randomly generated code is that it is difficult to evaluate properties or performance of a randomly generated code such as minimum distance, weight distribution, ML decoding performance, etc.. To overcome this problem, we can take a *probabilistic approach*. In such an approach, an ensemble of parity check matrices is firstly assumed: i.e., probability is assigned to each matrix in the ensemble. In such a case, a property of a matrix (e.g., minimum distance, weight distributions) can be regarded as a random variable. It is natural to consider statistics of the random variable such as mean, variance, higher moments and covariance. In some cases, we can show that a property is strongly concentrated to its expectation. Such a concentration result justifies the use of the probabilistic approach.

Recent advance of analysis on average weight distributions of LDPC codes, such as Litsyn and Shevelev [3][4], Burshtein and Miller [5] Richardson and Urbanke [8] shows that the probabilistic approach is useful to investigate typical properties of codes and matrices, which are not easy to obtain from an instance. Furthermore, the second moment analysis on weight distribution of LDPC codes [6][7] can be utilized to prove some concentration results on weight distributions.

To evaluate the error detection probability of a given code (or given parity check matrix) is a classical problem in coding theory [2] and some results on this topic have been derived from the view point of the probabilistic approach. For example, the inequality for a linear code ensemble, $P_U < 2^{-m}$, has been long known where P_U is the undetected error probability and m is the number of rows of a parity check matrix. Since the undetected error probability can be expressed as a linear combination of weight distribution of a code, there are natural connection between expectation of weight distribution and expectation of undetected error probability.

In this paper, analysis on undetected error probability of ensembles of $m \times n$ binary matrices is presented. Two ensembles are considered: One is an ensemble of dense matrices, called a *random ensemble*, and another is an ensemble of sparse matrices, called a *sparse matrix ensemble*. An error detection scheme is a crucial part of a feedback error correction scheme such as ARQ. Detailed knowledge on error detection performance of a matrix ensemble would be useful for performance assessment for a feedback error correction scheme.

The contents of this paper are as follows: Firstly, we will focus on the error exponent of average undetected error probability. It will be shown that the asymptotic growth rate of the weight distribution determines the exponent. Then, the variance of undetected error probability will be discussed. To derive the variance, we need to know the covariance of the weight distribution. Simple covariance formulas for the random ensemble and the sparse matrix ensemble are derived based on a combinatorial approach.

II. Average undetected error probability

In this section, ensemble average of undetected error probability of a given matrix ensemble is discussed.

A. Notations

For a given $m \times n (m, n \geq 1)$ binary parity check matrix H, let $C(H)$ be the binary linear code of length n defined by H, namely,

$$C(H) \stackrel{\triangle}{=} \{\boldsymbol{x} \in F_2^n : H\boldsymbol{x}^t = \boldsymbol{0}\}, \tag{1}$$

where F_2 is the Galois field with two elements $\{0,1\}$ (the addition over F_2 is denoted by $\oplus$). In this paper, a boldface letter, like $\boldsymbol{x}$, denotes a binary row vector.

Throughout the paper, a binary symmetric channel (BSC) with the crossover probability ϵ ($0 < \epsilon < 1/2$) is assumed. We assume the conventional scenario for error detection. A transmitter send a codeword $\boldsymbol{x} \in C(H)$ to a receiver via BSC with crossover probability ϵ. The receiver obtain a received word $\boldsymbol{y} = \boldsymbol{x} \oplus \boldsymbol{e}^t$, where $\boldsymbol{e}$ denotes an error vector. The receiver firstly computes the syndrome $\boldsymbol{s} = H\boldsymbol{y}^t$ and then check whether $\boldsymbol{s} = \boldsymbol{0}$ holds or not.

An undetected error event occurs when $H\boldsymbol{e}^t = 0$ and $\boldsymbol{e} \neq \boldsymbol{0}$. This means that the error vector $\boldsymbol{e} \in C(\boldsymbol{e} \neq \boldsymbol{x})$ causes an undetected error event. Thus, the undetected error probability $P_U(H)$ can be expressed as

$$P_U(H) = \sum_{\boldsymbol{e} \in C(H), \boldsymbol{e} \neq \boldsymbol{0}} \epsilon^{w(\boldsymbol{e})}(1-\epsilon)^{n-w(\boldsymbol{e})} \tag{2}$$

where $w(\boldsymbol{x})$ denotes the Hamming weight of vector $\boldsymbol{x}$. The above equation can be rewritten as

$$P_U(H) = \sum_{w=1}^{n} A_w(H)\epsilon^w(1-\epsilon)^{n-w}, \tag{3}$$

where $A_w(H)$ is defined by

$$A_w(H) \stackrel{\triangle}{=} \sum_{\boldsymbol{x} \in Z^{(n,w)}} I[H\boldsymbol{x}^t = \boldsymbol{0}]. \tag{4}$$

The set $\{A_w(H)\}_{w=0}^n$ is usually called the *weight distribution* of $C(H)$. The notation $Z^{(n,w)}$ denotes the set of n-tuples with weight w. The notation $I[condition]$ is the indicator function such that $I[condition] = 1$ if $condition$ is true; otherwise, it gives 0.

Suppose that $\mathcal{G}$ is a set of binary $m \times n$ matrices ($m, n \geq 1$). Note that we allow that $\mathcal{G}$ contains some matrices with same configuration. Such matrices should be distinguished as distinct matrices. A matrix H in $\mathcal{G}$ is associated with probability $P(H)$. Thus, $\mathcal{G}$ can be considered as an *ensemble* of binary matrices. Let $f(H)$ be a real-valued function which depends on $H \in \mathcal{G}$. The expectation of $f(H)$ with respect to the ensemble $\mathcal{G}$ is defined by

$$E_{\mathcal{G}}[f(H)] \stackrel{\triangle}{=} \sum_{H \in \mathcal{G}} P(H)f(H). \tag{5}$$

The average weight distribution of a given ensemble $\mathcal{G}$ is given by $E_{\mathcal{G}}[A_w(H)]$. This quantity is very useful for performance analysis of binary linear codes including the undetected error probability.

B. Binary matrix ensembles

Two ensembles are mainly considered in this paper: the random ensemble and the sparse matrix ensemble. In this subsection, the definition and the average weight distribution of both ensembles are briefly reviewed.

1) Random ensemble: The random ensemble $\mathcal{R}_{m,n}$ includes all the binary $m \times n$ matrices ($m, n \geq 1$). From this definition, it is evident that the size of $\mathcal{R}_{m,n}$ is 2^{mn}. For each matrix in $\mathcal{R}_{m,n}$, equal probability $P(H) = 1/2^{mn}$ is assigned. It is well known [1] that the average weight distribution of $\mathcal{R}_{m,n}$ is given by

$$E_{\mathcal{R}_{m,n}}[A_w(H)] = 2^{-m}\binom{n}{w} \tag{6}$$

for $w \in [0, n]$. The notation $[a, b]$ means the set of consecutive integers from a to b. Because a typical instance of this ensemble contains $O(n^2)$ ones, the ensemble can be regarded as an ensemble of dense matrices.

2) Sparse matrix ensemble: The *sparse matrix ensemble* $\mathcal{T}_{m,n,k}$ contains all the binary $m \times n$ matrices ($m, n \geq 1$). The elements in a $m \times n$ binary matrix included in the sparse matrix ensemble are regarded as i.i.d. binary random variables such that an element takes the value 1 with probability $p \stackrel{\triangle}{=} k/n$. The parameter $k (0 < k \leq n/2)$ is a positive real number which represents the average number of ones for each row. In other words, a matrix $H \in \mathcal{T}_{m,n,k}$ can be considered as an output from the Bernoulli source such that symbol 1 occurs with probability p.

From the above definition, it is clear that a matrix $H \in \mathcal{T}_{m,n,k}$ is associated with the probability

$$P(H) = p^{\bar{w}(H)}(1-p)^{mn-\bar{w}(H)}, \tag{7}$$

where $\bar{w}(H)$ is the number of ones in H (i.e., Hamming weight of H). The average weight distribution of sparse matrix ensemble is given by

$$E_{\mathcal{T}_{m,n,k}}[A_w(H)] = \left(\frac{1+x^w}{2}\right)^m \binom{n}{w} \tag{8}$$

for $w \in [0, 1n]$ where $x \stackrel{\triangle}{=} 1 - 2p$. The average weight distribution of this ensemble was firstly discussed by Litsyn and Shevelev [3]. If k is a constant (i.e., not a function of n), a typical matrix in the ensemble contains $O(n)$ ones. Thus, this ensemble can be considered as an ensemble of sparse matrices.

C. Average undetected error probability of an ensemble

For a given $m \times n$ matrix H, the evaluation of the undetected error probability $P_U(H)$ is computationally difficult in general because we need to know the weight distribution of $C(H)$ for such evaluation. On the other hand, in some cases, we can evaluate the average of $P_U(H)$ for a given ensemble. Such a average probability is useful to estimate the undetected error probability of a matrix which belongs to the ensemble.

Taking ensemble average of the undetected error probability over a given ensemble $\mathcal{G}$, we have

$$\begin{aligned} E_{\mathcal{G}}[P_U(H)] &= E_{\mathcal{G}}\left[\sum_{w=1}^{n} A_w(H)\epsilon^w(1-\epsilon)^{n-w}\right] \\ &= \sum_{w=1}^{n} E_{\mathcal{G}}[A_w(H)]\epsilon^w(1-\epsilon)^{n-w}. \end{aligned} \tag{9}$$

In the above equations, H can be regarded as a random variable. From this equation, it is evident that the average of $P_U(H)$ can be evaluated if we know the average weight distribution of the ensemble. For example, in the case of the random ensemble $\mathcal{R}_{m,n}$, the average undetected error probability has the simple closed form.

Lemma 1: The average undetected error probability of random ensemble $\mathcal{R}_{m,n}$ is given by

$$E_{\mathcal{R}_{m,n}}[P_U(H)] = 2^{-m}(1-(1-\epsilon)^n). \tag{10}$$

(Proof) Combining (6) and (9), we have

$$\begin{aligned} E_{\mathcal{R}_{m,n}}[P_U(H)] &= \sum_{w=1}^{n} E_{\mathcal{R}_{m,n}}[A_w(H)]\epsilon^w(1-\epsilon)^{n-w} \\ &= \sum_{w=1}^{n} 2^{-m}\binom{n}{w}\epsilon^w(1-\epsilon)^{n-w} \\ &= 2^{-m}(1-(1-\epsilon)^n). \end{aligned} \tag{11}$$

The last equality is due to the binomial theorem. □

D. Error exponent of undetected error probability

For a given sequence of $(1-R)n \times n$ matrix ensembles $(n = 1, 2, 3, \ldots,)$, the average undetected error probability is usually exponentially decreasing function of n, where R is a real number satisfying $0 < R < 1$ (called *design rate*). Thus, the exponent of the undetected error probability has prime importance to grasp the asymptotic behavior of the undetected error probability.

1) Definition of error exponent: Let $\{\mathcal{G}_n\}_{n>0}$ be a series of ensembles such that $\mathcal{G}_n$ consists of $(1-R)n \times n$ binary matrices. In order to see the asymptotic behavior of the undetected error probability of this sequence of ensembles, it is reasonable to define the error exponent of undetected error probability in the following way.

Definition 1: The asymptotic error exponent of the average undetected error probability for a series of ensembles $\{\mathcal{G}_n\}_{n>0}$ is defined by

$$T_{\mathcal{G}_n} \stackrel{\triangle}{=} \lim_{n\to\infty} \frac{1}{n}\log_2 E_{\mathcal{G}_n}[P_U] \tag{12}$$

if the limit exists. □

Note that, from here, dependency to H is often omitted as P_U instead of $P_U(H)$ if there are no fear of confusion.

The following example shows the exponent of the random ensemble.

Example 1: Consider the series of random ensembles $\{\mathcal{R}_{n,(1-R)n}\}_{n>0}$. It is easy to evaluate $T_{\mathcal{R}_{(1-R)n,n}}$:

$$\begin{aligned} T_{\mathcal{R}_{(1-R)n,n}} &= \lim_{n\to\infty} \frac{1}{n}\log_2 E_{\mathcal{R}_{(1-R)n,n}}[P_U] \\ &= \lim_{n\to\infty} \frac{1}{n}\log_2 2^{-(1-R)n}(1-(1-\epsilon)^n) \\ &= -(1-R). \end{aligned} \tag{13}$$

This equality implies that the average undetected error probability of the sequence of random ensembles behaves like

$$E_{\mathcal{R}_{(1-R)n,n}}[P_U] \simeq 2^{-n(1-R)} \tag{14}$$

if n is sufficiently large. Note that the exponent $-(1-R)$ is independent from the crossover probability ϵ. □

2) Error exponent and asymptotic growth rate: The *asymptotic growth rate* of the average weight distribution (for simplicity, it is abbreviated as asymptotic growth rate), which is a basis of the derivation of the error exponent, is defined as follows.

Definition 2: Suppose that a series of ensembles $\{\mathcal{G}_n\}_{n>0}$ is given. If

$$\lim_{n\to\infty} \frac{1}{n}\log_2 E_{\mathcal{G}_n}[A_{\ell n}]$$

exists for $0 \le \ell \le 1$, then we define the *asymptotic growth rate* $f(\ell)$ by

$$f(\ell) \stackrel{\triangle}{=} \lim_{n\to\infty} \frac{1}{n}\log_2 E_{\mathcal{G}_n}[A_{\ell n}]. \tag{15}$$

The parameter ℓ is called *normalized weight*. □

From this definition, it is obvious that

$$E_{\mathcal{G}_n}[A_{\ell n}] = 2^{n(f(\ell)+o(1))} \tag{16}$$

holds. The notation $o(1)$ represents the term which converges to 0 as n goes to infinity. The asymptotic growth rate of some ensembles of binary matrices can be found in [3][4][5].

The next theorem gives the error exponent of the undetected error probability for a series of ensembles $\{\mathcal{G}_n\}_{n>0}$.

Theorem 1: The error exponent of $\{\mathcal{G}_n\}_{n>0}$ is given by

$$T_{\mathcal{G}_n} = \sup_{0<\ell\le 1}[f(\ell) + \ell\log_2\epsilon + (1-\ell)\log_2(1-\epsilon)], \tag{17}$$

where $f(\ell)$ is the asymptotic growth rate of $\{\mathcal{G}_n\}_{n>0}$.

(Proof) Based on the definition of asymptotic growth rate, we can rewrite $T_{\mathcal{G}_n}$ in the following form:

$$\begin{aligned} T_{\mathcal{G}_n} &= \lim_{n\to\infty} \frac{1}{n}\log_2 E_{\mathcal{G}_n}[P_U] \\ &= \lim_{n\to\infty} \frac{1}{n}\log_2 \sum_{w=1}^{n} E_{\mathcal{G}_n}[A_w]\epsilon^w(1-\epsilon)^{n-w} \\ &= \lim_{n\to\infty} \frac{1}{n}\log_2 \sum_{w=1}^{n} 2^{n(f(\frac{w}{n})+K(\epsilon,n,w)+o(1))}, \end{aligned}$$

where $K(\epsilon, n, w)$ is defined by

$$K(\epsilon,n,w) \stackrel{\triangle}{=} \frac{w}{n}\log_2\epsilon + \left(1-\frac{w}{n}\right)\log_2(1-\epsilon). \tag{18}$$

Using a conventional technique for bounding summation, we have the following upper bound on $T_{\mathcal{G}_n}$:

$$\begin{aligned} T_{\mathcal{G}_n} &= \lim_{n\to\infty} \frac{1}{n}\log_2 \sum_{w=1}^{n} 2^{n(f(\frac{w}{n})+K(\epsilon,n,w)+o(1))} \\ &\le \lim_{n\to\infty} \frac{1}{n}\log_2 n \max_{w=1}^{n} 2^{n(f(\frac{w}{n})+K(\epsilon,n,w)+o(1))} \\ &= \lim_{n\to\infty} \max_{w=1}^{n} \frac{1}{n}\log_2 2^{n(f(\frac{w}{n})+K(\epsilon,n,w)+o(1))} \\ &= \lim_{n\to\infty} \max_{w=1}^{n} \left[f\left(\frac{w}{n}\right) + K(\epsilon,n,w) + o(1)\right] \\ &= \sup_{0<\ell\le 1}[f(\ell) + \ell\log_2\epsilon + (1-\ell)\log_2(1-\epsilon)]. \end{aligned} \tag{19}$$

We can also show that $T_{\mathcal{G}_n}$ is greater than or equal to the right hand side of the above inequality (19) in a similar manner. This means that the right hand side of the inequality is asymptotically equal to $T_{\mathcal{G}_n}$. □

The next example discuss the case of random ensemble.

Example 2: Let us consider the series of the random ensemble $\{\mathcal{R}_{(1-R)n,n}\}_{n>0}$ again. This ensembles has the asymptotic growth rate $f(\ell) = h(\ell) - (1-R)$. The function $h(x)$ is the binary entropy function defined by

$$h(x) \stackrel{\triangle}{=} -x\log_2 x - (1-x)\log_2(1-x). \tag{20}$$

In this case, due to Theorem 1, we have

$$T_{\mathcal{R}_{(1-R)n,n}} = \sup_{0<\ell\leq 1}[h(\ell)-(1-R)+\ell\log_2\epsilon+(1-\ell)\log_2(1-\epsilon)]. \tag{21}$$

Let

$$D_{\ell,\epsilon} \stackrel{\triangle}{=} \ell\log_2\left(\frac{\ell}{\epsilon}\right) + (1-\ell)\log_2\left(\frac{1-\ell}{1-\epsilon}\right). \tag{22}$$

By using $D_{\ell,\epsilon}$, we can rewrite (21) as follows:

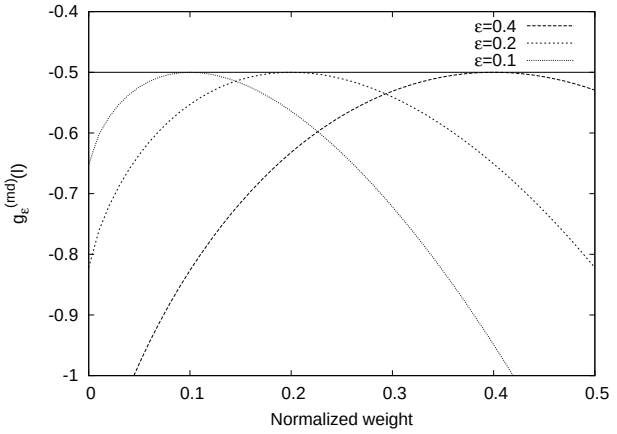
$$T_{\mathcal{R}_{(1-R)n,n}} = \sup_{0<\ell\leq 1}[-(1-R) - D_{\ell,\epsilon}]. \tag{23}$$

Since $D_{\ell,\epsilon}$ can be considered as Kullback-Libler divergence between two probability distributions $(\epsilon, 1-\epsilon)$ and $(\ell, 1-\ell)$, $D_{\ell,\epsilon}$ is always non-negative and $D_{\ell,\epsilon} = 0$ holds if and only if $\ell = \epsilon$. Thus, we get

$$\sup_{0<\ell\leq 1}[-(1-R) - D_{\ell,\epsilon}] = -(1-R), \tag{24}$$

which is identical to the exponent obtained in (13).

Let $g_\epsilon^{(rnd)}(\ell) \stackrel{\triangle}{=} h(\ell)-(1-R)+\ell\log_2\epsilon+(1-\ell)\log_2(1-\epsilon)$. Figure 1 presents the behavior of $g_\epsilon^{(rnd)}(\ell)$ when $R = 0.5$. This figure indicates that the maximum ($\sup_{0<\ell\leq 1} g_\epsilon^{(rnd)}(\ell) = -0.5$) is certainly attained at $\ell = \epsilon$. □



The curves of $g_\epsilon^{(rnd)}(\ell)$ correspond to the parameters $\epsilon = 0.1, 0.2, 0.4$ from left to right are presented. As a reference, line of $-(1-R) = -0.5$ is also included in the figure.

Fig. 1. The curves of $g_\epsilon(\ell)$ for random ensembles with $R = 0.5$

E. Error exponent of sparse matrix ensemble

The asymptotic growth rate of the sparse matrix ensemble $\mathcal{T}_{m,n,k}$ [3] with a constant k and design rate R is given by

$$f(\ell) = h(\ell) + (1-R)\log_2\left(\frac{1+e^{-2k\ell}}{2}\right). \tag{25}$$

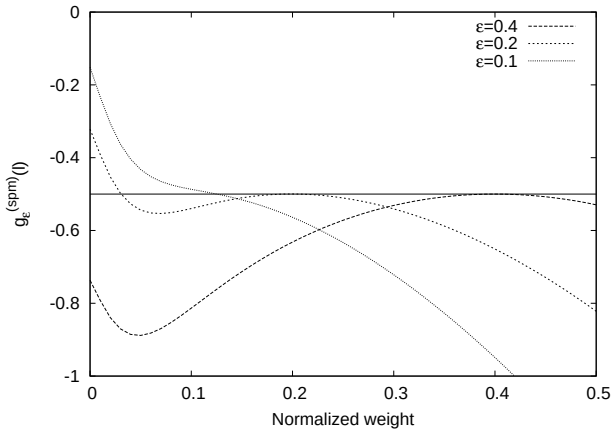
The error exponent of this ensemble shows quite different behavior from that of random ensemble.

Example 3: Consider the sparse matrix ensemble with parameters $R = 0.5$ and $k = 20$. Let

$$\begin{aligned} g_\epsilon^{(spm)}(\ell) &\stackrel{\triangle}{=} H(\ell) + (1-R)\log_2\left(\frac{1+e^{-2k\ell}}{2}\right) \\ &+ \ell\log_2\epsilon + (1-\ell)\log_2(1-\epsilon). \end{aligned} \tag{26}$$

Figure 2 includes the curves of $g_\epsilon^{(spm)}(\ell)$ where $\epsilon = 0.1, 0.2, 0.4$. We can see that $g_\epsilon^{(spm)}(\ell)$ is no more a concave function like $g_\epsilon^{(rnd)}(\ell)$ of random ensemble. The shape of the curve of $g_\epsilon^{(spm)}(\ell)$ depends on the crossover probability ϵ. For large ϵ, $g_\epsilon(\ell)$ takes the largest value around $\ell = \epsilon$. On the other hand, for small ϵ, $g_\epsilon^{(spm)}(\ell)$ have supremum at $\epsilon = 0$.

Figure 3 presents the error exponent of sparse matrix ensemble with parameters $R = 0.3, 0.5, 0.7, 0.9$ and $k = 20$. An an example, consider the exponent for $R = 0.5$. In the regime where ϵ is smaller than (around) 0.3, the error exponent is a monotonically decreasing function of ϵ.

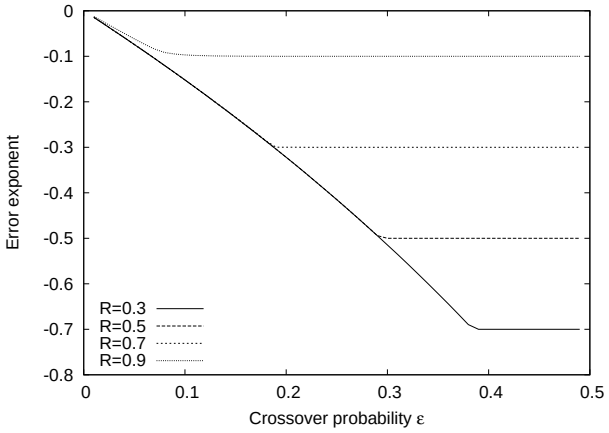


The curves of $g_\epsilon^{(spm)}(\ell)$ correspond to the parameters $\epsilon = 0.1, 0.2, 0.4$ are presented. The parameters $R = 0.5, k = 20$ are assumed. As a reference, line of $-(1-R) = -0.5$ is also included in the figure.

Fig. 2. The curves of $g_\epsilon^{(spm)}(\ell)$ for sparse matrix ensembles

The example suggest that a sparse ensemble has less powerful error detection performance than that of a dense ensemble (such as random ensemble) in terms of the error exponent. However, the crossover probability is sufficiently large, the difference in exponent of sparse and dense ensembles is negligble. For example, the exponent of the sparse matrix ensemble in Fig. 3 is almost equal to that of random ensemble when ϵ is larger than (around) 0.3.

The above properties on the error exponents of sparse matrix ensemble can be explained by its average weight distributions (or asymptotic growth rate). Figure 4 presents the asymptotic



The curves of $T_{\mathcal{T}_{m,n,k}}$ correspond to the parameters $R = 0.3, 0.5, 0.7, 0.9$ and $k = 20$. are presented.

Fig. 3. Error exponent of sparse matrix ensemble

growth rates of random ensemble and sparse matrix ensemble.

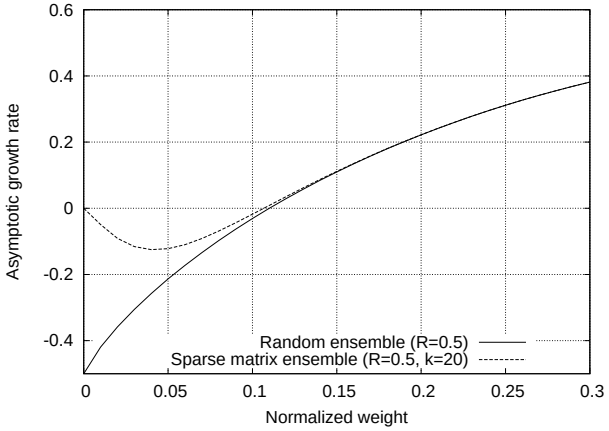


Fig. 4. Asymptotic growth rate of random ensemble and sparse matrix ensemble

The weight of typical error vectors is very close to ϵn when n is sufficiently large. If ϵ is high such as $\epsilon = 0.4$, the average weight distribution around $w = 0.4n$, namely $E_{\mathcal{G}}[A_{0.4n}]$, dominates the undetected error probability. In such a range, the difference of the average weight distribution between random ensemble and sparse matrix ensemble is very small. On the other hand, if the crossover probability is small, weight distributions of low weight become the most influential parameter. The difference in the average weight distributions of small weight results in the difference in the error exponent.

Note that time complexity of error detection operation (multiplication of received vector and a parity check matrix) takes $O(n^2)$-time for a typical instance of random ensemble and $O(n)$-time for that of sparse matrix ensemble. The linear time error detection with a sparse matrix achieves almost same error detection performance of a dense matrix if ϵ is sufficiently large.

The next lemma is useful to grasp the behavior of error exponent without detailed numerical optimization.

Lemma 2: If $f(\ell)$ has the form $f(\ell) = h(\ell) + \alpha(\ell)$, then the following lower bound on $T_{\mathcal{G}_n}$ holds:

$$T_{\mathcal{G}_n} \geq \max\left\{\alpha(0) + \log_2(1-\epsilon), \alpha(\epsilon)\right\}. \tag{27}$$

(Proof) Let

$$g_\epsilon(\ell) \stackrel{\triangle}{=} f(\ell) + \ell \log_2 \epsilon + (1-\ell)\log_2(1-\epsilon). \tag{28}$$

It is obvious that

$$\sup_{0<\ell\leq 1}[f(\ell) + \ell\log_2\epsilon + (1-\ell)\log_2(1-\epsilon)] \geq \max\{g_\epsilon(0), g_\epsilon(\epsilon)\} \tag{29}$$

holds. Since $h(0) = 0$, we have

$$g(0) = \alpha(0) + \log_2(1-\epsilon). \tag{30}$$

On the other hand, $g(\epsilon)$ is obtained in such a way:

$$\begin{aligned} g(\epsilon) &= h(\epsilon) + \alpha(\epsilon) + \epsilon\log_2\epsilon + (1-\epsilon)\log_2(1-\epsilon) \\ &= h(\epsilon) + \alpha(\epsilon) - H(\epsilon) \\ &= \alpha(\epsilon). \end{aligned} \tag{31}$$

Combining Theorem 1 and these results, we get the claim of the lemma. □

III. Variance of undetected error probability

In this section, we first discuss variance of undetected error probability for random ensemble. We then discuss the case of sparse matrix ensemble.

A. Variance of undetected error probability: random ensemble

1) Covariance formula: In the previous section, we have seen that the average weight distribution plays an important role in the derivation of average undetected error probability. Similarly, we need *covariance of weight distribution* to handle the variance of undetected error probability.

Definition 3: For $0 \leq w_1, w_2 \leq n$ and a given ensemble $\mathcal{G}$, covariance of weight distribution is defined by

$$\mathrm{Cov}_{\mathcal{G}}(A_{w_1}, A_{w_2}) \stackrel{\triangle}{=} E_{\mathcal{G}}[A_{w_1}A_{w_2}] - E_{\mathcal{G}}[A_{w_1}]E_{\mathcal{G}}[A_{w_2}]. \tag{32}$$

□

The next lemma is a basis of the derivation of variance of the undetected error probability for random ensemble.

Lemma 3: For random ensemble $\mathcal{R}_{m,n}$, covariance of A_{w_1} and A_{w_2} is given by

$$\begin{aligned} &\mathrm{Cov}_{\mathcal{R}_{m,n}}(A_{w_1}, A_{w_2}) \\ &= \begin{cases} 0, & 0 < w_1, w_2 \leq n, w_1 \neq w_2 \\ (1-2^{-m})2^{-m}\binom{n}{w}, & 0 < w_1 = w_2 \leq n. \end{cases} \end{aligned} \tag{33}$$

(Proof) See the forthcoming full paper version. □

Remark 1: The variance of weight distribution, namely $\mathrm{Cov}_{\mathcal{R}_{m,n}}(A_w, A_w) = (1-2^{-m})2^{-m}\binom{n}{w}$, has already shown in [8]. Thus, the new contribution of this lemma is the part $\mathrm{Cov}_{\mathcal{R}_{m,n}}(A_{w_1}, A_{w_2}) = 0$ when $w_1 \neq w_2$. □

Remark 2: Covariance of weight distribution for a given ensemble $\mathcal{G}$ is useful not only for evaluation of variance of P_U. Let X be a random variable represented by

$$X = \sum_{w=0}^{n} \alpha(w) A_w, \tag{34}$$

where $\alpha(w)$ is a real-valued function of w. The covariance of weight distribution is required for evaluation of variance of X, which is given by

$$\sigma_X^2 = \sum_{w_1=0}^{n} \sum_{w_2=0}^{n} \mathrm{Cov}_{\mathcal{G}}(A_{w_1}, A_{w_2}) \alpha(w_1) \alpha(w_2). \tag{35}$$

A specialized version (the case where $X = P_U$) of this equation will be derived in the proof of Theorem 2. For example, if $a(w) = 1(w \in [0, n])$, X denotes the number of codewords in $C(H)$. Based on the covariance, we can derive variance of number of codewords for a given ensemble $\mathcal{G}$. □

2) Variance of undetected error probability: The variance of the undetected error probability P_U is given by

$$\sigma^2_{\mathcal{R}_{m,n}} \stackrel{\triangle}{=} E_{\mathcal{R}_{m,n}}[(P_U - \mu)^2]. \tag{36}$$

The next theorem gives a closed formula of the variance $\sigma^2_{\mathcal{R}_{m,n}}$.

Theorem 2: For random ensemble $\mathcal{R}_{m,n}$, variance of the undetected error probability P_U is given by

$$\sigma^2_{\mathcal{R}_{m,n}} = (1 - 2^{-m}) 2^{-m} \left((\epsilon^2 + (1-\epsilon)^2)^n - (1-\epsilon)^{2n} \right). \tag{37}$$

(Proof) We first consider the second moment of the undetected error probability:

$$\begin{aligned} & E_{\mathcal{R}_{m,n}}[P_U^2] \\ &= E_{\mathcal{R}_{m,n}} \left[\left(\sum_{w=1}^{n} A_w \epsilon^w (1-\epsilon)^{n-w} \right)^2 \right] \\ &= E_{\mathcal{R}_{m,n}} \left[\sum_{w_1=1}^{n} \sum_{w_2=1}^{n} A_{w_1} A_{w_2} \epsilon^{w_1+w_2} (1-\epsilon)^{2n-w_1-w_2} \right] \\ &= \sum_{w_1=1}^{n} \sum_{w_2=1}^{n} E_{\mathcal{R}_{m,n}} \left[A_{w_1} A_{w_2} \right] \epsilon^{w_1+w_2} (1-\epsilon)^{2n-w_1-w_2}. \end{aligned} \tag{38}$$

The squared average undetected error probability can be expressed as

$$\begin{aligned} E_{\mathcal{R}_{m,n}}[P_U]^2 &= E_{\mathcal{R}_{m,n}} \left[\left(\sum_{w=1}^{n} A_w \epsilon^w (1-\epsilon)^{n-w} \right) \right]^2 \\ &= \sum_{w_1=1}^{n} \sum_{w_2=1}^{n} E_{\mathcal{R}_{m,n}} \left[A_{w_1} \right] E_{\mathcal{R}_{m,n}} \left[A_{w_2} \right] \\ &\times \epsilon^{w_1+w_2} (1-\epsilon)^{2n-w_1-w_2}. \end{aligned} \tag{39}$$

Combining these equalities and the covariance of the weight distribution (Lemma 3), variance of undetected error probability $\sigma^2_{\mathcal{R}_{m,n}}$ can be obtained in the following way:

$$\begin{aligned} & \sigma^2_{\mathcal{R}_{m,n}} \\ &= E_{\mathcal{R}_{m,n}}[P_U^2] - E_{\mathcal{R}_{m,n}}[P_U]^2 \\ &= \sum_{w_1=1}^{n} \sum_{w_2=1}^{n} \mathrm{Cov}_{\mathcal{R}_{m,n}} \left[A_{w_1}, A_{w_2} \right] \epsilon^{w_1+w_2} (1-\epsilon)^{2n-w_1-w_2} \\ &= \sum_{w=1}^{n} \mathrm{Cov}_{\mathcal{R}_{m,n}} \left[A_w, A_w \right] \epsilon^{2w} (1-\epsilon)^{2n-2w} \\ &= \sum_{w=1}^{n} (1 - 2^{-m}) 2^{-m} \binom{n}{w} \epsilon^{2w} (1-\epsilon)^{2n-2w}. \end{aligned} \tag{40}$$

The last equalities are due to Lemma 3. We can further simplify the expression using the binomial theorem in the following way:

$$\begin{aligned} \sigma^2_{\mathcal{R}_{m,n}} &= (1 - 2^{-m}) 2^{-m} \sum_{w=0}^{n} \binom{n}{w} (\epsilon^2)^w ((1-\epsilon)^2)^{n-w} \\ &- (1 - 2^{-m}) 2^{-m} (1-\epsilon)^{2n} \\ &= (1 - 2^{-m}) 2^{-m} \\ &\times \left((\epsilon^2 + (1-\epsilon)^2)^n - (1-\epsilon)^{2n} \right). \end{aligned} \tag{41}$$

The last equality is the claim of the theorem. □

Example 4: Table I includes the weight distributions and undetected error probabilities for the 4-instances in $\mathcal{R}_{1,2}$. Since

TABLE I
WEIGHT DISTRIBUTIONS AND UNDETECTED ERROR PROBABILITIES

H	$C(H)$	$A_1(H)$	$A_2(H)$	$P_U(H)$
(0,0)	$\{00, 01, 10, 11\}$	2	1	$2\epsilon - \epsilon^2$
(0,1)	$\{00, 10\}$	1	0	$\epsilon - \epsilon^2$
(1,0)	$\{00, 01\}$	1	0	$\epsilon - \epsilon^2$
(1,1)	$\{00, 11\}$	0	1	ϵ^2

equal probability is assigned to each matrix, the average of P_U can be written as

$$\begin{aligned} E_{\mathcal{R}_{1,2}}[P_U] &= \frac{(2\epsilon - \epsilon^2) + 2(\epsilon - \epsilon^2) + \epsilon^2}{4} \\ &= \epsilon - \frac{1}{2}\epsilon^2. \end{aligned} \tag{42}$$

On the other hand, from Lemma 1, we have

$$\begin{aligned} E_{\mathcal{R}_{1,2}}[P_U] &= 2^{-1}(1 - (1-\epsilon)^2) \\ &= \epsilon - \frac{1}{2}\epsilon^2, \end{aligned} \tag{43}$$

which is identical to (42).

We then consider the variance. From Table I, it is easy to compute the second moment of P_U in such a way:

$$\begin{aligned} E_{\mathcal{R}_{1,2}}[P_U^2] &= \frac{(2\epsilon - \epsilon^2)^2 + 2(\epsilon - \epsilon^2)^2 + (\epsilon^2)^2}{4} \\ &= \frac{3}{2}\epsilon^2 - 2\epsilon^3 + \epsilon^4. \end{aligned} \tag{44}$$

Subtracting the squared first moment from the second moment, we get the variance:

$$\begin{aligned}\sigma^2_{\mathcal{R}_{1,2}} &= E_{\mathcal{R}_{1,2}}[P_U^2] - E_{\mathcal{R}_{1,2}}[P_U]^2 \\ &= \frac{3}{2}\epsilon^2 - 2\epsilon^3 + \epsilon^4 - \left(\epsilon - \frac{1}{2}\epsilon^2\right)^2 \\ &= \frac{1}{2}\epsilon^2 - \epsilon^3 + \frac{3}{4}\epsilon^4. \end{aligned} \quad (45)$$

Note that Theorem 2 gives

$$\begin{aligned}\sigma^2_{\mathcal{R}_{1,2}} &= (1-2^{-1})2^{-1}\left((\epsilon^2+(1-\epsilon)^2)^2 - (1-\epsilon)^4\right) \\ &= \frac{1}{2}\epsilon^2 - \epsilon^3 + \frac{3}{4}\epsilon^4, \end{aligned} \quad (46)$$

which is exactly the same result as (45). □

3) Concentration to average: The variance derived Theorem 2 can be used to show the following concentration result.

Corollary 1: The ratio of P_U and $E_{\mathcal{R}_{m,n}}[P_U]$ converges to 1 in probability, namely,

$$\frac{P_U}{E_{\mathcal{R}_{m,n}}[P_U]} \to 1 \quad \text{in probability} \quad (47)$$

as n goes to infinity if $\epsilon (0 < \epsilon < 1/2)$ satisfies

$$1 - R + \log_2(\epsilon^2 + (1-\epsilon)^2) < 0. \quad (48)$$

(Proof) Let $\mu \triangleq E_{\mathcal{R}_{m,n}}[P_U]$ and $\sigma \triangleq \sigma_{\mathcal{R}_{m,n}}$. From Chebyshev inequality, we have

$$Pr\left[\frac{P_U}{\mu} \in (1-\alpha, 1+\alpha)\right] \le \frac{\sigma^2}{\alpha^2\mu^2}, \quad (49)$$

where α is a positive real number. If the equation

$$\lim_{n\to\infty} \frac{\sigma^2}{\mu^2} = 0 \quad (50)$$

holds, then the right hand side of inequality (49) converges to 0 as n goes to infinity regardless of choice of α. This means P_U/μ converges to 1 in probability.

We now are going to discuss the asymptotic behavior of the ratio σ^2/μ^2. The ratio can be rewritten into the following form:

$$\begin{aligned}\frac{\sigma^2}{\mu^2} &= \frac{(1-2^{-m})2^{-m}\left((\epsilon^2+(1-\epsilon)^2)^n - (1-\epsilon)^{2n}\right)}{2^{-2m}(1-(1-\epsilon)^n)^2} \\ &= \frac{(2^m-1)\left((\epsilon^2+(1-\epsilon)^2)^n - (1-\epsilon)^{2n}\right)}{(1-(1-\epsilon)^n)^2} \\ &\le \frac{2^{(1-R)n}(\epsilon^2+(1-\epsilon)^2)^n}{(1+o(1))^2}. \end{aligned} \quad (51)$$

From the above inequatliy, we get

$$\begin{aligned}\lim_{n\to\infty}\frac{\sigma^2}{\mu^2} &\le \lim_{n\to\infty} 2^{(1-R)n}(\epsilon^2+(1-\epsilon)^2)^n \quad (52) \\ &= \lim_{n\to\infty} 2^{n(1-R+\log_2(\epsilon^2+(1-\epsilon)^2))}. \quad (53)\end{aligned}$$

We can see that σ^2/μ^2 converges to zero if the exponent $1 - R + \log_2(\epsilon^2 + (1-\epsilon)^2)$ takes a negative value. □

Let ϵ^* be the root of the equation

$$1 - R + \log_2(\epsilon^{*2} + (1-\epsilon^*)^2) = 0. \quad (54)$$

Table II presents some values of ϵ^*. When $\epsilon > \epsilon^*$, we have $1 - R + \log_2(\epsilon^{*2} + (1-\epsilon^*)^2) < 0$. In such a region, P_U concentrates its average as n grows to infinity.

TABLE II
ROOTS OF $1 - R + \log_2(\epsilon^{*2} + (1-\epsilon^*)^2) = 0$

R	ϵ^*
0.1	0.366047
0.2	0.307193
0.3	0.259613
0.4	0.217375
0.5	0.178203
0.6	0.140933
0.7	0.104872
0.8	0.069564
0.9	0.034687

B. Variance of undetected error probability: sparse matrix ensemble

1) Covariance formula: The covariance of the weight distribution for the sparse matrix ensemble is given in the following lemma.

Lemma 4: The covariance of weight distribution for sparse matrix ensemble $\mathcal{T}_{m,n,k}$ is given by

$$\mathrm{Cov}_{\mathcal{T}_{m,n,k}}(A_{w_1}, A_{w_2}) = \psi(w_1, w_2), \quad (55)$$

for $1 \le w_1, w_2 \le n$. The function $\psi(w_1, w_2)$ is defined by

$$\begin{aligned}&\psi(w_1,w_2) \\ &\triangleq \left(\frac{1+x^{w_1}}{2}\right)^m \left(\frac{1+x^{w_2}}{2}\right)^m \\ &\times \sum_{j=1}^{w_1} \binom{n}{w_1}\binom{w_1}{j}\binom{n-w_1}{w_2-j}\left(\xi^m_{w_1,w_2,j} - 1\right), \end{aligned} \quad (56)$$

if $1 \le w_1 \le w_2 \le n$. If $1 \le w_2 < w_1 \le n$, $\psi(w_1, w_2)$ is defined by

$$\psi(w_1, w_2) \triangleq \psi(w_2, w_1). \quad (57)$$

The symbol $\xi_{w_1,w_2,j}$ represents

$$\xi_{w_1,w_2,j} \triangleq 1 - \frac{x^{w_1+w_2} - x^{w_1+w_2-2j}}{(1+x^{w_1})(1+x^{w_2})} \quad (58)$$

for $1 \le w_1 \le w_2 \le n$, $0 \le j \le w_1$.

(Proof) See the forthcoming full paper version. □

Remark 3: When $k = n/2$, sparse matrix ensemble coincides with random ensemble because $p = 1/2$ means $P(H) = 1/2^{mn}$ for any H. We discuss this case here.

To simplify the discussion, we assume that $1 \le w_1 \le w_2 \le n$. Let $p = 1/2$ (i.e., $k = n/2$). In such a case, we have $x = 1 - 2p = 0$ and $\xi_{w_1,w_2,j}$ takes the following value:

$$\xi_{w_1,w_2,j} = \begin{cases} 1 & w_1 < w_2 \\ 1 & w_1 = w_2, j < w_1 \\ 2 & w_1 = w_2, j = w_1. \end{cases} \quad (59)$$

Substituting $x = 0$ into (56), we get

$$\mathrm{Cov}(A_{w_1}, A_{w_2}) = \begin{cases} 0, & 1 \le w_1 < w_2 \le n \\ 2^{-2m}\binom{n}{w}(2^m - 1), & 1 \le w_1 = w_2 \le n. \end{cases} \tag{60}$$

These equations coincide with the covariance of random ensemble given in Lemma 3. □

2) Variance of undetected error probability: The variance of undetected error probability is a straightforward consequence of Lemma 4.

Theorem 3: The variance of undetected error probability of sparse matrix ensemble, $\sigma^2_{\mathcal{T}_{m,n,k}}$ is given by

$$\sigma^2_{\mathcal{T}_{m,n,k}} = \sum_{w_1=1}^{n}\sum_{w_2=1}^{n} \psi(w_1, w_2)\epsilon^{w_1+w_2}(1-\epsilon)^{2n-w_1-w_2}. \tag{61}$$

(Proof) From Lemma 4, the claim of the lemma follows in the following way:

$$\begin{aligned} &\sigma^2_{\mathcal{T}_{m,n,k}} \\ &= \sum_{w_1=1}^{n}\sum_{w_2=1}^{n} \mathrm{Cov}_{\mathcal{T}_{m,n,k}}(A_{w_1}, A_{w_2})\epsilon^{w_1+w_2}(1-\epsilon)^{2n-w_1-w_2} \\ &= \sum_{w_1=1}^{n}\sum_{w_2=1}^{n} \psi(w_1, w_2)\epsilon^{w_1+w_2}(1-\epsilon)^{2n-w_1-w_2}. \end{aligned} \tag{62}$$

□

Example 5: Let us consider the sparse matrix ensemble with $m = 1, n = 2$ and $k = 1/2 (p = 1/4)$. From the definition of sparse matrix ensemble, the following probability is assigned for each matrix: $P((0,0)) = 9/16, P((0,1)) = 3/16, P((1,0)) = 3/16, P((1,1)) = 1/16$. Combining the undetected error probabilities presented in Table I and the above probability assignment, we immediately have the first and second moment:

$$E_{\mathcal{T}_{1,2,1/2}}[P_U] = \frac{2}{3}\epsilon - \frac{7}{8}\epsilon^2 \tag{63}$$

$$E_{\mathcal{T}_{1,2,1/2}}[P_U^2] = \frac{21}{8}\epsilon^2 - \frac{3}{8}\epsilon^3 + \epsilon^4. \tag{64}$$

From these moments, the variance can be derived:

$$\begin{aligned} \sigma^2_{\mathcal{T}_{1,2,1/2}} &= E_{\mathcal{T}_{1,2,1/2}}[P_U^2] - E_{\mathcal{T}_{1,2,1/2}}[P_U]^2 \\ &= \frac{3}{8}\epsilon^2 - \frac{3}{8}\epsilon^3 + \frac{15}{64}\epsilon^4. \end{aligned} \tag{65}$$

We then consider another route to derive variance. From the definition of ψ in (56), we have

$$\psi(1,1) = 3/8 \tag{66}$$

$$\psi(1,2) = \psi(2,1) = 3/16 \tag{67}$$

$$\psi(2,2) = 15/64. \tag{68}$$

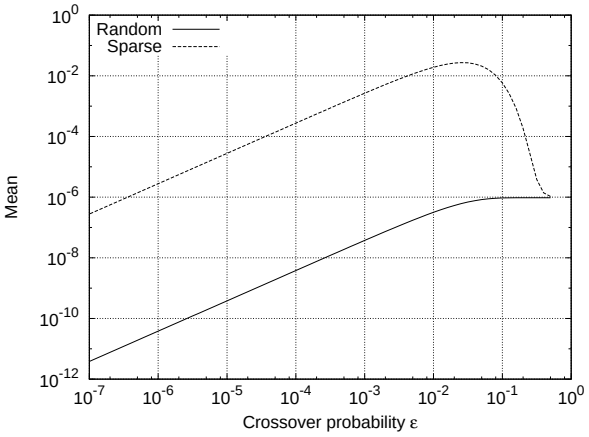
From Theorem 3, we obtain the variance

$$\begin{aligned} \sigma^2_{\mathcal{T}_{1,2,1/2}} &= \sum_{w_1=1}^{2}\sum_{w_2=1}^{2} \psi(w_1, w_2)\epsilon^{w_1+w_2}(1-\epsilon)^{4-w_1-w_2} \\ &= \psi(1,1)\epsilon^2(1-\epsilon)^2 + \psi(1,2)\epsilon^3(1-\epsilon)^1 \\ &+ \psi(2,1)\epsilon^3(1-\epsilon)^1 + \psi(2,2)\epsilon^4(1-\epsilon)^0 \\ &= (3/8)\epsilon^2(1-\epsilon)^2 + (3/16)\epsilon^3(1-\epsilon) \\ &+ (3/16)\epsilon^3(1-\epsilon) + (15/64)\epsilon^4 \\ &= \frac{3}{8}\epsilon^2 - \frac{3}{8}\epsilon^3 + \frac{15}{64}\epsilon^4, \end{aligned}$$

which is identical to (65). □

The next example would help us to understand how mean and variance of P_U behave.

Example 6: We here consider random ensemble with $m = 20, n = 40$ and sparse matrix ensemble with $m = 20, n = 40, k = 5$. Figure 5 presents average (mean) undetected error probabilities of two ensembles. It can be observed that the average undetected error probability of random ensemble monotonically decreases as ϵ gets small. On the other hand, the curve for the sparse matrix ensemble has peak around $\epsilon = 0.025$. Figure 6 shows variance of P_U for the above



Random ensemble: $m = 20$, $n = 40$, Sparse matrix ensemble: $m = 20$, $n = 40$, $k = 5$.

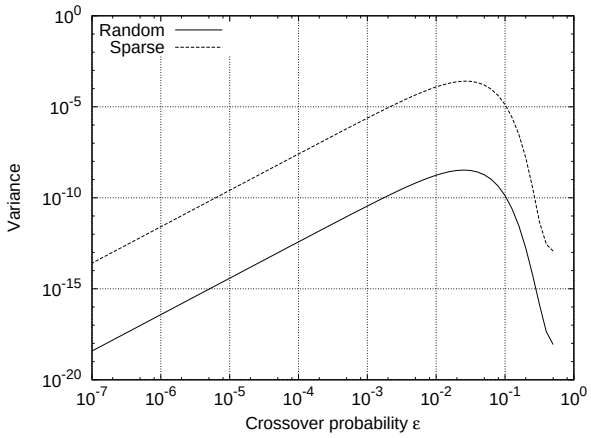
Fig. 5. Average undetected error probabilities

two ensembles. Two curves have similar shape but variance of the sparse ensemble is always larger than that of the random ensemble. □

3) Asymptotic behavior: We here discuss the asymptotic behaviors of covariance of weight distribution and variance of P_U for sparse matrix ensemble. The following corollary explains the asymptotic behavior of covariance of weight distribution which is a consequence of Lemma 4.

Corollary 2: For $0 < \ell_1 \le \ell_2 \le 1$, the equality

$$\lim_{n\to\infty} \frac{1}{n}\log_2 \psi(\ell_1 n, \ell_2 n) = \sup_{0<\kappa\le\ell_1} L(\ell_1, \ell_2, \kappa), \tag{69}$$



Random ensemble: $m = 20$, $n = 40$, Sparse matrix ensemble: $m = 20$, $n = 40$, $k = 5$.

Fig. 6. Variance of undetected error probability

holds where $L(\ell_1, \ell_2, \kappa)$ is defined by

$$\begin{aligned} & L(\ell_1, \ell_2, \kappa) \\ &= -2(1-R) + h(\ell_1) + h\left(\frac{\kappa}{\ell_1}\right) + h\left(\frac{\ell_2 - \kappa}{1-\ell_1}\right) \\ &+ (1-R)\log_2\left(1 + e^{-2k\ell_1} + e^{-2k\ell_2} + e^{-2k(\ell_1+\ell_2-2\kappa)}\right). \end{aligned}$$

(Proof) Let us assume that $0 < w_1 \leq w_2$. In this case, $\psi(w_1, w_2)$ defined in (56) can be rewritten into the following form:

$$\begin{aligned} & \psi(w_1, w_2) \\ &= \left(\frac{1+x^{w_1}}{2}\right)^m \left(\frac{1+x^{w_2}}{2}\right)^m \\ &\times \sum_{j=1}^{w_1} \binom{n}{w_1}\binom{w_1}{j}\binom{n-w_1}{w_2-j}\left(\xi_{w_1,w_2,j}^m - 1\right) \\ &= 2^{-2m}\sum_{j=1}^{w_1} \binom{n}{w_1}\binom{w_1}{j}\binom{n-w_1}{w_2-j} \\ &\times (1 + x^{w_1} + x^{w_2} + x^{w_1+w_2-2j})^m(1-\delta), \end{aligned} \tag{70}$$

where δ is defined by

$$\delta \triangleq \left(\frac{1 + x^{w_1} + x^{w_2} + x^{w_1+w_2}}{1 + x^{w_1} + x^{w_2} + x^{w_1+w_2-2j}}\right)^m. \tag{71}$$

In the above derivation, the following identity was used:

$$\begin{aligned} & \xi_{w_1,w_2,j} \\ &= 1 - \frac{x^{w_1+w_2} - x^{w_1+w_2-2j}}{(1+x^{w_1})(1+x^{w_2})} \\ &= \frac{(1+x^{w_1})(1+x^{w_2}) - x^{w_1+w_2} + x^{w_1+w_2-2j}}{(1+x^{w_1})(1+x^{w_2})} \\ &= \frac{1 + x^{w_1} + x^{w_2} + x^{w_1+w_2-2j}}{(1+x^{w_1})(1+x^{w_2})}. \end{aligned} \tag{72}$$

Note that

$$\frac{1 + x^{w_1} + x^{w_2} + x^{w_1+w_2}}{1 + x^{w_1} + x^{w_2} + x^{w_1+w_2-2j}} < 1 \tag{73}$$

holds when $j > 0$. This is because $x = 1 - 2k/n < 1$.

Letting $w_1 = \ell_1 n, w_2 = \ell_2 n, m = (1-R)n$ and using (70) we have an upper bound $(1/n)\log_2 \psi(\ell_1 n, \ell_2 n)$:

$$\begin{aligned} & \frac{1}{n}\log_2 \psi(\ell_1 n, \ell_2 n) \\ &\leq -2(1-R) + \frac{\log_2(\ell_1 n)}{n} \\ &+ \max_{j=1}^{\ell_1 n} \frac{1}{n}\log_2\left(\binom{n}{\ell_1 n}\binom{\ell_1 n}{j}\binom{n-\ell_1 n}{\ell_2 n - j}\right) \\ &+ (1-R)\log_2(1 + x^{\ell_1 n} + x^{\ell_2 n} + x^{\ell_1 n + \ell_2 n - 2j}) \\ &+ \frac{1}{n}\log_2(1-\delta). \end{aligned} \tag{74}$$

It is obvious that the following equations hold:

$$\lim_{n\to\infty} \frac{\log_2(\ell_1 n)}{n} = 0, \tag{75}$$

$$\begin{aligned} & \lim_{n\to\infty} \frac{1}{n}\log_2\left(\binom{n}{\ell_1 n}\binom{\ell_1 n}{j}\binom{n-\ell_1 n}{\ell_2 n - j}\right) \\ &= h(\ell_1) + h\left(\frac{\kappa}{\ell_1}\right) + h\left(\frac{\ell_2 - \kappa}{1-\ell_1}\right), \end{aligned} \tag{76}$$

where κ is a real number satisfying $0 < \kappa \leq \ell_1$ and $j = \kappa n$. If k is a constant and $0 \leq \ell \leq 1$, the equation

$$\begin{aligned} \lim_{n\to\infty}\left(1 - 2\left(\frac{k}{n}\right)\right)^{\ell n} &= \lim_{n\to\infty} x^{\ell n} \\ &= e^{-2k\ell}, \end{aligned} \tag{77}$$

holds [3] and it gives the following equality:

$$\begin{aligned} & \lim_{n\to\infty}(1-R)\log_2(1 + x^{\ell_1 n} + x^{\ell_2 n} + x^{\ell_1 n + \ell_2 n - 2j}) \\ &= (1-R) \\ &\times \log_2(1 + e^{-2k\ell_1} + e^{-2k\ell_2} + e^{-2k(\ell_1+\ell_2-2\kappa)}). \end{aligned} \tag{78}$$

Finally, from inequality (73), we have

$$\frac{1}{n}\log_2(1-\delta) = 0. \tag{79}$$

Applying these equations to inequality (74), we get

$$\lim_{n\to\infty}\frac{1}{n}\log_2\psi(\ell_1 n, \ell_2 n) \leq \sup_{0<\kappa\leq\ell_1} L(\ell_1, \ell_2, \kappa). \tag{80}$$

On the other hand, in a similar way, we can prove the the opposite direction inequality as well:

$$\lim_{n\to\infty}\frac{1}{n}\log_2\psi(\ell_1 n, \ell_2 n) \geq \sup_{0<\kappa\leq\ell_1} L(\ell_1, \ell_2, \kappa). \tag{81}$$

Combining these two inequalities, we obtain the claim of the corollary. □

We here extend the definition of $L(\ell_1, \ell_2, \kappa)$ in order to make it consistent with the definition of $\psi(w_1, w_2)$:

$$L(\ell_1, \ell_2, \kappa) \triangleq L(\ell_2, \ell_1, \kappa) \tag{82}$$

if $\ell_1 > \ell_2$. The following corollary gives the asymptotic growth rate of the $\sigma^2_{\mathcal{I}_{(1-R)n,n,k}}$.

Corollary 3: The asymptotic growth rate of the variance of undetected error is given by

$$\lim_{n\to\infty}\frac{1}{n}\log_2\sigma^2_{\mathcal{T}_{(1-R)n,n,k}} = \sup_{0<\ell_1\le 1}\sup_{0<\ell_2\le 1}\sup_{0<\kappa\le\ell_1} U(\ell_1,\ell_2,\kappa), \tag{83}$$

where $U(\ell_1,\ell_2,\kappa)$ is given by

$$\begin{aligned} U(\ell_1,\ell_2,\kappa) &= (\ell_1+\ell_2)\log_2\epsilon + (2-\ell_1-\ell_2)\log_2(1-\epsilon) \\ &+ L(\ell_1,\ell_2,\kappa). \end{aligned} \tag{84}$$

(Proof) Applying Corollary 2 to Theorem 3, we obtain

$$\begin{aligned} &\lim_{n\to\infty}\frac{1}{n}\log_2\sigma^2_{\mathcal{T}_{m,n,k}} \\ &= \sup_{0<\ell_1\le 1}\sup_{0<\ell_2\le 1}[\lim_{n\to\infty}\frac{1}{n}\log_2\psi(\ell_1 n,\ell_2 n) \\ &+ \lim_{n\to\infty}\frac{1}{n}\log_2\epsilon^{\ell_1 n+\ell_2 n}(1-\epsilon)^{2n-\ell_1 n-\ell_2 n} \\ &= \sup_{0<\ell_1\le 1}\sup_{0<\ell_2\le 1}[\sup_{0\kappa\le\ell_1} L(\ell_1,\ell_2,\kappa) \\ &+ (\ell_1+\ell_2)\log_2\epsilon + (2-\ell_1-\ell_2)\log_2(1-\epsilon)]. \end{aligned} \tag{85}$$

□

ACKNOWLEDGMENT

This work was partly supported by the Ministry of Education, Science, Sports and Culture, Japan, Grant-in-Aid for Scientific Research on Priority Areas (Deepening and Expansion of Statistical Informatics) 180790091.

REFERENCES

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

[2] T. Klove and V. Korzhik, *"Error Detecting Codes: General Theory and Their Application in Feedback Communication Systems"*, Kluwer Academic, 1995.

[3] S.Litsyn and V. Shevelev, "On ensembles of low-density parity-check codes: asymptotic distance distributions," *IEEE Trans. Inform. Theory*, vol.48, pp.887–908, Apr. 2002.

[4] S.Litsyn and V. Shevelev, "Distance distributions in ensembles of irregular low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol.49, pp.3140–3159, Nov. 2003.

[5] D.Burshtein and G. Miller, "Asymptotic enumeration methods for analyzing LDPC codes," *IEEE Trans. Inform. Theory*, vol.50, pp.1115–1131, June 2004.

[6] O. Barak, D. Burshtein, "Lower bounds on the spectrum and error rate of LDPC code ensembles," in Proceedings of International Symposium on Information Theory, 2005.

[7] V. Rathi, "On the Asymptotic Weight Distribution of Regular LDPC Ensembles," in Proceedings of International Symposium on Information Theory, 2005.

[8] T. Richardson, R. Urbanke, "Modern Coding Theory," online: http://lthcwww.epfl.ch/