On Undetected Error Probability of Binary Matrix Ensembles

Tadashi Wadayama†

arXiv:0705.3995v1 [cs.IT] 28 May 2007

***Abstract*— In this paper, an analysis of the undetected error probability of ensembles of $m \times n$ binary matrices is presented. Two ensembles are considered: One is an ensemble of dense matrices, while the other is an ensemble of sparse matrices. The main contributions of this work are (i) derivation of the error exponent of the average undetected error probability and (ii) closed form expressions for the variance of the undetected error probability. It is shown that the behavior of the exponent for a sparse ensemble is somewhat different from that for a dense ensemble. The variance of undetected error probability leads to a concentration result. Furthermore, as a byproduct of the proof of the variance formulae, simple covariance formulae of the weight distribution are derived.**

I. Introduction

Random coding is an extremely powerful technique to show the existence of a code satisfying certain properties. It has been used for proving the direct part (achievability) of many types of coding theorems. Recently, the idea of random coding has also come to be regarded as important from a 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 is a growing interest in randomly generated codes.

One of the main difficulties associated with the use of randomly generated codes is the difficulty in evaluating the properties or performance of such codes. For example, it is difficult to evaluate minimum distance, weight distribution, ML decoding performance, etc. for these codes. To overcome this problem, we can take a *probabilistic approach*. In such an approach, we consider an ensemble of parity check matrices: i.e., probability is assigned to each matrix in the ensemble. A property of a matrix (e.g., minimum distance, weight distributions) can then 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 around its expectation. Such a concentration result justifies the use of the probabilistic approach.

Recent advances in the analysis of average weight distributions of LDPC codes, such as those described by Litsyn and Shevelev [3][4], Burshtein and Miller [5] Richardson and Urbanke [8], show that the probabilistic approach is a useful technique for investigating typical properties of codes and matrices, which are not easy to obtain. Furthermore, the second moment analysis of the weight distribution of LDPC codes [6][7] can be utilized to prove concentration results for weight distributions.

The evaluation of 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 a probabilistic approach. For example, for a linear code ensemble the inequality $P_U < 2^{-m}$, has long been 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 the weight distribution of a code, there is a natural connection between the expectation of the weight distribution and the expectation of the undetected error probability.

In this paper, an analysis of the undetected error probability of ensembles of binary matrices of size $m \times n$ is presented. Two ensembles are considered: One is an ensemble of dense matrices, called a *random ensemble*, while the other 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(Automatic Repeat reQuest). Detailed knowledge of the error detection performance of a matrix ensemble would be useful for assessing the performance of a feedback error correction scheme.

The contents of this paper are arranged 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 formulae for the random ensemble and the sparse matrix ensemble are derived based on a combinatorial approach.

II. Average undetected error probability

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

A. Notation

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, such as $\boldsymbol{x}$ for example, denotes a binary row vector.

†Nagoya Institute of Technology, email:wadayama@nitech.ac.jp. A part of this work was presented at ITA workshop in UCSD, Feb. 2007.

Throughout the paper, a binary symmetric channel (BSC) with crossover probability ϵ ($0 < \epsilon < 1/2$) is assumed. We assume the conventional scenario for error detection: A transmitter sends a codeword $\boldsymbol{x} \in C(H)$ to a receiver via a BSC with crossover probability ϵ. The receiver obtains 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 checks 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 evaluates to 0.

Suppose that $\mathcal{G}$ is a set of binary $m \times n$ matrices $(m, n \geq 1)$. Note that $\mathcal{G}$ may contain some matrices with all elements identical. Such matrices should be distinguished as distinct matrices. A probability $P(H)$ is associated with each matrix H in $\mathcal{G}$. 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 analyzing the performance of binary linear codes, including analysis of the undetected error probability.

B. Binary matrix ensembles

Attention is focused on two types of ensemble 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 matrices of size $m \times n$ for $(m, n \geq 1)$. From this definition, it is evident that the size of $\mathcal{R}_{m,n}$ is 2^{mn}. For each matrix in $\mathcal{R}_{m,n}$, an 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 $[a, b]$ denotes the set of consecutive integers from a to b. Since 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 $(m, n \geq 1)$, whose elements 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 first 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 in general computationally difficult, 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 an average probability is useful for the estimation of the undetected error probability of a matrix which belongs to the ensemble.

Taking the 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 a 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 an exponentially decreasing function of n, where R is a real number satisfying $0 < R < 1$ (called the *design rate*). Thus, the exponent of the undetected error probability is of prime importance in understanding 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

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

if the limit exists. □

Henceforth we will not explicitly express the dependence of P_U on H, writing instead P_U to denote $P_U(H)$ in all cases where there is no fear of confusion.

The following example describes the exponent of one 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 henceforth abbreviated as the asymptotic growth rate), which is the 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

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

The parameter ℓ is called *normalized weight.* □

From this definition, it is clear that

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

where the notation $o(1)$ denotes terms which converge to 0 in the limit 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 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

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

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 the asymptote of $T_{\mathcal{G}_n}$ in the limit as n tends to infinity.

□

The next example discusses the case of a random ensemble.

Example 2: Let us again consider the series of random ensembles given by $\{\mathcal{R}_{(1-R)n,n}\}_{n>0}$. These ensembles have the asymptotic growth rate $f(\ell) = h(\ell) - (1-R)$, where 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, with respect 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

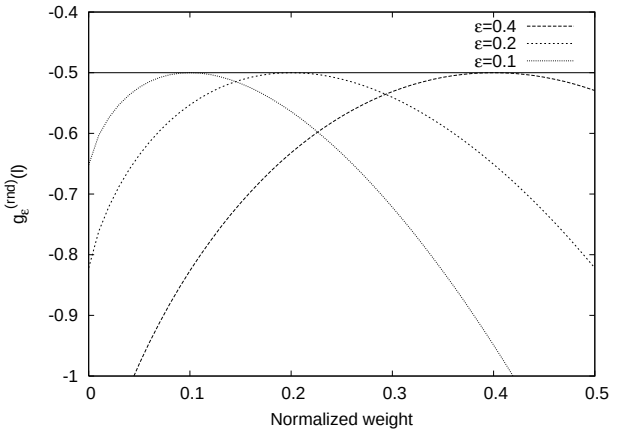
$$T_{\mathcal{R}_{(1-R)n,n}} = \sup_{0<\ell\leq 1}[-(1-R) - D_{\ell,\epsilon}]. \tag{23}$$

Since $D_{\ell,\epsilon}$ can be considered as the 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 obtain

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

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

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

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

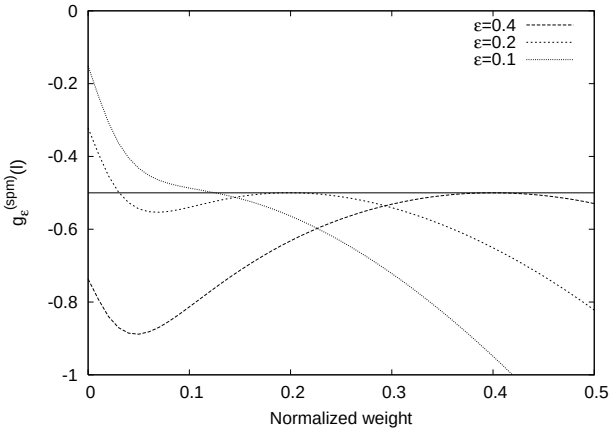
The error exponent of this ensemble shows quite a different behavior from that for random ensembles.

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

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

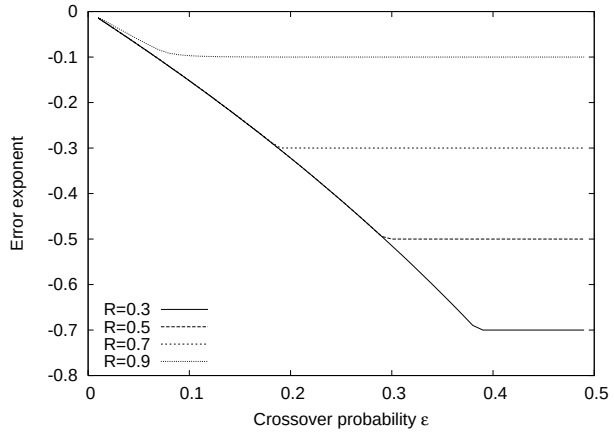
Figure 2 includes the curves of $g_\epsilon^{(spm)}(\ell)$ where $\epsilon = 0.1, 0.2, 0.4$. In contrast to $g_\epsilon^{(rnd)}(\ell)$ of a random ensemble, we can see that $g_\epsilon^{(spm)}(\ell)$ is not a concave function. The shape of the curve of $g_\epsilon^{(spm)}(\ell)$ depends on the crossover probability ϵ. For large ϵ, $g_\epsilon(\ell)$ takes its largest value around $\ell = \epsilon$. On the other hand, for small ϵ, $g_\epsilon^{(spm)}(\ell)$ has a supremum at $\epsilon = 0$.

Figure 3 presents the error exponent of sparse matrix ensembles with parameters $R = 0.3, 0.5, 0.7, 0.9$ and $k = 20$. As 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 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.

The examples suggest that a sparse ensemble has a less powerful error detection performance than that of a dense ensemble (such as random ensemble) in terms of the error exponent. However, if the crossover probability is sufficiently large, the difference in exponent of sparse and dense ensembles is negligible. 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 of the error exponents of sparse matrix ensembles can be explained with reference to their average weight distributions (or asymptotic growth rate). Figure 4 displays the asymptotic growth rates of a random ensemble and a sparse matrix ensemble.

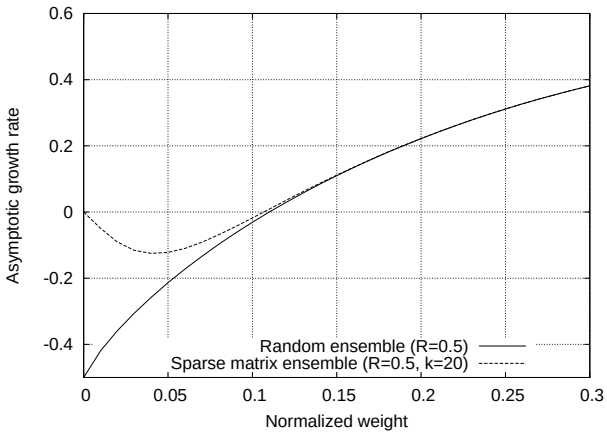


Fig. 4. Asymptotic growth rate of a random ensemble and a sparse matrix ensemble.

The weight of typical error vectors is very close to ϵn when n is sufficiently large. For a large value of ϵ, 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 distributions corresponding to random and sparse matrix ensembles is 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 a difference in the error exponent.

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

The following lemma is useful in understanding the behavior of the 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 the following 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 to examine the *covariance of weight distribution* in order to handle the variance of undetected error probability.

Definition 3: For $0 \leq w_1, w_2 \leq n$ and a given ensemble $\mathcal{G}$, the 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 forms the basis of the derivation of the variance of the undetected error probability for a random ensemble.

Lemma 3: For a random ensemble $\mathcal{R}_{m,n}$, the 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 Appendix. □

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 been shown in [8]. Thus, the new contribution of this lemma is the case $\mathrm{Cov}_{\mathcal{R}_{m,n}}(A_{w_1}, A_{w_2}) = 0$ when $w_1 \neq w_2$. □

Remark 2: The covariance of the weight distribution for a given ensemble $\mathcal{G}$ is useful not only for the evaluation of the 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 the weight distribution is required more generally for the evaluation of the variance of X, which is given by

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

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 the variance of the 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

$$\sigma^2_{\mathcal{R}_{m,n}} \triangleq E_{\mathcal{R}_{m,n}}[(P_U - \mu)^2]. \quad (36)$$

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

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

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

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

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

The squared average undetected error probability can be expressed as

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

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

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

The last equalities are due to Lemma 3. We can further simplify the expression using the binomial theorem,

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

The last equality is the claim of the theorem. □

Example 4: Table I displays the weight distributions and undetected error probabilities for the 4 matrices 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

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

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

On the other hand, from Lemma 1, we have

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

which is identical to expression (42).

We now consider the variance. From Table I, it is easy to compute the second moment of P_U,

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

Subtracting the squared first moment from the second moment, we obtain 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 yields

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

which is identical to expression (45). □

3) Concentration to average: The variance derived in 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,

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

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

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

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

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

where α is a positive real number. If the equation

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

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

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

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

From the above inequality, we get

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

Thus it is clear 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

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

Table II presents some values of ϵ^* for values of R from 0.1 to 0.9. When $\epsilon > \epsilon^*$, we have $1 - R + \log_2(\epsilon^{*2} + (1-\epsilon^*)^2) < 0$. In such a region, P_U concentrates around its average value in the limit as n tends 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 a sparse matrix ensemble is given in the following lemma.

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

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

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

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

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

$$\psi(w_1, w_2) \stackrel{\triangle}{=} \psi(w_2, w_1). \tag{57}$$

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

$$\xi_{w_1,w_2,j} \stackrel{\triangle}{=} 1 - \frac{x^{w_1+w_2} - x^{w_1+w_2-2j}}{(1+x^{w_1})(1+x^{w_2})} \tag{58}$$

for $1 \le w_1 \le w_2 \le n$, $0 \le j \le w_1$.
(Proof) See Appendix. □

Remark 3: When $k = n/2$, a sparse matrix ensemble coincides with a random ensemble because $p = 1/2$ implies $P(H) = 1/2^{mn}$ 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 values:

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

Substituting $x = 0$ into equation (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 a random ensemble as given in Lemma 3. □

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

Theorem 3: The variance of the undetected error probability of a 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 as

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

□

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 a sparse matrix ensemble, the following probability is assigned to 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 moments:

$$\begin{aligned}E_{\mathcal{T}_{1,2,1/2}}[P_U] &= \frac{2}{3}\epsilon - \frac{7}{8}\epsilon^2 \qquad (63)\\ E_{\mathcal{T}_{1,2,1/2}}[P_U^2] &= \frac{21}{8}\epsilon^2 - \frac{3}{8}\epsilon^3 + \epsilon^4. \qquad (64)\end{aligned}$$

From these moments, the variance can be derived,

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

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

$$\begin{aligned}\psi(1,1) &= 3/8 \qquad (66)\\ \psi(1,2) &= \psi(2,1) = 3/16 \qquad (67)\\ \psi(2,2) &= 15/64. \qquad (68)\end{aligned}$$

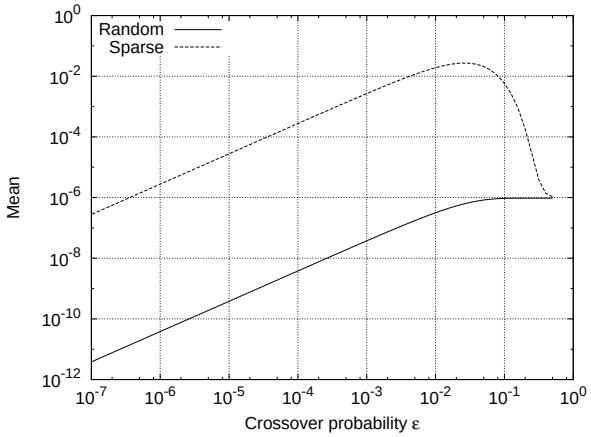
From Theorem 3, we obtain the variance

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

which is identical to expression (65). □

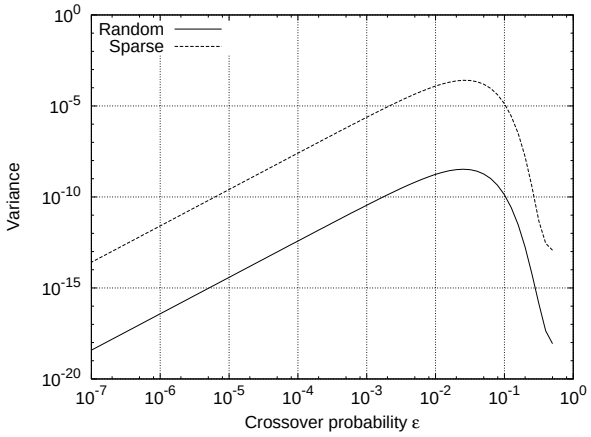
The next example facilitates an understanding of how the average and variance of P_U behave.

Example 6: We consider a random ensemble with $m = 20, n = 40$, and a sparse matrix ensemble with $m = 20, n = 40, k = 5$. Figure 5 depicts the average undetected error probabilities of the two ensembles. It can be observed that the average undetected error probability of the random ensemble monotonically decreases as ϵ decreases. In contrast, the curve for the sparse matrix ensemble has a peak around $\epsilon \simeq 0.025$. Figure 6 shows the variance of P_U for the above two ensembles. The two curves have a similar shape, but the variance of the sparse ensemble is always larger than that of the random ensemble. □



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

Fig. 5. Average undetected error probabilities.



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

Fig. 6. Variance of undetected error probability.

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

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

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

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 $0 < w_1 \le w_2$. In this case, $\psi(w_1, w_2)$ defined in equation (56) can be rewritten in the

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

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

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 equation (70), we can derive an upper bound for $(1/n)\log_2\psi(\ell_1 n,\ell_2 n)$:

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

It is straightforward to see that the following limits are obtained:

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

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

where κ is a real number satisfying $0<\kappa\le\ell_1$ and $j=\kappa n$. If k is a constant and $0\le\ell\le 1$, then, making use of Litsyn and Shevelev's [3] result that

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

we have

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

Finally, from inequality (73), we get

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

Applying these equations to inequality (74), we get

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

On the other hand, in a similar way, we can also prove that

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

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

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

$$L(\ell_1,\ell_2,\kappa)\stackrel{\triangle}{=}L(\ell_2,\ell_1,\kappa) \tag{82}$$

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

Corollary 3: The asymptotic growth rate of the variance of the 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

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

□

IV. Appendix

A. The proof of Lemma 3(Covariance for random ensemble)

1) Preparation of the proof: The second moment of the weight distribution for a given ensemble $\mathcal{G}$ is given by

$$\begin{aligned}E_{\mathcal{G}}&\left[A_{w_1}A_{w_2}\right]\\ &= E_{\mathcal{G}}\left[\sum_{\boldsymbol{x}\in Z^{(n,w_1)}}\sum_{\boldsymbol{y}\in Z^{(n,w_2)}}I[H\boldsymbol{x}^t=\boldsymbol{0}]I[H\boldsymbol{y}^t=\boldsymbol{0}]\right]. \end{aligned}$$

for $0 < w_1, w_2 \leq n$. Since

$$I[H\boldsymbol{x}^t = \boldsymbol{0}]I[H\boldsymbol{y}^t = \boldsymbol{0}] = I[H\boldsymbol{x}^t = \boldsymbol{0}, H\boldsymbol{y}^t = \boldsymbol{0}],$$

we have

$$\begin{aligned} &E_{\mathcal{G}}\left[A_{w_1}A_{w_2}\right] \\ &= E_{\mathcal{G}}\left[\sum_{\boldsymbol{x}\in Z^{(n,w_1)}}\sum_{\boldsymbol{y}\in Z^{(n,w_2)}} I[H\boldsymbol{x}^t = \boldsymbol{0}, H\boldsymbol{y}^t = \boldsymbol{0}]\right] \\ &= \sum_{\boldsymbol{x}\in Z^{(n,w_1)}}\sum_{\boldsymbol{y}\in Z^{(n,w_2)}} E_{\mathcal{G}}\left[I[H\boldsymbol{x}^t = \boldsymbol{0}, H\boldsymbol{y}^t = \boldsymbol{0}]\right] \\ &= \sum_{\boldsymbol{x}\in Z^{(n,w_1)}}\sum_{\boldsymbol{y}\in Z^{(n,w_2)}}\sum_{H\in\mathcal{G}} P(H)I[H\boldsymbol{x}^t = \boldsymbol{0}, H\boldsymbol{y}^t = \boldsymbol{0}]. \end{aligned} \quad (86)$$

For the case where $\mathcal{G} = \mathcal{R}_{m,n}$, we obtain

$$\begin{aligned} &E_{\mathcal{R}_{m,n}}[A_{w_1}A_{w_2}] \\ &= \sum_{\boldsymbol{x}\in Z^{(n,w_1)}}\sum_{\boldsymbol{y}\in Z^{(n,w_2)}} \frac{\#\{H : H\boldsymbol{x}^t = \boldsymbol{0}, H\boldsymbol{y}^t = \boldsymbol{0}\}}{2^{mn}}. \end{aligned} \quad (87)$$

We here encounter a problem of counting the matrices which satisfy both $H\boldsymbol{x}^t = \boldsymbol{0}$ and $H\boldsymbol{y}^t = \boldsymbol{0}$. In preparation to solve this counting problem, we will introduce some notation:

Definition 4: For a given pair $(\boldsymbol{x}, \boldsymbol{y}) \in Z^{(n,w_1)} \times Z^{(n,w_2)}$, the index sets I_1, I_2, I_3, I_4 are defined as follows:

$$I_1 \stackrel{\triangle}{=} \{k \in [1,n] : x_k = 1, y_k = 0\} \quad (88)$$

$$I_2 \stackrel{\triangle}{=} \{k \in [1,n] : x_k = 1, y_k = 1\} \quad (89)$$

$$I_3 \stackrel{\triangle}{=} \{k \in [1,n] : x_k = 0, y_k = 1\} \quad (90)$$

$$I_4 \stackrel{\triangle}{=} \{k \in [1,n] : x_k = 0, y_k = 0\}, \quad (91)$$

where $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$. These regions are illustrated in Fig.7. The size of each index set is denoted by $i_k = \#I_k (k = 1, 2, 3, 4)$. Let $\boldsymbol{h} = (h_1, h_2, \ldots, h_n)$ be a binary n-tuple. The partial weight of $\boldsymbol{h}$ corresponding to an index set $I_k (k = 1, 2, 3, 4)$ is denoted by $w_k(\boldsymbol{h})$, namely

$$w_k(\boldsymbol{h}) = \#\{j \in I_k : h_j = 1\}. \quad (92)$$

□

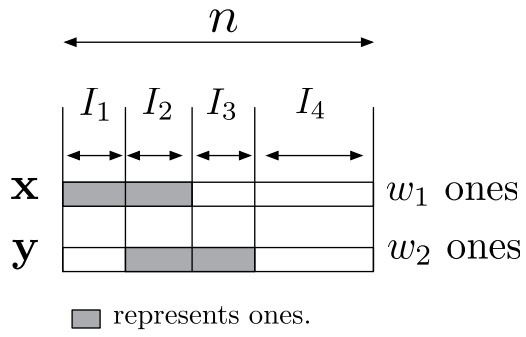


Fig. 7. The 4 regions I_1, I_2, I_3, I_4.

Since the index sets are mutually exclusive, the equation $i_1 + i_2 + i_3 + i_4 = n$ holds and i_2 can take the integer values in the following range:

$$\max\{w_1 + w_2 - n, 0\} \leq i_2 \leq \min\{w_1, w_2\}. \quad (93)$$

The size of each index set can be expressed as $i_1 = w_1 - i_2$, $i_3 = w_2 - i_2$, $i_4 = n - (w_1 + w_2 - i_2)$.

The next lemma forms the basis of the proof of Lemma 3.

Lemma 5: For any $\boldsymbol{x} \in Z^{(n,w_1)}$ and $\boldsymbol{y} \in Z^{(n,w_2)} (0 < w_1, w_2 \leq n)$, the following equalities hold:

$$\#\{\boldsymbol{h} \in F_2^n : \boldsymbol{h}\boldsymbol{x}^t = 0, \boldsymbol{h}\boldsymbol{y}^t = 0\} = \begin{cases} 2^{n-2} & \boldsymbol{x} \neq \boldsymbol{y} \\ 2^{n-1} & \boldsymbol{x} = \boldsymbol{y}. \end{cases} \quad (94)$$

(Proof) In the following, we are going to prove the claim of the lemma for the conditions $0 < w_1 \leq w_2 \leq n$. The proof for the final case $0 < w_2 \leq w_1 \leq n$ then follows immediately upon exchanging the variables w_2 and w_1 in the proof.

Firstly, we will show that

$$\#\{\boldsymbol{h} \in F_2^n : \boldsymbol{h}\boldsymbol{x}^t = 0, \boldsymbol{h}\boldsymbol{y}^t = 0\} = 2^{n-2} \quad (95)$$

if $0 < w_1 \leq w_2 \leq n$ and $\boldsymbol{x} \neq \boldsymbol{y}$. Let the support sets of $\boldsymbol{x}$ and $\boldsymbol{y}$ be $S(\boldsymbol{x}) \stackrel{\triangle}{=} \{i \in [1,n] : x_i = 1\}$ and $S(\boldsymbol{y}) \stackrel{\triangle}{=} \{i \in [1,n] : y_i = 1\}$, respectively. We need to consider the following three cases:

- Case (i): $0 < i_2 < w_1$ (i.e., $S(\boldsymbol{x})$ and $S(\boldsymbol{y})$ overlap but $S(\boldsymbol{y})$ does not include $S(\boldsymbol{x})$)
- Case (ii): $i_2 = 0$ (i.e., $S(\boldsymbol{x})$ and $S(\boldsymbol{y})$ do not overlap)
- Case (iii): $i_2 = w_1$(i.e., $S(\boldsymbol{y})$ includes $S(\boldsymbol{x})$)

These 3-cases are depicted in Fig.8.

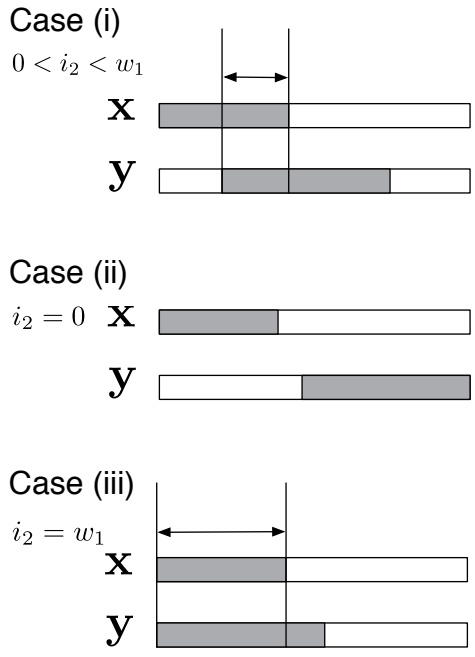


Fig. 8. The 3 cases.

Consider first Case (i). From the assumption $0 < i_2 < w_1$, it is evident that $I_1 \neq \emptyset$ (since $i_2 < w_1$), $I_2 \neq \emptyset$ (since $i_2 > 0$),

$I_3 \neq \emptyset$ (since $w_2 \geq w_1 > i_2$). For any $\boldsymbol{h} \in F_2^n$, the equations $\boldsymbol{h}\boldsymbol{x}^t = 0$ and $\boldsymbol{h}\boldsymbol{y}^t = 0$ hold if and only if

$$(w_1(\boldsymbol{h}) \text{ is even}), (w_2(\boldsymbol{h}) \text{ is even}) \text{ and } (w_3(\boldsymbol{h}) \text{ is even})$$

or

$$(w_1(\boldsymbol{h}) \text{ is odd}) , (w_2(\boldsymbol{h}) \text{ is odd}) \text{ and } (w_3(\boldsymbol{h}) \text{ is odd}).$$

Thus, the number of vectors satisfying the above condition is given by

$$\begin{aligned} N_{\boldsymbol{h}} &= 2 \times 2^{i_1-1} \times 2^{i_2-1} \times 2^{i_3-1} \times 2^{i_4} \\ &= 2 \times 2^{w_1-i_2-1} \times 2^{i_2-1} \\ &\times \ 2^{w_2-i_2-1} \times 2^{n-w_1-w_2+i_2} \\ &= 2^{n-2}, \end{aligned} \tag{96}$$

where $N_{\boldsymbol{h}}$ is defined by

$$N_{\boldsymbol{h}} \stackrel{\triangle}{=} \#\{\boldsymbol{h} \in F_2^n : \boldsymbol{h}\boldsymbol{x}^t = 0, \boldsymbol{h}\boldsymbol{y}^t = 0\}.$$

In the above derivation, we used the equalities $w_1 = i_1 + i_2, w_2 = i_2 + i_3, i_4 = n - (w_1 + w_2 - i_2)$.

Note that equation (96)(and equations (97)(98)(100) to be presented below) holds regardless of the size of I_4($i_4 = 0$ or $i_4 > 0$).

We now consider Case (ii). For this case, $I_1 \neq \emptyset$ (since $w_1 > 0$), $I_2 = \emptyset$(since $i_2 = 0$) and $I_3 \neq \emptyset$ (since $w_2 > 0$). The equalities $\boldsymbol{h}\boldsymbol{x}^t = 0$ and $\boldsymbol{h}\boldsymbol{y}^t = 0$ hold if and only if

$$(w_1(\boldsymbol{h}) \text{ is even}) \text{ and } (w_3(\boldsymbol{h}) \text{ is even}).$$

The number of vectors satisfying the condition is given by

$$\begin{aligned} N_{\boldsymbol{h}} &= 2^{i_1-1} \times 2^{i_3-1} \times 2^{i_4} \\ &= 2^{w_1-1} \times 2^{w_2-1} \times 2^{n-w_1-w_2} \\ &= 2^{n-2}. \end{aligned} \tag{97}$$

The final case is Case (iii). For this case, $I_1 = \emptyset$ (since $i_2 = w_1$), $I_2 \neq \emptyset$(since $i_2 = w_1 > 0$) and $I_3 \neq \emptyset$ (since $\boldsymbol{x} \neq \boldsymbol{y}$ and $w_1 \leq w_2$). These conditions lead to the following condition

$$(w_2(\boldsymbol{h}) \text{ is even}) \text{ and } (w_3(\boldsymbol{h}) \text{ is even})$$

for $\boldsymbol{h}\boldsymbol{x}^t = 0, \boldsymbol{h}\boldsymbol{y}^t = 0$. Again, 2^{n-2} n-tuples satisfy the above condition, namely

$$\begin{aligned} N_{\boldsymbol{h}} &= 2^{i_2-1} \times 2^{i_3-1} \times 2^{i_4} \\ &= 2^{i_2-1} \times 2^{w_2-i_2-1} \times 2^{n-w_2} \\ &= 2^{n-2}. \end{aligned} \tag{98}$$

Combining the above results for Cases (i)(ii)(iii), we obtain expression (95).

We then show that

$$N_{\boldsymbol{h}} = 2^{n-1} \tag{99}$$

holds if $0 < w_1 = w_2 \leq n$ and $\boldsymbol{x} = \boldsymbol{y}$. For this case, we have $I_1 = \emptyset, I_2 \neq \emptyset, I_3 = \emptyset$(since $\boldsymbol{x} = \boldsymbol{y}$). The equations $\boldsymbol{h}\boldsymbol{x}^t = 0, \boldsymbol{h}\boldsymbol{y}^t = 0$ hold if and only if

$$w_2(\boldsymbol{h}) \text{ is even.}$$

The number of n-tuples satisfying the above condition is given by

$$\begin{aligned} N_{\boldsymbol{h}} &= 2^{i_2-1} \times 2^{i_4} \\ &= 2^{i_2-1} \times 2^{n-i_2} \\ &= 2^{n-1}. \end{aligned} \tag{100}$$

The proof is completed. □

2) Proof of Lemma 3: The proof of Lemma 3 consists of two parts: The first part corresponds to the case where the covariance becomes zero. The second part corresponds to the case where the covariance becomes non-zero.

We commence with the first part of the proof: Assume that $0 < w_1, w_2 \leq n, \boldsymbol{x} \neq \boldsymbol{y}$. From Lemma 5 we obtain

$$\begin{aligned} &\#\{H : H\boldsymbol{x}^t = \boldsymbol{0}, H\boldsymbol{y}^t = \boldsymbol{0}\} \\ &\quad = \prod_{k=1}^{m} \#\{\boldsymbol{h} \in F_2^n : \boldsymbol{h}\boldsymbol{x}^t = 0, \boldsymbol{h}\boldsymbol{y}^t = 0\} \\ &\quad = \prod_{k=1}^{m} 2^{n-2} \\ &\quad = 2^{m(n-2)}. \end{aligned} \tag{101}$$

Substituting into (87) we obtain

$$\begin{aligned} &E_{\mathcal{R}_{m,n}}[A_{w_1}A_{w_2}] \\ &= \sum_{\boldsymbol{x} \in Z^{(n,w_1)}} \sum_{\boldsymbol{y} \in Z^{(n,w_2)}} \frac{\#\{H : H\boldsymbol{x}^t = \boldsymbol{0}, H\boldsymbol{y}^t = \boldsymbol{0}\}}{2^{mn}} \\ &= \sum_{\boldsymbol{x} \in Z^{(n,w_1)}} \sum_{\boldsymbol{y} \in Z^{(n,w_2)}} \frac{2^{m(n-2)}}{2^{mn}} \\ &= 2^{-2m} \sum_{\boldsymbol{x} \in Z^{(n,w_1)}} \sum_{\boldsymbol{y} \in Z^{(n,w_2)}} 1 \\ &= 2^{-2m} \binom{n}{w_1} \binom{n}{w_2} \\ &= E_{\mathcal{R}_{m,n}}[A_{w_1}] E_{\mathcal{R}_{m,n}}[A_{w_2}]. \end{aligned} \tag{102}$$

The last equality is equivalent to $\mathrm{Cov}_{\mathcal{R}_{m,n}}(A_{w_1}, A_{w_2}) = 0$.

We now consider the second part of the proof: Assume that $\boldsymbol{x} = \boldsymbol{y}$. From Lemma 5 we have

$$\#\{H : H\boldsymbol{x}^t = \boldsymbol{0}, H\boldsymbol{y}^t = \boldsymbol{0}\} = 2^{m(n-1)}, \tag{103}$$

and so

$$\begin{aligned} &E_{\mathcal{R}_{m,n}}[A_w^2] \\ &= \sum_{\boldsymbol{x} \in Z^{(n,w)}} \sum_{\boldsymbol{y} \in Z^{(n,w)}} \frac{\#\{H : H\boldsymbol{x}^t = \boldsymbol{0}, H\boldsymbol{y}^t = \boldsymbol{0}\}}{2^{mn}} \\ &= \sum_{\boldsymbol{x} \in Z^{(n,w)}} \sum_{\boldsymbol{y} \in Z^{(n,w)}} \frac{I[\boldsymbol{x} = \boldsymbol{y}] 2^{m(n-1)}}{2^{mn}} \\ &+ \sum_{\boldsymbol{x} \in Z^{(n,w)}} \sum_{\boldsymbol{y} \in Z^{(n,w)}} \frac{I[\boldsymbol{x} \neq \boldsymbol{y}] 2^{m(n-2)}}{2^{mn}} \\ &= 2^{-m} \sum_{\boldsymbol{x} \in Z^{(n,w)}} \sum_{\boldsymbol{y} \in Z^{(n,w)}} I[\boldsymbol{x} = \boldsymbol{y}] \\ &+ 2^{-2m} \sum_{\boldsymbol{x} \in Z^{(n,w)}} \sum_{\boldsymbol{y} \in Z^{(n,w)}} I[\boldsymbol{x} \neq \boldsymbol{y}] \end{aligned}$$

$$= 2^{-m}\binom{n}{w} + 2^{-2m}\left(\binom{n}{w}\binom{n}{w} - \binom{n}{w}\right)$$
$$= 2^{-2m}\binom{n}{w}\binom{n}{w} + 2^{-m}\binom{n}{w} - 2^{-2m}\binom{n}{w}$$
$$= E_{\mathcal{R}_{m,n}}[A_w]^2 + 2^{-m}\binom{n}{w} - 2^{-2m}\binom{n}{w}. \quad (104)$$

The last equality is equivalent to

$$\mathrm{Cov}_{\mathcal{R}_{m,n}}(A_w, A_w) = (1-2^{-m})2^{-m}\binom{n}{w}. \quad (105)$$

The proof is now completed. □

B. Proof of Lemma 4 (Covariance of sparse matrix ensemble)

Consider two binary n-tuples $\boldsymbol{x} \in Z^{(n,w_1)}$ and $\boldsymbol{y} \in Z^{(n,w_2)}$. As in the proof of Lemma 3, we need to consider 3-cases: Case (i) $0 < i_2 < w_1$, Case (ii) $i_2 = 0$, and Case (iii) $i_2 = w_1$.

We first study Case (i). Suppose that a binary n-tuple $\boldsymbol{h}$ is generated from a Bernoulli source with $Pr[h_i = 1] = p(i \in [1,n])$. Recall that p is defined by $p = k/n$. We denote the probability that $\boldsymbol{h}$ satisfies $\boldsymbol{h}\boldsymbol{x}^t = 0$ and$\boldsymbol{h}\boldsymbol{y}^t = 0$ under the condition $0 < i_2 < w_1$ by Q_1, that is

$$Q_1 \stackrel{\triangle}{=} Pr[\boldsymbol{h}\boldsymbol{x}^t = 0, \boldsymbol{h}\boldsymbol{y}^t = 0]. \quad (106)$$

As in the proof of Lemma 3, we need to consider the condition:

$(w_1(\boldsymbol{h})$ is even$)$, $(w_2(\boldsymbol{h})$ is even$)$ and $(w_3(\boldsymbol{h})$ is even$)$

or

$(w_1(\boldsymbol{h})$ is odd$)$, $(w_2(\boldsymbol{h})$ is odd$)$ and $(w_3(\boldsymbol{h})$ is odd$)$.

It is well known that a binary vector $(t_1, t_2, \ldots, t_u)$ generated from a Bernoulli source has even weight with probability $(1+(1-2q)^u)/2$, where q is the probability that $t_i (i \in [1,n])$ takes 1 [1]. The probability that $(t_1, t_2, \ldots, t_u)$ has an odd weight is given by $(1-(1-2q)^u)/2$. For example, the probability that $w_1(\boldsymbol{h})$ becomes even is $(1+x^{w_1})/2$. where $x = 1-2p$.

Based on the above argument, we can write the probability Q_1 as a function of x,

$$\begin{aligned} Q_1 &= \frac{(1+x^{i_1})(1+x^{i_2})(1+x^{i_3})}{8} \\ &+ \frac{(1-x^{i_1})(1-x^{i_2})(1-x^{i_3})}{8} \\ &= \frac{1+x^{i_1+i_2}+x^{i_2+i_3}+x^{i_1+i_3}}{4} \\ &= \frac{1+x^{w_1}+x^{w_2}+x^{w_1+w_2-2i_2}}{4} \\ &= \frac{1+x^{w_1}+x^{w_2}+x^{w_1+w_2}-x^{w_1+w_2}+x^{w_1+w_2-2i_2}}{4} \\ &= \left(\frac{1+x^{w_1}}{2}\right)\left(\frac{1+x^{w_2}}{2}\right) - \frac{x^{w_1+w_2}-x^{w_1+w_2-2i_2}}{4}. \end{aligned} \quad (107)$$

From a combinatorial argument, we can see that the number of pairs $(\boldsymbol{x}, \boldsymbol{y})$ satisfying $0 < i_2 < w_1$, which is denoted by $A_1(w_1, w_2)$, is given by

$$\begin{aligned} &A_1(w_1, w_2) \\ &\stackrel{\triangle}{=} \#\{(\boldsymbol{x}, \boldsymbol{y}) \in Z^{(n,w_1)} \times Z^{(n,w_2)} : 0 < i_2 < w_1\} \\ &= \sum_{j=1}^{w_1-1} \binom{n}{w_1}\binom{w_1}{j}\binom{n-w_1}{w_2-j} \end{aligned} \quad (108)$$

for $w_1 \le w_2$.

We next consider Case (ii). For this case, i_2 is assumed to be zero. The probability that $\boldsymbol{h}$ satisfies $\boldsymbol{h}\boldsymbol{x}^t = 0$ and $\boldsymbol{h}\boldsymbol{y}^t = 0$ under the condition $i_2 = 0$ is given by

$$\begin{aligned} Q_2 &= \left(\frac{1+x^{i_1}}{2}\right)\left(\frac{1+x^{i_3}}{2}\right) \\ &= \left(\frac{1+x^{w_1}}{2}\right)\left(\frac{1+x^{w_2}}{2}\right). \end{aligned} \quad (109)$$

The number of pairs $(\boldsymbol{x}, \boldsymbol{y})$ satisfying $i_2 = 0$ is given by

$$\begin{aligned} A_2(w_1, w_2) &\stackrel{\triangle}{=} \#\{(\boldsymbol{x}, \boldsymbol{y}) \in Z^{(n,w_1)} \times Z^{(n,w_2)} : i_2 = 0\} \\ &= \binom{n}{w_1}\binom{n-w_1}{w_2} \end{aligned} \quad (110)$$

for $w_1 \le w_2$.

Finally we consider Case (iii). We first consider the case $i_2 = w_1, \boldsymbol{x} \neq \boldsymbol{y}$ The probability that $\boldsymbol{h}$ satisfies $\boldsymbol{h}\boldsymbol{x}^t = 0, \boldsymbol{h}\boldsymbol{y}^t = 0$ under the condition $i_2 = w_1, \boldsymbol{x} \neq \boldsymbol{y}$ is

$$\begin{aligned} Q_3 &= \left(\frac{1+x^{i_2}}{2}\right)\left(\frac{1+x^{i_3}}{2}\right) \\ &= \left(\frac{1+x^{w_1}}{2}\right)\left(\frac{1+x^{w_2-w_1}}{2}\right) \\ &= \frac{1+x^{w_1}+x^{w_2-w_1}+x^{w_2}}{4} \\ &= \left(\frac{1+x^{w_1}}{2}\right)\left(\frac{1+x^{w_2}}{2}\right) - \frac{x^{w_1+w_2}-x^{w_2-w_1}}{4} \end{aligned} \quad (111)$$

We next consider the case $\boldsymbol{x} = \boldsymbol{y}$. For this case, we have

$$Q_3' = \frac{1+x^{w_1}}{2}. \quad (112)$$

In both cases, the number of pairs $(\boldsymbol{x}, \boldsymbol{y})$ satisfying $i_2 = w_1$ is given by

$$\begin{aligned} A_3(w_1, w_2) &\stackrel{\triangle}{=} \#\{(\boldsymbol{x}, \boldsymbol{y}) \in Z^{(n,w_1)} \times Z^{(n,w_2)} : i_2 = w_1\} \\ &= \binom{n}{w_1}\binom{n-w_1}{w_2-w_1}. \end{aligned} \quad (113)$$

We are now ready to derive the covariance of the weight distribution. Assume that $w_1 < w_2$. The second moment can be expressed as

$$\begin{aligned} &E_{\mathcal{T}_{m,n,k}}[A_{w_1}A_{w_2}] \\ &= \sum_{\boldsymbol{x}\in Z^{(n,w_1)}} \sum_{\boldsymbol{y}\in Z^{(n,w_2)}} Pr[H\boldsymbol{x}^t = \boldsymbol{0}, H\boldsymbol{y}^t = \boldsymbol{0}] \\ &= A_1(w_1, w_2)Q_1^m + A_2(w_1, w_2)Q_2^m \\ &+ A_3(w_1, w_2)Q_3^m. \end{aligned} \quad (114)$$

Substituting the Q_i and $A_i(w_1, w_2)(i = 1, 2, 3)$ obtained

above into equation 114, we immediately have

$$\begin{aligned}
&E_{\mathcal{T}_{m,n,k}}[A_{w_1}A_{w_2}] \\
&= \sum_{j=1}^{w_1-1}\binom{n}{w_1}\binom{w_1}{j}\binom{n-w_1}{w_2-j} \\
&\times \left(\left(\frac{1+x^{w_1}}{2}\right)\left(\frac{1+x^{w_2}}{2}\right)-\frac{x^{w_1+w_2}-x^{w_1+w_2-2j}}{4}\right)^m \\
&+ \binom{n}{w_1}\binom{n-w_1}{w_2}\left(\frac{1+x^{w_1}}{2}\right)^m\left(\frac{1+x^{w_2}}{2}\right)^m \\
&+ \binom{n}{w_1}\binom{n-w_1}{w_2-w_1} \\
&\times \left(\left(\frac{1+x^{w_1}}{2}\right)\left(\frac{1+x^{w_2}}{2}\right)-\frac{x^{w_1+w_2}-x^{w_2-w_1}}{4}\right)^m \\
&= \sum_{j=0}^{w_1}\binom{n}{w_1}\binom{w_1}{j}\binom{n-w_1}{w_2-j} \\
&\times \left(\left(\frac{1+x^{w_1}}{2}\right)\left(\frac{1+x^{w_2}}{2}\right)-\frac{x^{w_1+w_2}-x^{w_1+w_2-2j}}{4}\right)^m.
\end{aligned}$$

It is possible to retrieve $E_{\mathcal{T}_{m,n,k}}[A_{w_1}]E_{\mathcal{T}_{m,n,k}}[A_{w_2}]$ from the right-hand side of the above equation:

$$\begin{aligned}
&E_{\mathcal{T}_{m,n,k}}[A_{w_1}A_{w_2}] \\
&= \sum_{j=0}^{w_1}\binom{n}{w_1}\binom{w_1}{j}\binom{n-w_1}{w_2-j} \\
&\times \left(\frac{1+x^{w_1}}{2}\right)^m\left(\frac{1+x^{w_2}}{2}\right)^m\xi_{w_1,w_2,j}^m \\
&= \binom{n}{w_1}\binom{n}{w_2}\left(\frac{1+x^{w_1}}{2}\right)^m\left(\frac{1+x^{w_2}}{2}\right)^m \\
&+ \sum_{j=0}^{w_1}\binom{n}{w_1}\binom{w_1}{j}\binom{n-w_1}{w_2-j} \\
&\times \left(\frac{1+x^{w_1}}{2}\right)^m\left(\frac{1+x^{w_2}}{2}\right)^m\left(\xi_{w_1,w_2,j}^m-1\right) \\
&= E_{\mathcal{T}_{m,n,k}}[A_{w_1}]E_{\mathcal{T}_{m,n,k}}[A_{w_2}] \\
&+ \left(\frac{1+x^{w_1}}{2}\right)^m\left(\frac{1+x^{w_2}}{2}\right)^m \\
&\times \sum_{j=1}^{w_1}\binom{n}{w_1}\binom{w_1}{j}\binom{n-w_1}{w_2-j}\left(\xi_{w_1,w_2,j}^m-1\right). \quad (115)
\end{aligned}$$

In the last equality, the range of the summation on j was changed from $[0, w_1]$ to $[1, w_1]$. This is because $\xi_{w_1,w_2,j}^m - 1 = 0$ when $j = 0$, that is

$$\begin{aligned}
\xi_{w_1,w_2,0}^m - 1 &= \left(1-\frac{x^{w_1+w_2}-x^{w_1+w_2}}{(1+x^{w_1})(1+x^{w_2})}\right)^m - 1 \\
&= 0. \quad (116)
\end{aligned}$$

In the derivation of equation (115), the following identity was also used,

$$\sum_{j=0}^{w_1}\binom{n}{w_1}\binom{w_1}{j}\binom{n-w_1}{w_2-j} = \binom{n}{w_1}\binom{n}{w_2}. \quad (117)$$

From equation (115), we obtain the covariance

$$\mathrm{Cov}_{\mathcal{T}_{m,n,k}}(A_{w_1}, A_{w_2}) = \psi(w_1, w_2) \quad (118)$$

for $1 \le w_1 < w_2 \le n$. If $1 \le w_2 < w_1 \le n$ then

$$\mathrm{Cov}_{\mathcal{T}_{m,n,k}}(A_{w_1}, A_{w_2}) = \mathrm{Cov}_{\mathcal{T}_{m,n,k}}(A_{w_2}, A_{w_1}) = \psi(w_2, w_1). \quad (119)$$

Thus, it is reasonable to define $\psi(w_1, w_2) = \psi(w_2, w_1)$ if $1 \le w_2 < w_1 \le n$.

We now discuss the case $w = w_1 = w_2$. For this case, the second moment has the form

$$\begin{aligned}
&E_{\mathcal{T}_{m,n,k}}[A_w^2] \\
&= \sum_{\boldsymbol{x}\in Z^{(n,w)}}\sum_{\boldsymbol{y}\in Z^{(n,w)}} Pr[H\boldsymbol{x}^t = \mathbf{0}, H\boldsymbol{y}^t = \mathbf{0}] \\
&= A_1(w,w)Q_1^m + A_2(w,w)Q_2^m + A_3(w,w)Q_3'^m,
\end{aligned}$$

which can be written as

$$\begin{aligned}
&E_{\mathcal{T}_{m,n,k}}[A_w^2] \\
&= \sum_{j=1}^{w-1}\binom{n}{w}\binom{w}{j}\binom{n-w}{w-j} \\
&\times \left(\left(\frac{1+x^w}{2}\right)^2-\frac{x^{2w}-x^{2w-2j}}{4}\right)^m \\
&+ \binom{n}{w}\binom{n-w}{w}\left(\frac{1+x^w}{2}\right)^{2m} + \binom{n}{w}\left(\frac{1+x^w}{2}\right)^m \\
&= \sum_{j=0}^{w}\binom{n}{w}\binom{w}{j}\binom{n-w}{w-j} \\
&\times \left(\left(\frac{1+x^w}{2}\right)^2-\frac{x^{2w}-x^{2w-2j}}{4}\right)^m \\
&+ \binom{n}{w}\left(\frac{1+x^w}{2}\right)^m - \binom{n}{w}\left(\left(\frac{1+x^w}{2}\right)^2-\frac{x^{2w}-1}{4}\right)^m \\
&= E[A_w]^2 \\
&+ \left(\frac{1+x^w}{2}\right)^{2m}\sum_{j=1}^{w}\binom{n}{w}\binom{w}{j}\binom{n-w}{w-j}\left(\xi_{w,w,j}^m-1\right) \\
&+ \binom{n}{w}\left(\frac{1+x^w}{2}\right)^m - \binom{n}{w}\left(\frac{1+x^w}{2}\right)^m \\
&= E[A_w]^2 + \psi(w,w). \quad (120)
\end{aligned}$$

From the last equation, we obtain the variance

$$\mathrm{Cov}_{\mathcal{T}_{m,n,k}}(A_w, A_w) = \psi(w,w). \quad (121)$$

This completes the proof of Lemma 4. □

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/